# Variational Bayes Algorithms for mmWave and Massive MIMO-OFDM Systems with Low Resolution ADCs

A Thesis

Submitted for the Degree of

**Doctor of Philosophy**

in the **Faculty of Engineering**

by

**Sai Subramanyam Thoota**

under the Guidance of

**Prof. Chandra R. Murthy**


Electrical Communication Engineering
Indian Institute of Science
Bangalore – 560 012, INDIA

April 2022

TO

My Family

# Acknowledgments

I feel incredibly grateful to my advisor *Prof. Chandra R. Murthy*, for his tremendous support, patience, and advice throughout my time as a graduate student. His invaluable insights into our research work have molded my abilities to formulate and solve research problems and have taught me the art of writing papers and presenting the results. His pursuit of excellence and perfection has inspired me to be a better researcher and person.

I want to express my gratitude to *Dr. Ramesh Annavajjala* and *Dr. Prabhu Babu* for their invaluable technical discussions during my Ph.D.

My special thanks go to the members of the ECE office, especially *Mr. Srinivasa Murthy*. I also thank *Ms. Rajani* and *Ms. Suma* for their administrative support. I would also like to thank IISc for providing me an opportunity to pursue my doctoral studies along with good hostel and mess facilities. I am grateful to the doctors, nurses, and IISc health center staff, especially *Dr. R. Nirmala*, for her support during the Covid19 pandemic. I would also like to thank *Prof. Dipshika Chakravortty* for keeping the atmosphere light and cheerful during the pandemic. I appreciate the financial support from MHRD, Govt. of India, for pursuing my Ph.D. and conference travels.

A technically interactive and intellectually vibrant lab is crucial in the life of any Ph.D. scholar. I have been fortunate to have one during my Ph.D. I thank my lab members *Akshay, Ashok, Bala, Bhawesh, Chethan, Himani, Kanchan, Mohan, Mohit, Monika, Nandan, Prabhasa, Raksha, Ramu, Ribhu, Rupam, Sangeeta Bhattacharjee, Sangeeta Bind, Sanjhi, Shilpa, Shivani, Sireesha, Srinivas, Sunny, Swetha, Vaibhav, Varun, Venu, Yashvanth*. They have been a part of my journey in IISc, during different phases of my Ph.D. I would also like to thank *Arun, Chandrasekhar, Geethu, Lekshmi, Rubin,*

*Sameera, Saurabh, Unni* for their technical discussions in individual and group meetings. My special thanks to *Anubhab, Chirag, Dheeraj, Pradip*, and *Soumendu* for the technical interactions and weekend fun, which kept the mood upbeat during my Ph.D. I would also like to thank the Next Generation Wireless lab colleagues, *Ramakiran, Reneeta, Saikiran Bulusu, Sayan*, and *Vineeth*, for the fun trips and discussions.

I am indebted to my family, who have supported me throughout my Ph.D. I begin by thanking my parents, *Hemavathy* and *Vinayagam*, and my sisters, *Lalitha* and *Uma*, who have been supportive in all my endeavors. I would also like to thank my brothers-in-law, *Srihari* and *Murali*, for their support. I also thank my nieces *Kheerthi, Litika*, and nephew *Milan* for keeping me joyful and entertained during this journey of my Ph.D. I would also like to take this opportunity to thank *Ms. Sujatha* and *Mr. Madhavudu* for their invaluable help in my life. I also thank *Mr. Yathindra Babu* for his support before joining my Ph.D.

# Abstract

In this thesis, we develop novel low-complexity algorithms for massive multiple-input multiple-output (MIMO) systems under practical non-idealities and theoretically analyze their performance. The first problem we consider is that of joint channel estimation and data decoding in uplink massive multiple-input-multiple-output (MIMO) systems with low-resolution analog-to-digital converters (ADCs) at the base station. The nonlinearities introduced by the ADCs make the problem challenging: in particular, the existing linear detectors perform poorly. Also, the channel coding used in commercial wireless systems necessitates soft symbol detection to obtain satisfactory performance. In this part of the thesis, we present a low-complexity variational Bayesian (VB) inference procedure to jointly solve the (possibly correlated) channel estimation and soft symbol decoding problem. We present the approach in progressively more complex scenarios, including the case where even the channel statistics are not available at the receiver. Then, we combine the VB procedure with a belief propagation (BP) based channel decoder, which further enhances the performance without any additional complexity. We numerically evaluate the bit error rate (BER) and the normalized mean squared error (NMSE) in the channel estimates obtained by our algorithm as a function of various system parameters, and benchmark the performance against genie-aided and state-of-the-art receivers. The results show that the VB procedure is a promising approach for developing low-complexity advanced receivers in low-resolution ADC based systems.

In the second problem, we consider the delay-domain sparse channel estimation and data decoding problems in a massive MIMO orthogonal frequency division multiplexing (MIMO-OFDM) wireless communication system with low-resolution ADCs. The high

non-linear distortion due to coarse quantization leads to severe performance degradation in conventional OFDM receivers, which necessitates novel receiver techniques. Firstly, we derive the Bayesian Cramér-Rao lower bound (CRLB) on the mean squared error (MSE) in recovering jointly compressible vectors from quantized noisy underdetermined measurements. Secondly, we formulate the pilot-assisted channel estimation as a multiple measurement vector (MMV) sparse recovery problem, and develop a VB algorithm to infer the posterior distribution of the channel. We benchmark the MSE performance of our algorithm with that of the CRLB, and numerically show that the VB algorithm meets the CRLB. Thirdly, we present a soft symbol decoding algorithm that infers the posterior distributions of the data symbols given the quantized observations. We utilize the posterior statistics of the detected data symbols as virtual pilots, and develop an iterative soft symbol decoding and data-aided channel estimation procedure. Finally, we present a variant of the iterative algorithm that utilizes the output bit log-likelihood-ratios (LLRs) of the channel decoder to adapt the data prior to further improve the performance. We provide interesting insights into the impact of the various system parameters on the MSE and BER of the developed algorithms, and benchmark them against the state-of-the-art.

In the third problem, we present a novel model-and-data-driven channel estimation procedure in a millimeter-wave MIMO-OFDM wireless communication system. The transceivers employ a hybrid analog-digital architecture. We adapt techniques from a wide range of signal processing methods, such as compressed sensing and Bayesian inference, to learn the unknown sparsifying dictionary in the beamspace domain, as well as the delay-and-beamspace sparse channel. We train the model-based algorithm with a site-specific training dataset generated using a realistic ray tracing-based wireless channel simulation tool. We assess the performance of the developed channel estimation algorithm with the same site's test data. We benchmark the performance of our procedure in terms of NMSE error against an existing fast greedy method and two state-of-the-art algorithms, and empirically show that model-based approaches combined with data-driven customization outperform purely model based techniques by a large margin. This algorithm was

selected as one of the top three solutions in the "ML5G-PHY Channel Estimation Global Challenge 2020" organized by the International Telecommunication Union.

In the last problem considered in this thesis, we study the problem of downlink (DL) sum rate maximization in codebook-based multiuser (MU) MIMO systems. The user equipments (UEs) estimate the DL channels using pilot symbols sent by the access point (AP) and feedback the estimates to the AP over a control channel. We present a closed form expression for the achievable sum rate of the MU-MIMO broadcast system with codebook constrained precoding based on the estimated channels, where multiple data streams are simultaneously transmitted to all users. Next, we present novel, computationally efficient, minorization-maximization (MM) based algorithms to determine the selection of beamforming vectors and power allocation to each beam that maximizes the achievable sum rate. Our solution involves multiple uses of MM in a nested fashion. Based on this approach, we present and contrast two algorithms, which we call the square-root-MM (SMM) and inverse-MM (IMM) algorithms. The algorithms are iterative and converge to a locally optimal beamforming vector selection and power allocation solution from any initialization. We evaluate the performance and complexity of the algorithms for various values of the system parameters, compare them with existing solutions, and provide further insights into how they can be used in system design.

# Contents

# List of Figures

# List of Tables

# Glossary

| | |
|---|---|
| 3GPP | Third Generation Partnership Project |
| AMP | Approximate Message Passing |
| AWGN | Additive White Gaussian Noise |
| BiGAMP | Bilinear Generalized Approximate Message Passing |
| BER | Bit Error Rate |
| BPSK | Binary Phase Shift Keying |
| CP | Cyclic Prefix |
| CRLB | Cramér Rao Lower Bound |
| CS | Compressed Sensing |
| EM | Expectation Maximization |
| FIM | Fisher Information Matrix |
| IMM | Inverse Minorization Maximization |
| LLR | Log-Likelihood Ratio |
| LTE-A | Long Term Evolution - Advanced |
| MAP | Maximum a Posteriori |
| MIMO | Multiple Input Multiple Output |
| MISO | Multiple Input Single Output |
| ML | Maximum Likelihood |
| MLGS | Multi-level Greedy Search |
| MM | Minorization Maximization |
| MMSE | Minimum Mean Squared Error |

| | |
|---|---|
| MMV | Multiple Measurement Vector |
| MSE | Mean Squared Error |
| NMSE | Normalized Mean Squared Error |
| OFDM | Orthogonal Frequency Division Multiplexing |
| OMP | Orthogonal Matching Pursuit |
| PC-OMP | Projection Cuts Orthogonal Matching Pursuit |
| PCSBL | Pattern Coupled Sparse Bayesian Learning |
| QAM | Quadrature Amplitude Modulation |
| QPSK | Quadrature Phase Shift Keying |
| RZF | Regularized Zero Forcing |
| SBL | Sparse Bayesian Learning |
| SMM | Square-root Minorization Maximization |
| SMV | Single Measurement Vector |
| SNR | Signal-to-Noise Ratio |
| SWOMP | Simultaneously Weighted Orthogonal Matching Pursuit |
| VB | Variational Bayes |
| WMMSE | Weighted Minimum Mean Squared Error |
| ZF | Zero Forcing |

# Notation

**Sets**

| | |
|---|---|
| $\mathbb{R}$ | The set of real numbers |
| $\mathbb{C}$ | The set of complex numbers |
| $[n]$ | The set $\{1, \ldots, n\}$ |

**Vectors and matrices**

| | |
|---|---|
| $\mathbf{I}_n$ | Identity matrix of size $n \times n$ |
| $\mathbf{A}^T$ | Transpose of a matrix $\mathbf{A}$ |
| $\bar{\mathbf{A}}$ | Complex conjugate of a matrix $\mathbf{A}$ |
| $\mathbf{A}^H$ or $\mathbf{A}^*$ | Conjugate transpose of a matrix $\mathbf{A}$ |
| $\lvert\mathbf{A}\rvert$ | Determinant of a matrix $\mathbf{A}$ |
| $\mathbf{A} \succcurlyeq \mathbf{B}$ | $\mathbf{A} - \mathbf{B}$ is positive semi-definite, where $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices |
| $\lVert\mathbf{A}\rVert_F$ | Frobenius norm of a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ |
| $\mathrm{tr}(\mathbf{A})$ | Trace of a matrix $\mathbf{A}$ |
| $\mathbf{A}^\dagger$ | Moore-Penrose pseudoinverse of a matrix $\mathbf{A}$ |
| $\mathtt{vec}(A)$ | Vectorized version of a matrix $\mathbf{A}$ |
| $\mathtt{diag}\,(\mathbf{a})$ | A diagonal matrix with the entries of vector $\mathbf{a}$ on the diagonal |
| $\mathtt{diag}\,(\mathbf{A})$ | A vector with its entries as the diagonal elements of a matrix $\mathbf{A}$ |
| $\mathtt{blkdiag}(\mathbf{A})$ | Block diagonal part of a matrix $\mathbf{A}$ |
| $\mathbf{A} \otimes \mathbf{B}$ | Kronecker product of matrices $\mathbf{A}$ and $\mathbf{B}$ |
| $\lVert\mathbf{x}\rVert_2$ | $\ell_2$ norm of a vector $\mathbf{x} \in \mathbb{C}^n$ |

**Random Variables and Events**

| | |
|---|---|
| i.i.d. | Independent and identically distributed |
| $P(\mathcal{E})$ or $p(\mathcal{E})$ | Probability of an event $\mathcal{E}$ |
| $\mathbb{E}[X]$ | Expectation of a random variable $X$ |
| $KL(\mathrm{P}\|\mathrm{Q})$ | Kullback-Leibler divergence between distributions P and Q |
| $X \sim \mathcal{CN}(\mu, \sigma^2)$ | $X$ is a complex Gaussian random variable with mean $\mu$ and variance $\sigma^2$ |

**Order notation**

| | |
|---|---|
| $f(n) = O(g(n))$ | $\exists c > 0$ such that $f(n) < cg(n)$ for all sufficiently large $n$ |

**Miscellaneous**

| | |
|---|---|
| $\Gamma(u)$ | Gamma function at $u > 0$, i.e. $\int_0^\infty x^{u-1} e^{-x} \mathtt{d}x$ |

# Chapter 1

# Introduction

Massive multiple-input-multiple-output (MIMO) antenna systems is one of the most spectrally efficient wireless communication technologies devised to date, and plays an integral role in the 5G wireless standard [2–6]. In massive MIMO, a base station (BS) is equipped with tens or hundreds of antennas and serves large number of single antenna users in the same time-frequency resource. It provides high beamforming and spatial multiplexing gains due to the large number of antennas used. When combined with millimeter-wave (mmWave) communications, which enables bandwidths of the order of GigaHertz (GHz), it can potentially offer a 1000-fold improvement in the system throughput [7, 8]. Further, massive MIMO assists in combating the large attenuation caused by mmWave frequencies, thereby improving the coverage area. Due to these advantages, both massive MIMO and mmWave technologies are expected to play a vital role in the sixth generation wireless technology as well [9–13]. Fig. 1.1 shows a block diagram of a typical uplink (UL) massive MIMO system where multiple users transmit data to the BS or access point (AP) in the same time-frequency resource.

Despite the advantages mentioned above, massive MIMO systems also pose new challenges in practical implementation, which need to be addressed in order to make them commercially viable. First, the massive number of antennas results in huge power consumption at the BS, which could make the system energy-inefficient. Second, in a massive

MIMO system, each antenna is connected to a radio frequency (RF) chain, which consists of components such as low noise amplifiers, analog filters, automatic gain control, etc. Each RF chain is connected to two ADCs, one each for the in-phase and quadrature components, whose power consumption increases exponentially with the bit-widths. For example, the power consumption of a commercial 20 Giga-samples per second, 8-12 bits ADC is around 500 mW. Therefore, for a massive MIMO system with 100 RF chains and 200 ADCs, the total power consumption will be as high as 100 W, which is unaffordable in a practical massive MIMO system [8].

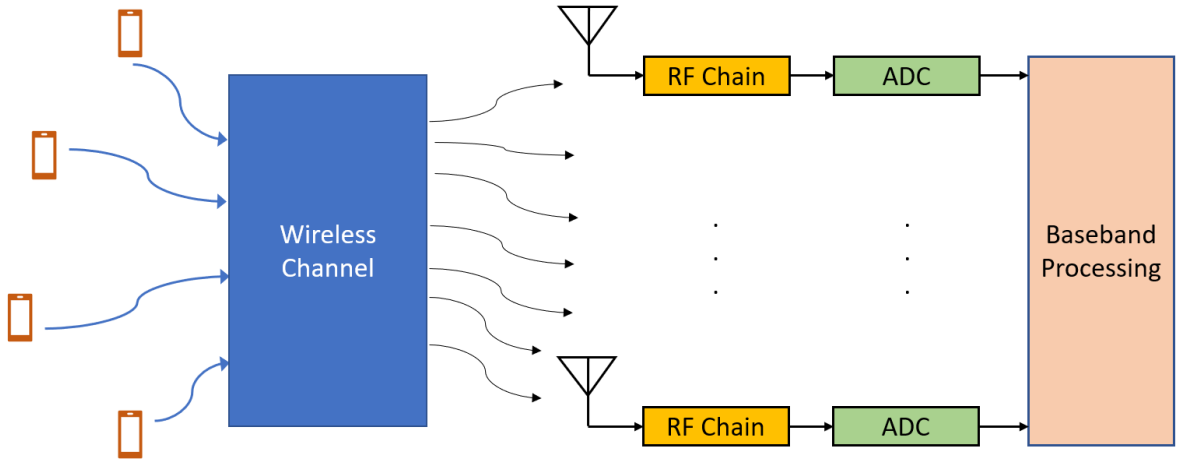


Figure 1.1: Block diagram of a massive MIMO system.

## Low-resolution ADC based Massive MIMO Systems

A promising solution to the high power consumption of massive MIMO systems is to use low-resolution ADCs at the BS [8]. Coarse quantization also enables the use of less expensive components in the RF chain, which further contributes to cost savings. However, a major demerit of using low-resolution ADCs is the non-linear distortion they introduce, which results in suboptimal performance of conventional receiver algorithms. For instance, channel estimation and data detection algorithms such as zero-forcing (ZF), minimum mean squared error (MMSE), etc. are designed for full precision inputs, so

applying them directly for coarsely quantized inputs leads to severe performance degradation [14, 15]. There are several approaches to handle the nonlinearities introduced by low-resolution quantization; we list two of them here. First, we can model the nonlinearity as an additive noise term using the Bussgang theorem [16]. That is, the output after the nonlinear distortion introduced by the ADC is equal to a scaled version of the input plus an uncorrelated distortion term [17]. Second, we can utilize the information of the specific quantizer used, and adopt a principled approach to develop receive-processing algorithms that explicitly account for the nonlinear distortion introduced by the ADCs. This second approach is the primary focus of this thesis.

In addition to the power and cost advantages due to the low-resolution ADCs, the existing literature on their capacity analysis in the context of MIMO systems provide further motivation to adopt them in future wireless technologies [18–23]. In [18, 19], it is shown that, under the assumption of perfect channel state information (CSI) at the receiver, at low signal-to-noise ratio (SNR) and up to a first order approximation in SNR, the mutual information of a MIMO system employing 1-bit ADCs is $2/\pi$ times that of a MIMO system without quantization. Also, quardrature phase shift keying (QPSK) is the optimal modulation scheme in terms of the achievable rate at low SNRs, and under i.i.d. Rayleigh fading, the ergodic capacity of a 1-bit quantized system increases linearly with the number of receive antennas. Reference [20] obtains the number of mass points in an input distribution to achieve the capacity of a discrete input Gaussian channel under an average power constraint for a particular choice of the quantizer. It is shown that a quantizer with $M$ bins needs no more than $M+1$ mass points to achieve the capacity. Also, reference [21] derives the capacity of real valued Gaussian channels using the cutting plane algorithm in multi-bit quantized systems. They infer that low-resolution ADCs incur a small loss in spectral efficiency compared to infinite precision ADC based systems. For instance, a 2 to 3 bits quantization results in at most 20% reduction in spectral efficiency at high SNRs. Accurate expressions for the channel capacity at high and infinite SNR with 1-bit ADCs in frequency flat MIMO and SIMO channels have been obtained in [22, 23]. These studies also develop a convex optimization based computationally efficient

algorithm to design the input alphabet that achieves the infinite SNR capacity.

With this motivation, we next discuss the architectural changes in the air interface that may be needed for designing efficient receiver algorithms for low-resolution ADC based massive MIMO systems. In wireless technologies such as long term evolution (LTE), LTE-advanced (LTE-A) or 5G-NR, pilot symbols are embedded within a resource block in specific subcarriers and OFDM symbols, along with the data symbols [24, 25]. The number of pilot symbols depend on various factors such as the number of layers, antennas, etc. These pilot symbols are used for channel estimation at those subcarriers, and interpolation is done in time and frequency to estimate the channels across all the other OFDM symbols and subcarriers. In full resolution ADC based MIMO-OFDM systems, after RF downconversion, we quantize the inphase and quadrature components of the complex baseband signal, and separate the subcarriers using a discrete Fourier transform (DFT) operation. This generates orthogonal frequency-flat fading channels, which enables the use of linear channel estimators. The inter-carrier interference (ICI) due to RF impairments, local oscillator mismatches, etc are usually corrected using frequency correction loops, which run once in many sub-frames. However, in low-resolution ADC based systems, the coarse quantization of the time-domain signal results in ICI in every sub-frame and OFDM symbol, which cannot be compensated using the frequency or gain correction loops. Moreover, the random nature of the input signal to the quantizers makes it difficult to exactly characterize the ICI. Due to this, one cannot separate the pilot and data subcarriers using a simple DFT operation. Hence, conventional frequency domain channel estimation and interpolation fail to perform well in low-resolution ADC based MIMO systems.

### Data Symbols as Virtual Training Symbols

To overcome the effect of ICI mentioned above, an alternate approach is to employ separate pilot and data OFDM symbols within a coherence interval of the channel. The channel estimated using the pilot symbols are then used for data detection. Here, the quality of channel estimates depends on the number of pilot OFDM symbols used. One

of the major drawbacks of using low-resolution ADC based architecture is the increase in training overhead in both massive MIMO and MIMO-OFDM systems. This reduces the data transmission duration, in turn adversely impacting the spectral efficiency of the system. Therefore, we need novel solutions to overcome the challenge of channel estimation and data detection in low-resolution ADC based systems.

One promising approach to reduce the training overhead is to exploit the unknown data symbols themselves as virtual pilot symbols. In this, we first obtain initial channel estimates using a reduced number of pilot symbols, use them to detect the data, and then reuse them as virtual pilots to refine the channel estimates. An obvious drawback of using the detected data as virtual pilots is the error propagation when there are data detection errors. In the approach explored in this thesis, the detected data symbols are not directly used for channel estimation. Instead, we account for the probability that the data symbols take on the different possible values, and use these probabilities in computing the channel estimates. To elaborate, instead of using the hard data decisions, we use the bit LLRs or the posterior beliefs of the data symbols obtained from an initial channel estimate for refining the channel estimates in subsequent iterations. Moreover, the channel decoder that follows the data detector offers better performance when their inputs are soft decisions or bit LLRs, rather than hard-detected data bits. In this context, developing computationally and analytically tractable solutions to estimate the bit LLRs is important in massive MIMO and MIMO-OFDM systems to reduce the training overhead and improve the spectral efficiency.

## Sparsity of Wireless Channels in OFDM Systems

Another approach to reduce the training overhead that has received significant research attention is to utilize any inherent structure in the wireless channels. This side-information about the channels can be used to reduce the number of measurements or training symbols needed for channel estimation and/or improve the estimation accuracy. Channel sounding experiments have shown that wireless channels exhibit sparsity in either delay/lag domain, beamspace domain or both [26–28]. Exploiting sparsity enables us to use

compressed sensing based approaches to estimate a high dimensional sparse signal from an underdetermined system of noisy linear or non-linear equations [29, 30]. Although this is an NP-hard problem in general, under certain conditions on the measurement matrix (e.g., when it satisfies the restricted isometry property), convex relaxation based methods, greedy methods, or Bayesian inference can be used to obtain polynomial-time algorithms [31]. Extensive research has been carried out in recovering sparse signals from lower dimensional compressive measurements. Some of the key references on compressed sensing and sparse signal recovery are [32–44].

In this thesis, we utilize the delay domain as well as beamspace domain sparsity of the channel in the sub-6 GHz and mmWave frequencies, respectively, and design high performing polynomial time algorithms for massive MIMO-OFDM channel estimation with both infinite resolution and low-resolution ADCs. We also derive a Bayesian Cramér-Rao lower bound on the MSE performance of an estimator of sparse signals from quantized compressive measurements. This serves as a benchmark to evaluate the performance of the novel Bayesian sparse channel estimation algorithm developed in this thesis.

## Millimeter-Wave MIMO-OFDM Systems

In the previous subsections, we discussed the use of low-resolution ADCs to address the high power consumption and cost issues in massive MIMO and MIMO-OFDM systems. We now discuss another promising technology for future generations of wireless communication systems. Due to the large and ever-increasing demand for various communication services, there is a tremendous need for high bandwidth and data rates. This motivates the usage of mmWave frequencies, from 30 to 300 GHz, where the use of bandwidths exceeding 2 GHz is feasible. The mmWave communications are used in a wide variety of applications such as wireless local and personal area networks, automotive radar, vehicular communications, wearables, internet of things, etc., where data rate and stringent latency requirements need to be met [45–48].

Millimeter-wave channel models are different from the sub-6 GHz channels due to the signal propagation characteristics at smaller wavelengths. For example, diffraction effects

are lower due to the reduced Fresnel zone, scattering effects are higher and depend on type of the surface of different materials, penetration and propagation losses are higher, blockage effects (e.g., due to the human body) are higher, etc. [49]. Also, the antenna aperture and size decreases with the center frequency, which in turn limits the amount of signal energy a single antenna can capture. These factors necessitate the use of large antenna arrays in mmWave communications in order to achieve reliable communication at high data rates. Also, the presence of line-of-sight, non-line-of-sight, and clustered paths due to highly directional antennas at mmWave frequencies lead to structure in the channels such as sparsity in the lag/delay or beamspace domains. Considering all the aspects, several research papers have discussed the design of novel signal processing algorithms to improve the data rate and spectral efficiency of a mmWave communication system equipped with large antenna arrays [50–54]. Moreover, new hardware constraints arise due to the power consumption at mmWave frequencies. A popular approach to address the power consumption in large antenna array mmWave communication system is to employ the so-called *hybrid* analog-digital beamforming architecture. Here, we partition the antennas into sub-arrays and connect only one RF chain to each sub-array, thereby reducing the number of RF chains that need to be deployed. In a hybrid beamforming based transceiver, the signal at each transmit/receive antenna is phase shifted and connected to an RF chain [55, 56]. There are several architectures in a hybrid beamforming system, such as fully connected or partially connected networks, which provide different tradeoffs between the complexity and performance. Each of these architectures come with their own design challenges, some of which are discussed below.

One of the most important engineering problems to solve in a hybrid mmWave MIMO systems is the design of digital and analog precoders/combiners for data transmission and reception. To do that, we need reliable channel state information at the transmitter and receiver. So, designing novel channel estimation algorithms is a crucial task to be done to obtain high data rate and spectral efficiency in a mmWave MIMO communication system. In this thesis, we consider a site-specific channel estimation algorithm, where the channel statistics such as the angles of arrival and departure distributions, delay domain

and beamspace domain sparsity can be potentially learnt from a training dataset. This information can be used to estimate the channel for a test dataset. We numerically show that utilizing the information learned from a training dataset leads to improvement in the NMSE performance of our channel estimator compared to state-of-the-art [57].

In the next subsection, we discuss the problems addressed in this thesis. More detailed mathematical discussion of each problem and its solution is provided in the corresponding chapter.

## 1.1 Problems Addressed in this Thesis

In this thesis, we primarily focus on channel estimation and data decoding problems when low-resolution ADCs are employed at the BS of massive MIMO and MIMO-OFDM wireless communication systems [58–62]. We also consider mmWave MIMO-OFDM channel estimation when we have prior information of the angles-of-departure (AoD), angles-of-arrival (AoA), and multipath intensity profiles (MIPs) of multiple channel instantiations in a specific target environment [63]. The overall theme is to develop Bayesian techniques to solve the channel estimation and data decoding problems in both full resolution and low-resolution ADC based MIMO and MIMO-OFDM systems. The novelty in our solutions lies in the formulation of statistical inference problems to obtain the posterior distributions of the channel and data symbols as non-convex optimization problems. We then derive a concave surrogate function to bound the non-convex objective function from below at every point of the feasible set and is tight at the current iterate. Furthermore, the surrogate function is relatively easy to optimize. We maximize this surrogate function to find a local optimum, which becomes the new iterate.[1] We again find a new surrogate function which is tight at the new iterate, and repeat this process iteratively till a suitable convergence criterion is satisfied. This approach falls into the category of minorization-maximization (MM) (or majorization-minimization in the case of non-convex minimization problems), which provides a general and flexible framework to solve

[1] We seek a convex upper bound surrogate function in the case of minimization problems.

non-convex optimization problems [64, 65].

Another important problem in multiuser MIMO communications, apart from channel estimation and data decoding, is the precoding/beamforming in UL or downlink (DL) to maximize a metric such as sum rate or minimum rate under a set of given constraints. Along with the development of MM based algorithms for channel estimation and data detection, we also explore the application of this framework for beamforming problems as well. In this direction, we consider a single cell DL multiuser MIMO system with finite rate downlink control channels, and formulate a codebook based sum rate maximization problem [66, 67]. The finite rate downlink control channel limits the feedback of channel state information to a finite number of bits. So, once the AP optimizes the sum-rate, it feeds back only the beamforming vector indices to the users to set their precoding matrices for data reception. In this context, we focus on the codebook based precoding for sum rate maximization in a multiuser MIMO broadcast system to select the beamforming vectors and allocate powers to each of them.

We provide an overview of the problems addressed in this thesis below.

### 1.1.1 Joint Channel Estimation and Soft Symbol Decoding in Massive MIMO Single Carrier Systems with Low-Resolution ADCs

In this part of the thesis, we consider an UL massive MIMO single carrier (MIMO-SC) system equipped with low-resolution ADCs at the BS. Multiple single transmit antenna user equipments (UE) transmit their pilot symbols followed by data symbols within a channel coherence interval. The UEs encode their raw information bits using a channel code, modulate them to data symbols, upconvert to RF, and transmit them over flat fading wireless channels to the BS. Upon receiving the signal, the BS downconverts the signal to baseband, quantize the inphase and quadrature (IQ) components using low-resolution ADCs to obtain the complex baseband signal. Our aim is to estimate the UEs' channel using the received signal during the pilot phase, and use them to decode the information bits using the received signal in the data phase. Conventional receivers

in high resolution ADC based systems use linear processing using algorithms such as ZF, regularized zero-forcing (RZF), MMSE to detect the data [1, 15]. However, in low-resolution ADC based systems, these conventional receivers perform poorly due to the severe non-linearities introduced by the coarse quantization [68].

Mathematically, we represent the quantized received signal at the $N_r$ antennas during the pilot and data phases as

$$\begin{aligned} \mathbf{Y}_p &= \mathcal{Q}\left(\mathbf{H}\mathbf{X}_p + \mathbf{W}_p\right) \in \mathbb{C}^{N_r \times \tau_p}, \\ \mathbf{Y}_d &= \mathcal{Q}\left(\mathbf{H}\mathbf{X}_d + \mathbf{W}_d\right) \in \mathbb{C}^{N_r \times \tau_d}, \end{aligned} \tag{1.1}$$

respectively, where $\tau_p$ and $\tau_d$ are the pilot and data durations, respectively, and $\mathbf{X}_p \in \mathbb{C}^{K \times \tau_p}$ and $\mathbf{X}_d \in \mathbb{C}^{K \times \tau_d}$ are the pilot and data symbols, respectively. Here, $\mathbf{W}_p$ and $\mathbf{W}_d$ denote the complex additive white Gaussian noise (AWGN) during the pilot and data phases, respectively. The quantization operation denoted by $\mathcal{Q}(\cdot)$ is applied element-wise on the real and imaginary parts of its argument separately. Our goal is to estimate $\mathbf{H} \in \mathbb{C}^{N_r \times K}$ and decode $\mathbf{X}_d \in \mathbb{C}^{K \times \tau_d}$ using $\mathbf{Y}_p$ and $\mathbf{Y}_d$. More specifically, in coded communications, we need to provide the bit LLRs as input to the channel decoder. Therefore, our goal is not only to perform channel estimation and data detection, but also to obtain high-quality estimates of the bit LLRs to be input to the channel decoder, given the quantized received pilot and data signals.

Typically, in low-resolution analog-to-digital converter (ADC) based multiuser MIMO communication systems, we need to transmit a large number of pilot symbols (large $\tau_p$), to obtain reliable channel estimates [69, 70]. This overhead severely impacts the spectral efficiency of the system. One way to address this issue is to increase the resolution of the ADCs, but this again leads to increase in power consumption and hardware cost. An alternative approach is to utilize the decoded data symbols to re-estimate the channels in an iterative fashion. Further, we can jointly estimate both the channel and data symbols simultaneously by including both as unknown random parameters, and derive an algorithm to estimate them. In a Bayesian formulation, we include the channel and data symbols as latent variables, and infer their posterior distributions, given the quantized received

pilot and data signals $\mathbf{Y}_p$ and $\mathbf{Y}_d$. Due to the massive number of antennas employed at the BS, it is computationally intractable to find the exact posterior distributions of the channels and the data symbols. Therefore, our goal is to develop novel signal processing techniques of polynomial complexity for channel estimation and soft symbol decoding.

### 1.1.2 Sparse Channel Estimation and Soft Symbol Decoding in Massive MIMO-OFDM Systems with Low-Resolution ADCs

In the first problem discussed in subsection 1.1.1, we considered a massive MIMO-SC system in which the pilot and data signals undergo flat fading. However, in many commercial wireless technologies such as WiFi and LTE/LTE-A, multicarrier communication is used to improve the spectral efficiency and network capacity. Therefore, it is important to extend the idea of employing low-resolution ADCs to multicarrier MIMO systems also. In wideband orthogonal frequency division multiplexing (OFDM) communication systems, the transmit signals undergo frequency selective fading, which results in intersymbol interference (ISI). In conventional receivers operating in lag/delay domain, we use equalization algorithms such as maximum likelihood sequence estimation (MLSE), decision feedback equalization (DFE) etc., to remove the effects of the frequency selective channel and decode the data [71]. These are computationally expensive procedures which need dedicated hardware to meet the timing requirements. In full precision ADC based OFDM receivers, we adopt frequency domain equalization to convert the frequency selective channel to multiple orthogonal flat fading channels, and use single tap linear equalizers such as least-squares (LS) or MMSE for data detection.

In this part of the thesis, we consider a UL multiuser massive MIMO-OFDM communication system in which each user embeds the pilot and modulated data symbols onto multiple subcarriers, converts the frequency domain signal to time domain using an inverse DFT operation, adds cyclic prefix (CP), upconverts to RF, and transmits to the BS. At the BS, the pilot and data signal is downconverted and quantized using low-resolution

ADCs to obtain the complex baseband received signal. In full precision ADC based systems, we remove the CP, transform the time domain signal to frequency domain using a DFT operation. This decouples the subcarriers, resulting in multiple orthogonal flat fading channels, which simplifies both the channel estimation and data decoding problems. Mathematically, when $K$ single antenna users transmit separate pilot and data OFDM symbols, the unquantized received signals at the $n_r^{\text{th}}$ BS antenna during the $t^{\text{th}}$ pilot and data OFDM symbols are represented as

$$\mathbf{z}_{n_r}^{(p)}[t] = \sum_{k=1}^{K} \mathbf{F}_{N_c}^H \mathbf{X}_k^{(p)}[t] \mathbf{F}_{N_c,L} \mathbf{h}_{n_r,k} + \mathbf{w}_{n_r}^{(p)}[t], \tag{1.2}$$

$$\mathbf{z}_{n_r}^{(d)}[t] = \sum_{k=1}^{K} \mathbf{F}_{N_c}^H \mathbf{H}_{n_r,k}^{\text{freq}} \mathbf{x}_k^{(d)}[t] + \mathbf{w}_{n_r}^{(d)}[t], \tag{1.3}$$

respectively, where $N_r$ is the number of receive antennas, $N_c$ is the number of subcarriers, $L$ is the maximum delay spread in symbol intervals of all the users, $\mathbf{F}_{N_c} \in \mathbb{C}^{N_c \times N_c}$ is the DFT matrix, $\mathbf{F}_{N_c,L} \in \mathbb{C}^{N_c \times L}$ is the $L$ column truncated DFT matrix, $\mathbf{X}_k^{(p)}[t] \in \mathbb{C}^{N_c \times N_c}$ is a diagonal matrix containing the pilot symbols of the $k^{\text{th}}$ user transmitted during the $t^{\text{th}}$ pilot OFDM symbol, $\mathbf{x}_k^{(d)}[t] \in \mathbb{C}^{N_c \times 1}$ is the data transmitted by the $k^{\text{th}}$ user during the $t^{\text{th}}$ data OFDM symbol duration, $\mathbf{h}_{n_r,k} \in \mathbb{C}^{L \times 1}$ is the delay/lag domain channel between the $k^{\text{th}}$ user and the $n_r^{\text{th}}$ BS antenna, $\mathbf{H}_{n_r,k}^{\text{freq}} \in \mathbb{C}^{N_c \times N_c}$ is a diagonal matrix containing the DFT of $\mathbf{h}_{n_r,k}$. Also, $\mathbf{w}_{n_r}^{(p)}$ and $\mathbf{w}_{n_r}^{(d)}$ denote the complex AWGN during the pilot and data phases, respectively.

In full resolution ADC based receivers, we left multiply (1.2) and (1.3) by $\mathbf{F}_{N_c}$ to decouple the subcarriers, and obtain $N_c$ orthogonal flat fading MIMO multiple access channels (MAC). We can potentially use any linear receiver to estimate the channels and detect the data. However, in low-resolution ADC based systems, we quantize (1.2) and (1.3) to obtain the quantized complex baseband pilot and data received signals as

$$\mathbf{y}_{n_r}^{(p)}[t] = \mathcal{Q}\left(\mathbf{z}_{n_r}^{(p)}[t]\right), \tag{1.4}$$

$$\mathbf{y}_{n_r}^{(d)}[t] = \mathcal{Q}\left(\mathbf{z}_{n_r}^{(d)}[t]\right), \tag{1.5}$$

where $\mathcal{Q}$ denotes an element-wise quantization operation of the real and imaginary parts of its argument. If we apply DFT to (1.4) and (1.5), it results in severe ICI, which leads to suboptimal performance of linear receivers. This necessitates novel signal processing procedures for channel estimation and data detection. Similar to the discussion in subsection 1.1.1 in the context of massive MIMO-SC systems, our goal is to obtain the soft outputs or bit LLRs of the information bits here also. We solve the channel estimation and soft symbol decoding in this part of the thesis.

As discussed in the subsection 1.1.1, another crucial aspect to address is the large training overhead associated in low-resolution ADC based massive MIMO-OFDM systems, to obtain reliable channel estimates. One way to address this issue in frequency selective channels is to utilize the inherent delay (or lag) domain or angle domain sparsity in the wireless channels [27, 45, 72]. In our system model, the number of measurements available for estimating the users' channels is directly proportional to the number of pilot OFDM symbols. Thus, by reducing the number of measurements, we reduce the number of pilot symbols needed, but this leads to large channel estimation errors if conventional linear channel estimation schemes are used. Accounting for the sparsity of the channels in the channel estimation procedure enables us to use the theory of compressed sensing and sparse signal recovery [29] to obtain excellent performance with reduced training overhead. Moreover, across the receive antennas, the time-domain channels are independent but have the same support, which leads to a joint row sparse structure of the estimand. We use this structure to improve the support recovery performance of the sparsity-based channel estimation algorithm.

Along with deriving a solution for MIMO-OFDM channel estimation algorithm, we also analyze the NMSE performance of our algorithm, by deriving a Bayesian CRLB on the NMSE of an estimator of quantized MMV sparse signals [61].

### 1.1.3 Compressive Channel Estimation in Millimeter-Wave MIMO-OFDM Systems

In this part of the thesis, we consider a hybrid MIMO-OFDM system with a multi antenna transmitter and receiver communicating in mmWave frequencies ranging from 30 GHz to 300 GHz. At these frequencies, the large path loss due to factors such as blockages and oxygen absorption makes the received signal strength very low. One way to address this issue is to employ beamforming, which focuses the energy towards the intended receiver. To do this, we need reliable channel estimates, which are used to design the transmit precoders or beamformers. In mmWave transceivers, a large number of antennas are packed within a small area to obtain the array gain required to close the communication link. The high cost associated with employing a separate RF chain per antenna (i.e., a fully digital architecture) in this large MIMO scenario motivates one to adopt an alternate hybrid (digital+analog or fully analog) architecture. This results in a lower dimensional complex baseband signal at the receiver, which in turn necessitates the development of novel signal processing techniques to obtain high quality channel estimates, mainly because the channel is observable only through the lens of the analog beams used for signal transmission and reception [45].

Most of the current solutions for wireless communication problems such as channel estimation, beamforming and data detection follow a model driven approach, where a mathematical system model is used to develop an algorithm that optimizes a particular metric of interest. With the availability of large datasets, it is interesting to explore the option of using a data-driven approach to further improve the performance of a model driven algorithm. The primary benefit of such an approach is that it can potentially address model mismatches and result in more robust, better performing systems in real-world situations. In this direction, we consider a site-specific mmWave channel estimation problem, where a BS transmits pilot symbols to a user equipment (UE) via a mmWave channel. The channel is generated using a wireless channel generation tool called "Raymobtime", which uses raytracing and 3D scenarios to simulate a realistic mmWave wireless environment [73]. As the channels are generated in a specific environment, their statistics

are typically invariant within the measurement interval. So, if we can use the available information based on a training dataset, we can potentially design a channel estimation algorithm which outperforms a purely model driven algorithm. The solution to this problem was presented as part of an "**ML5G-PHY Channel Estimation Global Challenge 2020**" organized by the International Telecommunication Union (ITU).[2] Our algorithm was selected as one of the top three solutions in this challenge [63].

We present the receive signal model to introduce the problem here, and discuss more details in the corresponding chapter. In the mmWave OFDM receiver, after RF combining, down-conversion, cyclic/zero prefix removal and DFT, the complex baseband signal received during the $m^{\text{th}}$ training frame for the $k^{\text{th}}$ subcarrier, denoted by $\mathbf{y}^{(m)}[k] \in \mathbb{C}^{L_r \times 1}$, can be mathematically expressed as

$$\mathbf{y}^{(m)}[k] = \mathbf{W}_{\text{tr}}^{(m)^*}\big(\mathbf{H}[k]\mathbf{F}_{\text{tr}}^{(m)}\mathbf{q}^{(m)}t^{(m)}[k] + \mathbf{n}^{(m)}[k]\big), \quad m = 1, \ldots, M, \quad k = 1, \ldots, K, \tag{1.6}$$

where $L_t$ and $L_r$ are the number of RF chains at the transmitter and receiver, respectively, $K$ is the number of subcarriers, $M$ is the number of training frames, $\mathbf{W}_{\text{tr}}^{(m)}$ is the RF combining matrix during the $m^{\text{th}}$ training frame, $\mathbf{H}[k] \in \mathbb{C}^{N_r \times N_t}$ represents the frequency domain MIMO channel matrix for the $k^{\text{th}}$ subcarrier, $N_t$ and $N_r$ are the number of transmit and receive antennas, respectively. We choose the $m^{\text{th}}$ training signal as $\mathbf{s}^{(m)}[k] = \mathbf{q}^{(m)}t^{(m)}[k]$, where $t^{(m)}[k] \in \mathbb{C}$ is a subcarrier-dependent pilot symbol, and $\mathbf{q}^{(m)} \in \mathbb{C}^{L_t \times 1}$ is a frequency-flat vector whose entries are chosen as $\frac{1}{\sqrt{2L_t}}(a + jb)$, where $a, b \in \{-1, 1\}$ and are uniformly distributed. The noise vector $\mathbf{n}^{(m)}[k]$ is independently and identically distributed across $K$ subcarriers as $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_r})$. After compensating for $t^{(m)}[k]$ and vectorizing (1.6), we obtain

$$\mathsf{vec}(\mathbf{y}^{(m)}[k]) = \underbrace{\left(\mathbf{q}^{(m)\,T}\mathbf{F}_{\text{tr}}^{(m)\,T} \otimes \mathbf{W}_{\text{tr}}^{(m)^*}\right)}_{\mathbf{\Phi}^{(m)}} \mathsf{vec}(\mathbf{H}[k]) + \mathbf{W}_{\text{tr}}^{(m)^*}\mathbf{n}^{(m)}[k]. \tag{1.7}$$

Here, $\mathbf{H}[k]$ is the frequency domain representation of the frequency selective channel whose

[2] https://www.itu.int/en/ITU-T/AI/challenge/2020/Pages/default.aspx

$d^{\text{th}}$ delay domain tap is denoted by $\mathbf{H}_d$, and $N_c$ is the length of the channel. We adopt the extended virtual channel model in [74] to represent $\mathbf{H}[k]$ as

$$\mathbf{H}[k] \approx \tilde{\mathbf{A}}_{\text{R}} \boldsymbol{\Delta}^{v}[k] \tilde{\mathbf{A}}_{\text{T}}^{*}, \qquad k = 1, \ldots, K, \tag{1.8}$$

where $\boldsymbol{\Delta}^{v}[k]$ contains the complex channel gains, and the dictionary matrices $\tilde{\mathbf{A}}_{\text{R}} \in \mathbb{C}^{N_r \times G_r}$ and $\tilde{\mathbf{A}}_{\text{T}} \in \mathbb{C}^{N_t \times G_t}$ contain the Rx and Tx array steering vectors evaluated on a grid of size $G_r$ for the AoA and a grid of size $G_t$ for the AoD, respectively. With appropriate choices of $G_r$ and $G_t$, $\boldsymbol{\Delta}^{v}[k] \in \mathbb{C}^{G_r \times G_t}$ becomes a sparse matrix containing the channel path gains on the locations that match with the actual Angle Of Departures (AoDs) and Angle Of Arrivals (AoAs). By concatenating the RF combined signals of $M$ training frames, we get the receive equation for the $k^{\text{th}}$ subcarrier as

$$\underbrace{\begin{bmatrix} \mathbf{y}^{(1)}[k] \\ \vdots \\ \mathbf{y}^{(M)}[k] \end{bmatrix}}_{\mathbf{y}[k]} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi}^{(1)} \\ \vdots \\ \boldsymbol{\Phi}^{(M)} \end{bmatrix}}_{\boldsymbol{\Phi}} \boldsymbol{\Psi} \mathbf{h}^{\text{v}}[k] + \underbrace{\begin{bmatrix} \mathbf{n}_c^{(1)}[k] \\ \vdots \\ \mathbf{n}_c^{(M)}[k] \end{bmatrix}}_{\mathbf{n}_c[k]}, \tag{1.9}$$

where $\boldsymbol{\Phi}$ contains the pilot symbols, $\boldsymbol{\Psi}$ is the sparsifying dictionary, and $\mathbf{h}^{\text{v}}[k]$ is the sparse vector to be estimated. Now, by stacking the received signals of $K$ subcarriers, we get the final system equation

$$\begin{aligned} \mathbf{Y} &= \begin{bmatrix} \mathbf{y}[1] & \ldots & \mathbf{y}[K] \end{bmatrix} \\ &= \boldsymbol{\Phi}\boldsymbol{\Psi} \begin{bmatrix} \mathbf{h}^{\text{v}}[1] & \ldots & \mathbf{h}^{\text{v}}[K] \end{bmatrix} + \begin{bmatrix} \mathbf{n}_c[1] & \ldots & \mathbf{n}_c[K] \end{bmatrix} \\ &= \boldsymbol{\Phi}\boldsymbol{\Psi}\mathbf{H}^{\text{v}} + \mathbf{N}_c. \end{aligned} \tag{1.10}$$

Our goal is to learn the sparsifying dictionary $\boldsymbol{\Psi}$ and any other learnable parameters from the training dataset, and use them to estimate the channel $\mathbf{H}[k]$, $k = 1, \ldots, K$. We provide complete details of all the variables in (1.10) in the corresponding chapter.

### 1.1.4 Sum-Rate Maximization in Codebook based Multi-user MIMO Systems

In the last part of this thesis, we consider a single cell DL frequency division duplexing (FDD) multiuser MIMO system where the AP or BS equipped with $N_t$ transmit antennas communicates with $K$ user equipments with $N_r$ receive antennas each, via a MIMO broadcast channel. In the DL training phase, the AP transmits $\tau_p$ orthogonal pilot signals to the users for channel estimation. The received pilot sequence at the $k^{\text{th}}$ user, $\mathbf{Y}_k^{(p)} \in \mathbb{C}^{N_r \times \tau_p}$, is given by

$$\mathbf{Y}_k^{(p)} = \sqrt{\rho_{dl}^{(p)} \tau_p} \mathbf{H}_k \mathbf{X}_p + \mathbf{W}_k, \tag{1.11}$$

where $\rho_{dl}^{(p)}$ is the pilot signal to noise ratio (SNR), and $\mathbf{W}_k \in \mathbb{C}^{N_r \times \tau_p}$ is the complex AWGN whose columns are i.i.d. with mean $\mathbf{0}$ and covariance matrix $\mathbf{I}_{N_r}$, denoted $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$. The users estimate the CSI and feed them back to the AP via error free links in the UL. Upon receiving the CSI, the AP computes the precoding matrices to be used for DL data transmission.

In our problem, we consider a finite rate downlink control channel where the AP can transmit only limited information about the precoding matrices to the users [75]. Therefore, we limit the choice of beamforming vectors to a codebook, and choose them from this set to maximize a suitable performance metric. The AP then transmits only the precoding vector indices to the users to be used for their data reception [76]. We set the sum rate achieved by the system as our performance metric. One potential approach to maximize the sum rate is for the AP to compute the precoding matrices assuming an infinite precision feedback link, and then quantize them to the nearest beamforming vectors using any chosen distance metric. But this may result in suboptimal system performance. To alleviate this issue, we need to include the precoding codebook right from the problem formulation phase, and design algorithms to select the beamforming vectors from the codebook.

The received signal $\mathbf{y}_k \in \mathbb{C}^{N_r}$ at the $k^{\text{th}}$ user is given by

$$\mathbf{y}_k = \mathbf{H}_k \left( \sqrt{\rho_{dl}} \sum_{j=1}^{K} \mathbf{C}\mathbf{\Phi}_j^{\frac{1}{2}} \mathbf{s}_j \right) + \mathbf{w}_k, \tag{1.12}$$

where $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is the channel between the AP and $k^{\text{th}}$ user, $\mathbf{w}_k \in \mathbb{C}^{N_r}$ is the complex AWGN at the $k^{\text{th}}$ user with distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$. In (1.12), $\mathbf{C}$ is the codebook of beamforming vectors, which is known both to the users and AP. Details of the notation used will be explained in the corresponding chapter of this thesis.

Once the user estimates the channel $\widehat{\mathbf{H}}_k$, the received signal can be rewritten as

$$\mathbf{y}_k = \underbrace{\sqrt{\rho_{dl}} \widehat{\mathbf{H}}_k \mathbf{C}\mathbf{\Phi}_k^{\frac{1}{2}} \mathbf{s}_k}_{\text{Desired signal}} + \sqrt{\rho_{dl}} \widehat{\mathbf{H}}_k \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{C}\mathbf{\Phi}_j^{\frac{1}{2}} \mathbf{s}_j + \sqrt{\rho_{dl}} \widetilde{\mathbf{H}}_k \sum_{j=1}^{K} \mathbf{C}\mathbf{\Phi}_j^{\frac{1}{2}} \mathbf{s}_j + \mathbf{w}_k. \tag{1.13}$$

Our initial goal in this problem is to derive an expression for the achievable rate of each user. Subsequently, we use the derived achievable rate expressions as a performance metric to design algorithms to select the beamforming vectors to transmit the users' data in the downlink, and allocate powers to them, given a transmit power budget. Denoting the covariance matrix of the noise and interference of the $k^{\text{th}}$ user by $\mathbf{V}_k$ (which we will derive later), we write the achievable rate of the $k^{\text{th}}$ user as

$$R_k = \log \left| \mathbf{I}_{N_r} + \mathbf{V}_k^{-1} \widehat{\widehat{\mathbf{H}}}_k \mathbf{\Phi}_k \widehat{\widehat{\mathbf{H}}}_k^H \right|, \tag{1.14}$$

where $\widehat{\widehat{\mathbf{H}}}_k \triangleq \sqrt{\rho_{dl}} \widehat{\mathbf{H}}_k \mathbf{C}$. Now, we need to maximize the sum rate $R_{\text{tot}} = \sum_{k=1}^{K} R_k$ under a total power constraint. The optimization problem we solve is given by

$$\begin{aligned} \underset{\substack{\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_K \\ \mathbf{\Phi}_k \text{ diagonal, p.s.d.}}}{\text{maximize}} & \sum_{k=1}^{K} \log \left| \mathbf{I}_{N_r} + \mathbf{V}_k^{-1} \widehat{\widehat{\mathbf{H}}}_k \mathbf{\Phi}_k \widehat{\widehat{\mathbf{H}}}_k^H \right|, \\ \text{subject to } & \text{tr}\Big( \sum_{k=1}^{K} \mathbf{\Phi}_k \Big) = 1. \end{aligned} \tag{1.15}$$

where $\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_K$ are the transmit covariance matrices of the users need to be optimized. The interference and noise covariance matrices $\{\mathbf{V}_1, \ldots, \mathbf{V}_K\}$ in the objective function tie the power allocation of all the users together, which makes the problem difficult to solve. Moreover, the beamforming vector selection to maximize the sum rate makes it also a combinatorial problem, which complicates it further. Therefore, our subsequent goal in this problem is to design polynomial time optimization algorithms to solve (1.15), to obtain a locally optimal solution for the beamforming vector selection and power allocation problem.

## 1.2 Contributions and Outline of the Thesis

In this section, we provide a chapter-wise outline of the thesis. We divide the thesis into three parts. The first part deals with the channel estimation and data decoding problems in massive MIMO and MIMO-OFDM communication systems with low-resolution ADCs. The second part consists of the site-specific mmWave channel estimation problem, whereas the last part contains the codebook based precoding problem in multiuser MIMO systems. We provide a flow diagram of the contents of this thesis in Fig. 1.2.

1. In the second chapter, we solve the channel estimation and data decoding problems in the context of a massive MIMO-SC system discussed in the section 1.1.1. The contents of this chapter is published in [58–60]. We list our contributions in this chapter below:

   (a) We derive the posterior distributions of the channel and the data symbols obtained from the received pilot and data observations. Due to the analytical intractability caused because of the high dimensional nature of the problem, we resort to approximate inference methods to compute them. More specifically, we adopt a framework called variational Bayes (VB), which provides a principled approach to solve statistical inference problems.

   (b) We consider both the spatially correlated and uncorrelated Rayleigh fading channels in our problem setup. In the spatially correlated case, we exploit

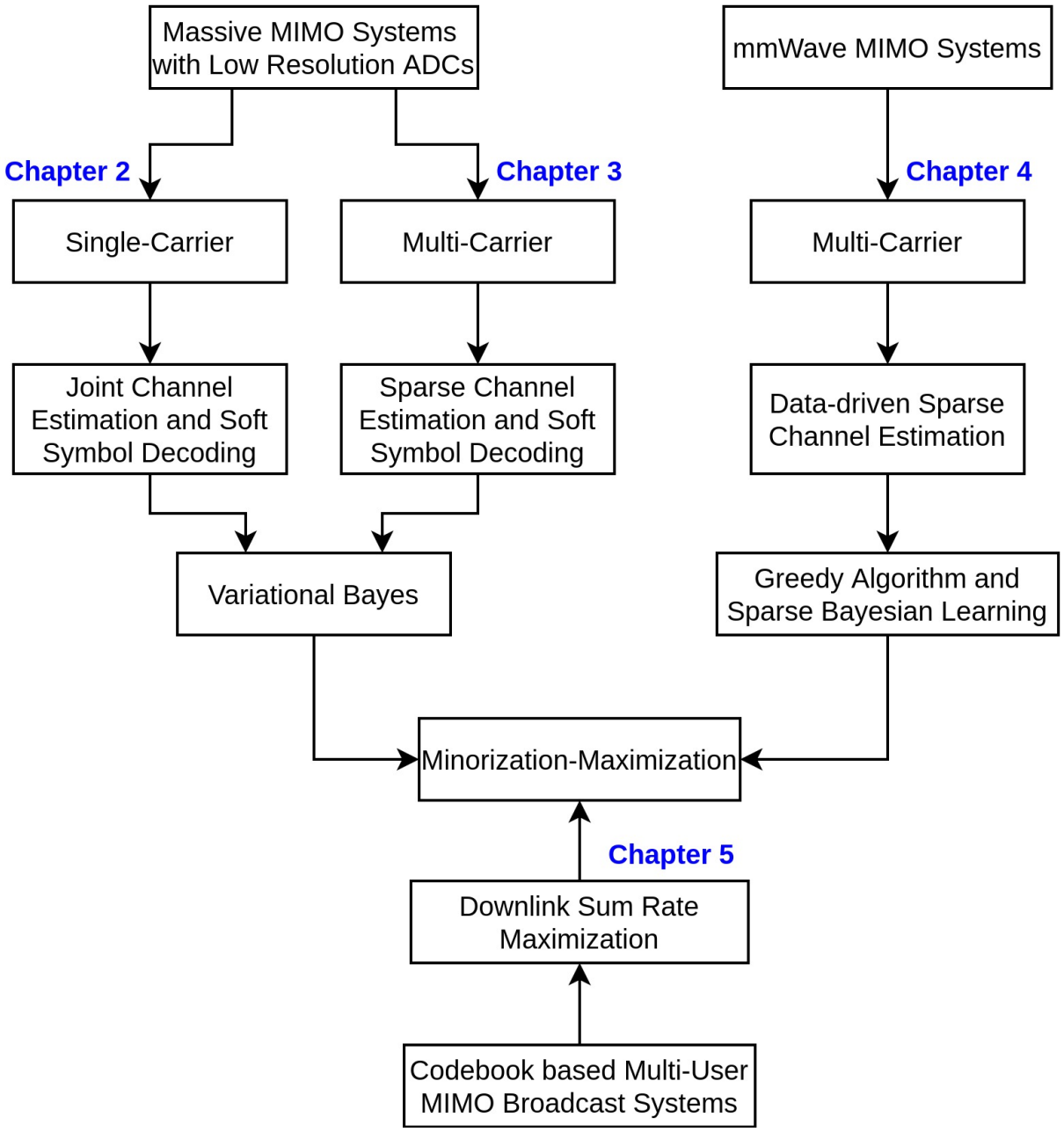


Figure 1.2: Organization and Chapter-wise Outline of the Thesis.

the channel correlation to improve the channel estimation and data detection performance. On the other hand, in the spatially uncorrelated scenario, we do not assume any knowledge of the channel statistics while deriving our solution. We also learn the channel statistics in the spatially uncorrelated case, which can be potentially used for link adaptation and downlink precoding.

(c) We combine the channel estimator and data detector with a channel decoder to improve the system performance further. We study the bit error rate (BER) and channel NMSE performance with respect to the various system parameters, and provide insights into system design. We evaluate the performance as a function of channel parameters, and benchmark against the state-of-the-art.

2. In the third chapter, we solve the channel estimation and data decoding problems in a massive MIMO-OFDM system with low-resolution ADCs discussed in the section 1.1.2. The contents of this chapter have been published in [61, 62]. We list our contributions in this chapter below:

   (a) We derive the Bayesian CRLB for the MSE incurred by an estimator for recovering jointly compressible vectors from quantized compressed sensing measurements. Specifically, we impose a hierarchical circularly symmetric complex Gaussian prior on the estimand, parameterized by a diagonal precision matrix. The precision matrix is in turn hyper-parameterized by a Gamma distribution. Although the CRLB is not available in closed-form, it can be evaluated by numerical methods.

   (b) We consider both deterministic and random cases for the precision matrix to obtain two different CRLBs on the MSE of jointly compressible vectors. Also, in the case of exactly-sparse signals, we derive a support-aware CRLB, which assumes the knowledge of the support set, to compute the bound.

   (c) We exploit the lag/delay domain sparsity of the channels to formulate channel estimation in a massive MIMO-OFDM system as a quantized MMV sparse signal recovery problem. We develop a VB algorithm to infer the posterior distributions of the channels. We benchmark the MSE performance of the VB channel estimator with the derived Bayesian CRLB, and empirically show that the estimator meets the CRLB.

   (d) We then develop a quantized VB soft symbol decoding algorithm that uses the estimated channels to obtain the posterior beliefs of the data symbols. We use

these posterior statistics to generate virtual pilots, and present a data-aided channel estimation procedure to refine the initial channel estimates. Based on this, we develop an iterative algorithm that alternately runs the soft symbol decoder and data-aided channel estimator steps.

(e) We also develop a variant of the iterative channel estimation and data decoding algorithm, which utilizes the aposteriori bit LLRs output from the channel decoder to adapt the prior used by the data detector. The resulting combined channel estimator, data detector and channel decoder further improves the system performance.

3. In the fourth chapter, we consider the site-specific mmWave channel estimation problem discussed in the section 1.1.3. The contents of this chapter has been published in [63]. We list our contributions in this chapter below:

   (a) We formulate the mmWave channel estimation problem in a MIMO-OFDM setting as a MMV sparse signal recovery problem. We integrate a fast greedy search with a high-performing Bayesian inference method in the first approach.

   (b) We present a multi-level greedy search (MLGS) procedure to learn the sparsifying virtual beamspace dictionary which reduces the dimensionality of the problem. We then use the learned dictionary to estimate the channel using a reduced dimensional sparse Bayesian learning (SBL) method.

   (c) We finally exploit the delay-domain sparsity learnt from the training dataset to de-noise the estimated channels. We compare the NMSE of our algorithm with the state-of-the-art simultaneously weighted orthogonal matching pursuit (OMP), pattern-coupled sparse Bayesian learning for channel estimation with dominating delay taps (PCSBL-DDT) and projection cuts OMP (PC-OMP).

4. In the fifth chapter, we solve the sum-rate maximization problem discussed in section 1.1.4. The contents of this chapter have been published in [66,67]. The contributions in this chapter are as follows:

(a) We derive the achievable sum rate of a codebook based precoding MU-MIMO broadcast system with minimum mean squared error (MMSE) channel estimation at the receiver and imperfect CSI at the transmitter. The achievable sum rate expression provides us with the objective function for beamforming vector selection and power allocation.

(b) We present two iterative algorithms for solving the non-convex and combinatorial sum rate maximization problem using MM procedure, which provide analytical expressions for the surrogate optimization problems. We show that the closed-form solutions are optimal with respect to their corresponding surrogate optimization problems.

(c) We empirically study the performance of the algorithms with respect to various system parameters. Further, we illustrate the performance advantage offered by the algorithms compared to the state-of-the-art approaches as well as a single-user-optimal codebook based precoding approach. The results demonstrate that jointly choosing beamforming vectors is necessary to realize the full potential of MU-MIMO transmission.

## 1.3 List of Publications

### 1.3.1 Journal Papers

1. S. S. Thoota, P. Babu, and C. R. Murthy, "Codebook-based precoding and power allocation for MU-MIMO systems for sum rate maximization," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8290–8302, Dec. 2019.

2. S. S. Thoota and C. R. Murthy, "Variational Bayes' joint channel estimation and soft symbol decoding for uplink massive MIMO systems with low resolution ADCs," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3467–3481, May 2021.

3. S. S. Thoota, D. G. Marti, O. T. Demir, R. Mundlamuri, J. Palacios, C. M. Yetis, C. K. Thomas, S. H. Bharadwaja, E. Björnson, P. Giselsson, M. Kountouris, C. R.

Murthy, N. González-Prelcic, J. Widmer, "Site-specific millimeter-wave compressive channel estimation algorithms with hybrid MIMO architectures," *ITU Journal on Future and Evolving Technologies (ITU J-FET) special issue on AI and machine learning solutions in 5G and future networks*, Volume 2 (2021), Issue 4, Jul. 2021.

4. S. S. Thoota and C. R. Murthy, "Massive MIMO-OFDM systems with low resolution ADCs: Cramér-Rao bound, sparse channel estimation, and soft symbol decoding," *IEEE Trans. Signal Process.*, 2022.

### 1.3.2 Conference Papers

1. S. S. Thoota, P. Babu, and C. R. Murthy, "Codebook based precoding for multiuser MIMO broadcast systems: An MM approach," in *Proc. Nat. Conf. Commun. (NCC)*, Feb. 2019, pp. 1–6.

2. S. S. Thoota, C. R. Murthy, and R. Annavajjala, "Quantized variational Bayesian joint channel estimation and data detection for uplink massive MIMO systems with low resolution ADCS," *Proc. IEEE 29th Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Oct. 2019, pp. 1–6.

3. S. S. Thoota and C. R. Murthy, "Variational Bayesian inference based soft-symbol decoding for uplink massive MIMO systems with low resolution ADCs," *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, Nov. 2019, pp. 2180–2184.

4. S. S. Thoota and C. R. Murthy, "Cramér-Rao lower bound for Bayesian estimation of quantized MMV sparse signals," *Proc. Signal Process. Adv. in Wireless Comm. (SPAWC)*, 2021.

# Part 1

# Chapter 2

# Variational Bayes' Joint Channel Estimation and Soft Symbol Decoding for Uplink Massive MIMO Systems with Low Resolution ADCs

## 2.1 Introduction

Massive multiple input multiple output (MIMO) wireless communication systems, where the base station (BS) or access point (AP) is equipped with hundreds or thousands of antennas, and simultaneously serves tens or hundreds of users, is one of the key enabling technologies to meet the increasing demand for the data rate and energy efficiency [77,78]. However, the advantages of massive MIMO come at the cost of increased power consumption and hardware complexity due to the large number of RF chains, high precision analog-to-digital converters (ADCs), etc. In particular, the power consumption of ADCs grow exponentially with the number of quantization bits per sample [69, 79, 80]. For example, a commercial 1 Gsamp/s 12-bit ADC from Texas Instruments consumes over 1 W of power [81]. Also, full precision ADCs require correspondingly high rate data processing at the receiver (for example, with 100 antennas, 500 Msamp/s, the data rate

at the remote radio head is more than 1 Tb/s.) This motivates the need for employing low-resolution ADCs in the BS of a massive MIMO system [8, 82–85]. On the other hand, the capacity analysis of coarsely quantized MIMO systems shows that, under the assumption of perfect channel state information at the receiver (CSIR), at low signal to noise ratio (SNR) and up to a first order approximation in SNR, the mutual information of a MIMO system employing 1-bit ADCs is $2/\pi$ times that of a MIMO system without quantization. Also, quardrature phase shift keying (QPSK) is the optimal modulation scheme in terms of the achievable rate at low SNRs, and under i.i.d. Rayleigh fading, the ergodic capacity of a 1-bit quantized system increases linearly with the number of receive (rx) antennas [18, 20, 23, 86]. Due to this promise of close to optimal performance despite the coarse quantization introduced by the ADCs, the receiver design in multiuser massive MIMO systems with low-resolution ADCs is of great practical interest.

The nonlinearities introduced by low resolution quantization leads to three challenges in receiver design. First, *linear* receive processing techniques like zero forcing (ZF), regularized ZF (RZF), and minimum mean square error (MMSE) become highly suboptimal [68]. Second, the received *training* symbols are also subject to low resolution quantization, leading to poor channel estimates. In practice, a large training overhead is needed to obtain reliable channel estimates [69, 70]. We note that although the primary task of the receiver is to correctly decode the data symbols, obtaining accurate CSIR is important for other tasks such as link adaptation. Third, practical channel decoders exhibit significantly lower codeword error rates when they are provided *soft symbol estimates* (i.e., log likelihood ratios of the coded data bits) as input, compared to the case where hard decisions are performed on the data symbols prior to channel decoding. Obtaining high-quality soft symbol estimates from coarsely quantized samples is challenging. These considerations motivate the need to devise novel techniques for joint channel estimation and soft symbol decoding with low-resolution ADCs, which is the focus of this chapter.

### 2.1.1 Prior Art

The problem of data detection in massive MIMO systems with low-resolution ADCs has been studied in the literature [87–99]. Most of these studies assume the availability of perfect CSIR, and perform data decoding in multiuser massive MIMO systems with low-resolution ADCs [87–93]. Techniques for joint channel estimation and data detection with low-resolution ADCs have been developed based on convex optimization [94] and approximate message passing (AMP) [95,96]. Under the assumption of perfect CSIR, [97] and [98] explore weighted Hamming distance based soft detection and minimum mean square error (MMSE) detection, respectively, along with successive interference cancellation (SIC) under 1-bit quantization, while [99] proposes approximate belief propagation (BP) based MIMO detection in coarsely quantized systems. Variational Bayesian inference (VBI) has been used to develop high-performing algorithms for channel estimation and data decoding, without considering the nonlinearities introduced by the low-resolution ADCs [100,101]. It is worth noting that algorithms based on AMP or expectation propagation (EP) are not numerically stable, requiring heuristic modifications to address such issues [100]. In contrast, in this work, we employ a more principled VBI based approach that is globally convergent to a local optimum.

To the best of our knowledge, most of the existing studies on massive MIMO systems do not consider joint channel estimation and soft symbol decoding, spatial correlation, and coarse quantization together. Our solution allows us to intrinsically learn the channel statistics from the quantized observations, which is potentially useful for power control and/or link adaptation also. Moreover, using VBI to compute the posterior distributions of the data leads to low complexity, high performing receivers. Preliminary versions of this work have appeared in [58,59]. Both these works were restricted to independent and identically distributed (i.i.d.) channels. While [58] considered soft symbol decoding with perfect CSIR, [59] extended the approach to joint channel estimation and soft symbol decoding. We present several advances in this work, including the extension to correlated Rayleigh fading channels, unknown channel statistics, etc. We also merge the VB receiver with a BP based LDPC channel decoder and utilize its extrinsic information to adapt the

data prior, which further improves the performance.

We formulate the channel estimation and data decoding as a statistical inference problem in a directed probabilistic graphical model and solve it using a VBI approach. The quantized received signal is the observed variable, while the wireless channel and the $M-$QAM modulated data symbols are the latent variables. The goal is to jointly infer the posterior distributions of both the channel state and the data symbols given the quantized received signal consisting of pilot and data symbols. Since direct computation of the posterior distribution is intractable, we approximate the posterior with a factorized variational approximation over the latent variables, and compute the factors by maximizing the evidence lower bound (ELBO). The ELBO is maximized by minimizing the Kullback-Leibler (KL) divergence between the exact and the factorized distributions. The resulting iterative algorithm converges to a stationary point of the ELBO. This method of using factorized distributions originates from the mean-field approximation in statistical physics; we refer the reader to [102] for an excellent introduction to VBI. A key novelty in our solution lies in the introduction of appropriately chosen latent variables. This leads to closed-form, computationally simple updates, and the resulting iterative algorithm has low complexity and is also fast.

### 2.1.2 Main Contributions

Our main contributions in this chapter are:

1. We model the channel estimation and data decoding problems in the uplink of a massive MIMO system with low-resolution ADCs as a statistical inference problem. We derive the posterior distributions of the channel and the data symbols obtained from the pilot and data observations using the VB framework. We consider the following two cases:

   (a) The *correlated Rayleigh fading* case, where UEs' channels are independent of each other, but the entries of the channel vector between a UE and the BS are correlated. In this case, our algorithm exploits the channel correlation

to improve the channel estimation and data detection performance. In the absence of correlation, the algorithm becomes computationally very simple, with no matrix inversions.

(b) The *no CSIR* case, in which the BS is unaware of the large scale fading coefficients (LSFCs) between the UEs and the BS. We infer the LSFCs using a non-informative inverse Gamma prior. We empirically show that the resulting receiver is only marginally inferior to a receiver that has perfect knowledge of the LSFCs. Learning the LSFCs is potentially useful for downlink precoding and link adaptation.

2. We provide theoretical insights on the structure of the solution by relating it to the MMSE channel estimation and successive interference cancellation based receiver. We show that, when the interference is perfectly cancelled, the resulting channel estimate is the Bayes' optimal MMSE estimator.

3. We merge the VB receiver with a (BP) channel decoder and utilize its extrinsic information to adapt the data prior, which further improves the performance.

4. We study the bit error rate (BER) and channel normalized mean squared error (NMSE) performance of the VB algorithms with respect to different system parameters such as the ADC resolution, data power, pilot power boosting etc. and provide key insights into system design. We also evaluate the performance for both the correlated and uncorrelated fading cases as a function of channel parameters such as the coherence interval, and benchmark it against state-of-the-art bilinear generalized approximate message passing (BiGAMP) based joint channel and data estimator [95] and MMSE soft decoder [1]. Our numerical results show that the VB soft symbol decoder offers excellent performance and fast convergence, and even outperforms an unquantized BiGAMP joint channel estimator and data detector, making it an attractive choice for high data rate applications.

One of the main takeaways from our work is that VBI is a powerful and flexible technique for designing receivers for massive MIMO systems, particularly when the BS

employs low-resolution ADCs. Using VBI to infer the required posterior distributions helps to reduce the pilot overhead required to achieve a given BER. Further, since the algorithm yields soft symbol estimates, it is well suited for coded communications. Another crucial takeaway is that the assumption of perfect CSI at the receiver significantly overestimates the system performance, which we illustrate through empirical studies in Sec. 2.6. Therefore, in a low-resolution ADC setup, since both the pilots and data are coarsely quantized, it is crucial to account for channel estimation errors while designing receivers and evaluating performance.

**Notation**

Matrices and vectors are denoted by boldface uppercase and boldface lowercase letters. $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. $\mathbb{E}$ denotes the expectation operator. $\langle f(\mathbf{X}) \rangle$ denotes the expectation of $f(\mathbf{X})$ with respect to an approximate distribution $q(\mathbf{X})$. $\phi(x)$ and $\Phi(x)$ denote the probability density and cumulative distribution functions of a standard normal random variable evaluated at $x$. $\mathbf{I}_M, \mathbf{0}_M$ denote an identity and a zero matrix of size $M \times M$, respectively. $\mathcal{Q}_b(\cdot)$ denotes an elementwise $b$-bit quantization operation of the real and imaginary components of the argument.

## 2.2 System Model and Problem Statement

We consider the uplink (UL) of a single cell massive MIMO system with $N_r$ rx antennas at the BS and $K$ single transmit antenna user equipments (UEs), where $N_r \geq K$. The UEs encode their raw data bits using a channel code, interleave the coded bits, map the interleaved data bits to the signal constellation, insert the pilot symbols, then up-convert the signal to the carrier frequency and transmit it to the BS. The signal received at the BS is down-converted to the baseband, sampled, and quantized using a $b$-bit ADC before passing it to the digital front-end of the receiver, as shown in Fig. 2.1.

The UEs transmit data simultaneously over a frame consisting of $\tau_p$ known pilot symbols ($\tau_p \geq K$) followed by $\tau_d$ unknown data symbols. The complex baseband pilot and

data symbols received at the BS, denoted by $\mathbf{Z}_p \in \mathbb{C}^{N_r \times \tau_p}$ and $\mathbf{Z}_d \in \mathbb{C}^{N_r \times \tau_d}$, respectively, are given by

$$\begin{aligned} \mathbf{Z}_p &= [\mathbf{z}_{p,1}, \ldots, \mathbf{z}_{p,\tau_p}] = \mathbf{H}\mathbf{X}_p + \mathbf{W}_p, \\ \mathbf{Z}_d &= [\mathbf{z}_{d,1}, \ldots, \mathbf{z}_{d,\tau_d}] = \mathbf{H}\mathbf{X}_d + \mathbf{W}_d, \end{aligned} \tag{2.1}$$

where $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K] \in \mathbb{C}^{N_r \times K}$, $\mathbf{X}_p = \left[\mathbf{x}_{p,1}, \ldots, \mathbf{x}_{p,\tau_p}\right] \in \mathbb{C}^{K \times \tau_p}$, and $\mathbf{X}_d = [\mathbf{x}_{d,1}, \ldots, \mathbf{x}_{d,\tau_d}] \in \mathbb{M}^{K \times \tau_d}$ denote the channel matrix, pilots, and the $M-$QAM modulated transmit data symbols of the $K$ users, respectively. Here, $\mathbf{h}_k \in \mathbb{C}^{N_r}$ is the channel between the $k^{\text{th}}$ UE and the BS, and $\mathbf{x}_{p,t} \in \mathbb{C}^K$ and $\mathbf{x}_{d,t} \in \mathbb{M}^K$ the pilot and data symbols transmitted by the $K$ UEs at the $t^{\text{th}}$ symbol, respectively, where $\mathbb{M}$ denotes the $M$-QAM constellation. Also, $\mathbf{W}_p \in \mathbb{C}^{N_r \times \tau_p}, \mathbf{W}_d \in \mathbb{C}^{N_r \times \tau_d}$ are the additive white Gaussian noise at the receiver during the pilot and data phases, with i.i.d. $\mathcal{CN}(0, \sigma_w^2)$ entries, where $\mathcal{CN}(\mu, \sigma^2)$ denotes the circularly symmetric complex normal distribution with mean $\mu$ and variance $\sigma^2$.

We assume that the data symbols are i.i.d. and are drawn from a uniform probability distribution $P_X$ defined on the signal constellation $\mathbb{M}$. The average data transmit power $\mathbb{E}(|x_{d,kt}|^2)$ of the $k^{\text{th}}$ UE is set to $P_{T,k}$. The pilot transmission power is set to be the same as the data transmit power. Mathematically, $P_{\mathbf{X}_d}(\mathbf{X}_d) = \prod_{k=1}^{K} \prod_{t=1}^{\tau_d} P_X(x_{d,kt})$.

The channel $\mathbf{h}_k$ is modeled as a correlated Rayleigh flat fading channel, which remains constant over a frame consisting of $\tau_p$ pilot symbols and $\tau_d$ data symbols. That is, $\mathbf{h}_k$ is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$, where $\mathbf{R}_k \in \mathbb{C}^{N_r \times N_r}$ is a positive semidefinite covariance matrix. Also, without loss of generality, we include the path loss component in $\mathbf{R}_k$. We denote $\mathbf{R} \triangleq \{\mathbf{R}_1, \ldots, \mathbf{R}_K\}$, and thus, $P_{\mathbf{H}}(\mathbf{H}; \mathbf{R}) = \prod_{k=1}^{K} \frac{1}{\pi^{N_r} \det(\mathbf{R}_k)} \exp\left(-\mathbf{h}_k^H \mathbf{R}_k^{-1} \mathbf{h}_k\right)$. The LSFC of the channel from the $k^{\text{th}}$ UE to the BS (i.e., the diagonal entries of $\mathbf{R}_k$) is denoted by $\beta_k$. We define the SNR of the system as $\frac{\sum_k P_{T,k} \beta_k}{K \sigma_w^2}$, where $P_{T,k}$ is the transmit power of the $k$th UE.

Under the low-resolution ADC architecture, the signals $\mathbf{Z}_p$ and $\mathbf{Z}_d$ are quantized using

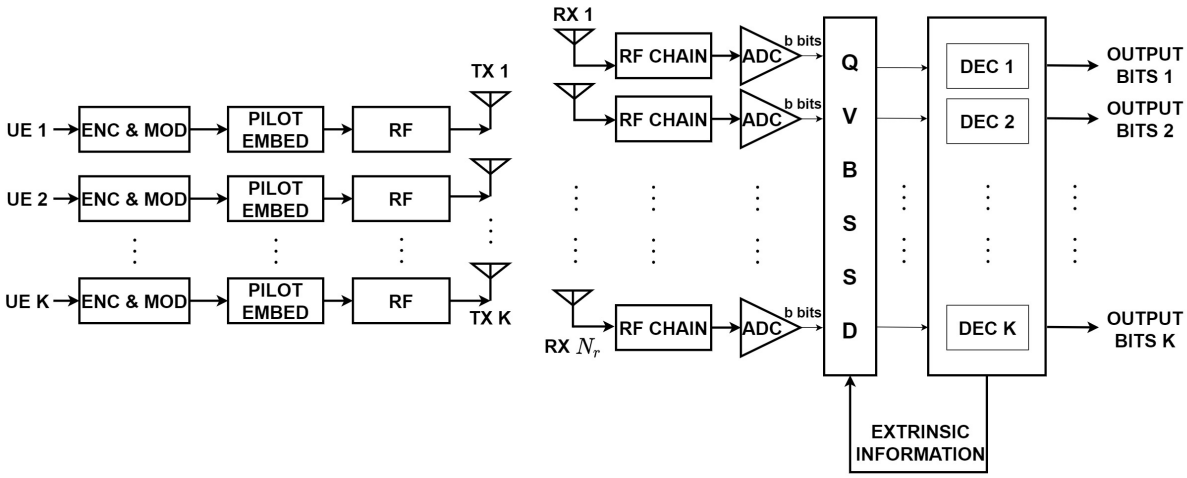


Figure 2.1: Massive MIMO system model with low-resolution ADCs.

a $b-$bit ADC to obtain the baseband observations $\mathbf{Y}_p$ and $\mathbf{Y}_d$ as follows:

$$\begin{aligned} \mathbf{Y}_p &= \mathcal{Q}_b\left(\mathbf{Z}_p\right) = \mathcal{Q}_b\left(\mathbf{H}\mathbf{X}_p + \mathbf{W}_p\right) \\ \mathbf{Y}_d &= \mathcal{Q}_b\left(\mathbf{Z}_d\right) = \mathcal{Q}_b\left(\mathbf{H}\mathbf{X}_d + \mathbf{W}_d\right). \end{aligned} \tag{2.2}$$

The function $\mathcal{Q}_b(\cdot)$ denotes a $b$-bit quantizer operated element-wise on both the real and imaginary parts of the complex input. A $b$-bit quantizer on a real valued input $z$ is defined as $\mathcal{Q}_b(z) = L_i,\ z \in [z_i, z_{i+1}),\ i = 0, 1, \ldots, B-1$, where $B = 2^b$ is the number of quantization levels, $z_0 < z_1 < \cdots < z_B$ are the quantization thresholds, and $L_0, L_1, \ldots, L_{B-1}$ are the quantizer outputs. In this work, for simplicity and concreteness, we consider a uniform quantizer, where $z_l = (-B/2 + l)\Delta, l = 0, \ldots, B$, $\Delta$ is the quantization step size, and $L_l = (z_l + z_{l+1})/2, l = 0, \ldots, B-1$. We set the dynamic range of the real and imaginary parts of the quantizer using the expected received signal power, $P_R$, as $z_0 = -2.5\sqrt{P_R/2}$, $z_B = 2.5\sqrt{P_R/2}$.[1] Our choice of $z_0$ and $z_B$ is motivated by the fact that the absolute value of a Gaussian distributed zero mean real-valued random variable with variance $P_R/2$ exceeds $2.5\sqrt{P_R/2}$ with probability less than 0.01, i.e., the quantizer

[1]In practice, we quantize any value below $z_0$ to $L_0$, and any value above $z_B$ to $L_{B-1}$. Also, in practical wireless systems, an automatic gain control (AGC) and a variable gain amplifier (VGA) are used to ensure that the power in the analog baseband signal is approximately equal to a predefined value, $P_R$, before quantization.

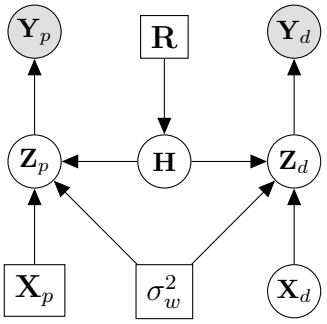


Figure 2.2: Bayesian network graphical model for the quantized massive MIMO wireless communication system.

gets overloaded with low probability. However, the design of quantizers for low-resolution ADC based systems is an interesting problem for future research.

Our goal is to recover the posterior probabilities of the transmitted data symbols from the quantized received signal $\mathbf{Y}_p$ and $\mathbf{Y}_d$. The posterior beliefs are fed to the channel decoder to obtain the decoded data bits. In practice, the equalizer (or detector) and channel decoder can be designed either jointly or as separate blocks. We adopt both approaches to obtain a decoder that is flexible, and comment on their relative merits.

## 2.3 Bayesian Estimation: An Overview

Before explaining our approach to the joint channel estimation and soft symbol decoding, we provide a brief overview about Bayesian estimation and discuss the challenges involved. We represent the quantized received signal model in (2.2) as a Bayesian network graphical model shown in Fig. 2.2. We treat the data symbols and channel as the latent variables and the quantized received samples as the observations. The pilot and data observations $\mathbf{Y}_p$ and $\mathbf{Y}_d$ are represented by shaded circles, while the latent variables $\mathbf{Z}_p$, $\mathbf{Z}_d$, $\mathbf{X}_d$, $\mathbf{H}$ are represented by transparent circles, and deterministic variables $\mathbf{X}_p$, $\mathbf{R}$, and $\sigma_w^2$ are represented by squares. Our goal is to infer the posterior distribution of the channel $\mathbf{H}$ and the data $\mathbf{X}_d$ given the observations $\mathbf{Y}_p, \mathbf{Y}_d$ and the pilots $\mathbf{X}_p$.

The posterior distribution of the channel $\mathbf{H}$ and the data $\mathbf{X}_d$ given the quantized

observations $\mathbf{Y}_p, \mathbf{Y}_d$ and the pilots $\mathbf{X}_p$ is given by[2]

$$P\left(\mathbf{H}, \mathbf{X}_d | \mathbf{Y}_d, \mathbf{Y}_p, \mathbf{X}_p\right) = \frac{P(\mathbf{Y}_p|\mathbf{X}_p, \mathbf{H})P(\mathbf{Y}_d|\mathbf{X}_d, \mathbf{H})P(\mathbf{H})P(\mathbf{X}_d)}{P(\mathbf{Y}_p)P(\mathbf{Y}_d)} \tag{2.3}$$

where

$$P(\mathbf{Y}_p) = \int P(\mathbf{Y}_p|\mathbf{X}_p, \mathbf{H})P(\mathbf{H})\,\mathrm{d}\mathbf{H}, \tag{2.4}$$

$$P(\mathbf{Y}_d) = \int P(\mathbf{Y}_d|\mathbf{X}_d, \mathbf{H})P(\mathbf{H})P(\mathbf{X}_d)\,\mathrm{d}\mathbf{H}\,\mathrm{d}\mathbf{X}_d \tag{2.5}$$

are the marginal likelihoods of $\mathbf{Y}_p$ and $\mathbf{Y}_d$, respectively.

Exact computation of the posterior distributions using the above is computationally intractable, as it requires solving high dimensional integrals over $\mathbf{H}$ and $\mathbf{X}_d$ to obtain the partition functions $P(\mathbf{Y}_p)$ and $P(\mathbf{Y}_d)$. To circumvent this problem, we explore alternative approximate inference approaches, which provide analytically and computationally tractable methods to solve the statistical inference problem.

There are two broad categories of approximate inference techniques: stochastic and deterministic. Stochastic approximate inference is based on sampling methods such as Markov chain Monte Carlo. Though they give exact solutions with sufficiently many samples, they are computationally demanding and therefore impractical for high dimensional statistical inference problems. Another drawback is the difficulty in assessing the convergence, which makes them less amenable for use in the high dimensional inference problems that arise in wireless communications. Deterministic approximate inference provides analytical approximations to the posterior distributions which may also have guaranteed convergence. Examples of deterministic approximate inference approaches include variational Bayes (VB), expectation propagation (EP), belief propagation (BP), loopy BP (LBP), Laplace approximation, etc [102].

Laplace approximation finds a mode of the posterior distribution and approximates it with a Gaussian distribution using a second order Taylor expansion. Since the Laplace

[2] $\mathbf{X}_d$ comes from a discrete $M$-QAM constellation, but we use integrals here for convenience. In the actual derivation, the integrals are replaced by summations.

approximation only considers the properties of the true posterior distribution around one of its modes, it can fail to capture the global properties. EP involves approximating the posterior with another distribution from an exponential family. It imposes a factorized structure on the posterior over the latent variables, and iteratively computes each factor. It minimizes the KL divergence between the original and approximate posterior distributions using a moment matching approach. A major demerit of EP is that it is not guaranteed to converge in general. Moreover, since it relies on moment matching, it is only limited to the models where it is possible to compute the required moments. BP is a message passing based approximate inference procedure which provides an efficient framework in tree-structured probabilistic graphical models. It can also be applied to graphs with loops (loopy BP), but there is no guaranteed convergence to a fixed point.

We approach the joint channel estimation and data detection using an approximate inference framework called VB, in which the posterior distributions of the latent variables are obtained by solving an optimization problem. We approximate the exact posterior using a structured factorized distribution that can be computed in polynomial time in each iteration. We explain this in detail in the forthcoming sections. In the next section, we provide a brief overview of VB.

## 2.4 Variational Bayesian Inference

VB inference is an iterative procedure to compute the approximate joint posterior distributions of the latent variables in a probabilistic graphical model. Consider a probabilistic graphical model with observations $\mathbf{Z} = \{\mathbf{Z}_1, \ldots, \mathbf{Z}_N\}$ and latent variables $\mathbf{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_N\}$. Our goal is to find the posterior distribution $p(\mathbf{X}|\mathbf{Z})$ and model evidence $p(\mathbf{Z})$. We write the logarithm of the model evidence as the sum of two terms as follows:

$$\ln p(\mathbf{Z}) = \mathcal{L}(q) + \mathrm{KL}(q \,\|\, p) \geq \mathcal{L}(q), \tag{2.6}$$

where

$$\mathcal{L}(q) \triangleq \int q(\mathbf{X}) \ln \left\{ \frac{p(\mathbf{Z}, \mathbf{X})}{q(\mathbf{X})} \right\} d\mathbf{X}, \tag{2.7}$$

$$\mathrm{KL}(q \| p) \triangleq - \int q(\mathbf{X}) \ln \left\{ \frac{p(\mathbf{X}|\mathbf{Z})}{q(\mathbf{X})} \right\} d\mathbf{X} \geq 0 \tag{2.8}$$

are the ELBO and the KL divergence terms. Here, $q(\mathbf{X})$ is the approximate posterior distribution which is arbitrary, and can be approximated and optimized. Note that $q(\mathbf{X})$ may depend on the observations $\mathbf{Y}$, but we do not explicitly include it here for the sake of notational brevity.

In the above, maximizing the ELBO $\mathcal{L}(q)$ would render an approximate distribution $q$ that is close to the original model evidence, because the KL divergence is non-negative. The maximum of $\mathcal{L}(q)$ occurs when $q(\mathbf{X}) = p(\mathbf{X}|\mathbf{Z})$, but $p(\mathbf{X}|\mathbf{Z})$ is computationally intractable. Therefore, we impose a fully factorized posterior structure on $q$, i.e., $q(\mathbf{X}) = \prod_{i=1}^{N} q_i(\mathbf{X}_i)$. Substituting this into the ELBO, and considering the optimization over one of the factors, say $q_j$, we get

$$\begin{aligned} \mathcal{L}(q) &= \int \prod_i q_i \left( \ln p(\mathbf{Z}, \mathbf{X}) - \sum_l \ln q_l \right) d\mathbf{X} \\ &= -\mathrm{KL}\left(q_j \| \tilde{p}(\mathbf{Z}, \mathbf{X}_j)\right) + \text{const.} \end{aligned} \tag{2.9}$$

where the constant terms do not depend on $q_j$, and $\tilde{p}(\mathbf{Z}, \mathbf{X}_j)$ is defined using $\ln \tilde{p}(\mathbf{Z}, \mathbf{X}_j) \triangleq \mathbb{E}_{i \neq j} \left[ \ln p(\mathbf{Z}, \mathbf{X}) \right] + \text{const.}$, where the notation $\mathbb{E}_{i \neq j}[.]$ denotes the expectation with respect to the distributions $q_1(\mathbf{X}_1)$, ..., $q_N(\mathbf{X}_N)$ except $q_j(\mathbf{X}_j)$. Now, $\mathcal{L}(q)$ is maximized with respect to $q_j$ when the KL divergence term in (2.9) is minimized, which in turn occurs when $q_j(\mathbf{X}_j) = \tilde{p}(\mathbf{Z}, \mathbf{X}_j)$. Therefore, the optimal marginal distribution is

$$q_j(\mathbf{X}_j) = \text{const} \times \exp\left( \mathbb{E}_{i \neq j} \left[ \ln p(\mathbf{Z}, \mathbf{X}) \right] \right), \tag{2.10}$$

where the constant is chosen such that $q_j$ becomes a probability distribution. From (2.10), the optimal $q_j$ depends on the distributions $\{q_i\}_{i \neq j}$, which also need to be evaluated. The

VBI algorithm proceeds by initializing $N-1$ latent distributions $\{q_i\}_{i=1,\neq j}^{N}$ and finding the other distribution $q_j$ in an iterative manner. This recipe falls in the category of minorization-maximization (MM), which solves a maximization problem by iteratively obtaining a lower bound to the objective function and maximizing it [64]. It is known that MM based optimization is guaranteed to converge to a stationary point of the original optimization problem from any initialization.

In addition to data detection, the soft input channel decoder also involves information bits as latent variables, which adds to the complexity of the channel estimation and symbol detection. We include the belief propagation based channel decoder as a separate block which seamlessly integrates with our solution via the extrinsic information. This architecture provides the flexibility of allowing different combinations for scheduling the exchange of messages between the VB (for channel estimation and symbol detection) and BP (for channel decoding) algorithms. For e.g., one can run several iterations of VB after which the BP can be scheduled, or both can run one after another after executing only one iteration each, etc. We now proceed to develop a VBI based algorithm for our joint channel estimation and data decoding problem in the uplink of a massive MIMO system with low-resolution ADCs.

## 2.5 Variational Bayes Joint Channel Estimation and Soft Symbol Decoding

In this section, we derive the joint channel estimation and soft symbol decoding algorithm for an uplink massive MIMO system with low-resolution ADCs. Typically, in iterative detection and decoding methods, point estimates of the channel and data symbols are computed and converted to bit LLRs using the input-output relationships of wireless channels. These LLRs are input to the channel decoder to obtain the decoded information bits. Iterative channel estimation and data decoding in *unquantized* MIMO systems has been well-studied in the literature. However, this thesis addresses the channel estimation and data decoding in massive MIMO systems with *low-resolution ADCs* being

employed at the receivers (and in addition, considering spatially correlated channels between the users and the base station.) We provide a unified framework for joint channel estimation and data decoding in massive MIMO single-carrier systems, in contrast to the conventional iterative approach. Moreover, the developed solution directly infers the posterior distributions of the channel and data symbols rather than only point estimates, and in a computationally and analytically tractable manner, using which bit-LLRs are computed. This method of obtaining the LLRs improves their quality, which translates to better channel decoding performance. Moreover, the guaranteed convergence compared to other approximate inference approaches makes it appealing from a practical standpoint. Therefore, the combined benefits of inferring the posterior distributions of the channel and data symbols using a unified framework, its guaranteed and fast convergence, motivates the use of a VB procedure for solving channel estimation and data decoding problems in massive MIMO systems.

We first present our algorithm in the simplest case of perfect CSIR with unquantized received data signal (Sec. 2.5.1), followed by the case with statistical CSIR and unquantized pilot and data signals (Sec. 2.5.2), then the case of statistical CSIR with quantized received signals (Sec. 2.5.3), and finally no CSIR with quantized pilot and data signals (Sec. 2.5.4). In Sec. 2.5.5, we combine the VB receiver with a belief propagation (BP) based channel decoder to further improve the coded BER. As explained earlier, this gives flexibility in configuring the VB and BP either jointly or separately. Our approach of presenting our VB based algorithms by adding each imperfection one after the other not only eases the exposition, it also allows us to benchmark results, analytically interpret the algorithms, and makes it convenient to apply the algorithms in a variety of settings. In Sec. 2.5.6, we elucidate the computational complexity of the variational Bayesian algorithm in the quantized case, and discuss ways to further reduce the complexity.

### 2.5.1 Perfect CSIR and Unquantized Observations

We assume that the receiver has perfect knowledge of the channel and has access to unquantized observations $\mathbf{Z}_d = \mathbf{H}\mathbf{X}_d + \mathbf{W}_d$. Here, the transmit data symbols $\mathbf{X}_d$ are the

latent variables whose posterior distribution needs to be inferred. The logarithm of the joint distribution of the observations and the latent variables can be written as

$$\ln p(\mathbf{Z}_d, \mathbf{X}_d|\mathbf{H}, \sigma_w^2) = \ln p(\mathbf{Z}_d|\mathbf{X}_d, \mathbf{H}, \sigma_w^2) + \ln p(\mathbf{X}_d). \tag{2.11}$$

Since the additive noise is complex Gaussian distributed,

$$p\left(\mathbf{Z}_d|\mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) = \frac{1}{(\pi\sigma_w^2)^{\tau_d N_r}} \times \exp\left(-\frac{1}{\sigma_w^2}\sum_{t=1}^{\tau_d} \|\mathbf{z}_{d,t} - \mathbf{H}x_{d,t}\|_2^2\right), \tag{2.12}$$

and since $\mathbf{X}_d$ is uniformly distributed over the $M$-QAM constellation, $p(\mathbf{X}_d) = \frac{1}{M^{(\tau_d K)}}$. Our goal is to find the posterior distribution $p(\mathbf{X}_d|\mathbf{Z}_d, \mathbf{H}, \sigma_w^2)$. We write the logarithm of the model evidence, $\ln p(\mathbf{Z}_d)$, as follows:

$$\ln p(\mathbf{Z}_d) = \mathcal{L}(q) + \mathrm{KL}(q \,\|\, p) \geq \mathcal{L}(q), \tag{2.13}$$

where

$$\mathcal{L}(q) \triangleq \int q(\mathbf{X}_d) \ln\left\{\frac{p(\mathbf{X}_d, \mathbf{Z}_d)}{q(\mathbf{X}_d)}\right\} d\mathbf{X}_d, \tag{2.14}$$

$$\mathrm{KL}(q\|p) \triangleq -\int q(\mathbf{X}_d) \ln\left\{\frac{p(\mathbf{X}_d|\mathbf{Z}_d)}{q(\mathbf{X}_d)}\right\} d\mathbf{X}_d \geq 0 \tag{2.15}$$

are the ELBO and the KL divergence terms, respectively. Here, $q(\mathbf{X}_d)$ is an approximate posterior distribution, which is arbitrary and can be optimized. We formally state the ELBO optimization problem as

$$\begin{aligned} q_{\text{opt}} &= \arg\max_{q\in\mathcal{P}} \ \mathcal{L}(q) \\ &= \arg\min_{q\in\mathcal{P}} \ \mathrm{KL}(q \,\|\, p), \end{aligned} \tag{2.16}$$

where $\mathcal{P}$ is the space of probability distributions.

In the above, maximizing the ELBO $\mathcal{L}(q)$ would render an approximate distribution $q$ that is close to the original model evidence, because the KL divergence is non-negative.

The maximum of $\mathcal{L}(q)$ occurs when $q(\mathbf{X}_d) = p(\mathbf{X}_d|\mathbf{Z}_d)$, but $p(\mathbf{X}_d|\mathbf{Z}_d)$ is computationally intractable. Therefore, we impose a factorized posterior structure on $q$, i.e., $q(\mathbf{X}_d) = \prod_{k=1}^{K}\prod_{t=1}^{\tau_d} q_{kt}(x_{d,kt})$. Substituting this into the ELBO, and considering one of the factors, say $q_{kt} \triangleq q(x_{d,kt})$, we get

$$\mathcal{L}(q) = -\mathrm{KL}\left(q_{kt} \| \tilde{p}(\mathbf{Z}_d, x_{d,kt})\right) + \text{const.} \tag{2.17}$$

where the `const.` terms do not depend on $q_{kt}$, and $\tilde{p}(\mathbf{Z}_d, x_{d,kt})$ is defined using $\ln \tilde{p}(\mathbf{Z}_d, x_{d,kt}) \triangleq \mathbb{E}_{(i,l)\neq(k,t)}\left[\ln p(\mathbf{Z}_d, \mathbf{X}_d)\right] + \text{const.}$, where the notation $\mathbb{E}_{(i,l)\neq(k,t)}[\cdot]$ denotes the expectation with respect to the distributions $q_{11}(x_{d,11}), \ldots, q_{K\tau_d}(x_{d,K\tau_d})$ except $q_{kt}(x_{d,kt})$. Now, $\mathcal{L}(q)$ is maximized with respect to $q_{kt}$ when the KL divergence term in (2.17) is minimized, which in turn occurs when $q_{kt}(x_{d,kt}) = \tilde{p}(\mathbf{Z}_d, x_{d,kt})$. Therefore, the optimal marginal distribution is

$$q_{kt}(x_{d,kt}) = \text{const.} \times \exp\left(\mathbb{E}_{(i,l)\neq(k,t)}\left[\ln p(\mathbf{Z}_d, \mathbf{X}_d)\right]\right), \tag{2.18}$$

where the constant is chosen such that $q_{kt}$ becomes a probability distribution. From (2.18), the optimal $q_{kt}$ depends on the distributions $\{q_{i\ell}\}_{(i,\ell)\neq(k,t)}$, which also need to be evaluated. The VBI algorithm proceeds by initializing latent distributions $\{q_{i\ell}\}_{(i,\ell)\neq(k,t)}$ and finding $q_{kt}$ in an iterative manner across all $k$ and $t$. This recipe falls in the category of minorization-maximization (MM), which solves a maximization problem by iteratively obtaining a lower bound to the objective function and maximizing it [64]. It is known that MM based optimization is guaranteed to converge to a stationary point of the original optimization problem from any initialization.

Thus, based on the above discussion, we impose a fully factorized structure on the approximate posterior, namely, $p\left(\mathbf{X}_d|\mathbf{Z}_d, \mathbf{H}, \sigma_w^2\right) \approx \prod_{k=1}^{K}\prod_{t}^{\tau_d} q_{x_{d,kt}}\left(x_{d,kt}\right)$. Now, we compute the approximate posterior $q(x_{d,kt})$, $k = 1, \ldots, K$, $t = 1, \ldots, \tau_d$. To this end, we compute the following expectation using the joint distribution in (2.11):

$$\ln q_{x_{d,kt}}\left(x_{d,kt}\right) = \left\langle \ln p\left(\mathbf{Z}_d|\mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) + \ln p\left(\mathbf{X}_d\right)\right\rangle, \tag{2.19}$$

where $\langle\cdot\rangle$ denotes the expectation with respect to all the latent variables except $x_{d,kt}$, and is computed using the approximate posterior distribution $q(\mathbf{X})$. Taking the exponential on both sides of (2.19) and normalizing to obtain a probability distribution, after some algebra, we get

$$q_{x_{d,kt}}(s_m) = \frac{\exp\Big(f_{k,t}(s_m)\Big)}{\sum_{s'\in\mathbb{M}} \exp\Big(f_{k,t}(s')\Big)}, \;\; m = 1, \ldots, M, \tag{2.20}$$

with $f_{k,t}(s)$ defined in (2.21).

$$f_{k,t}(s) \triangleq -\frac{1}{\sigma_w^2}\Big( \|\mathbf{h}_k\|^2 |s|^2 - 2\Re\Big[\mathbf{h}_k^H\Big(\mathbf{z}_{d,t} - \sum_{\substack{k'=1\\k'\neq k}}^{K} \mathbf{h}_{k'}\big\langle x_{d,k't}\big\rangle\Big)s^*\Big]\Big) + \ln p(x_{d,kt} = s). \tag{2.21}$$

$$s_{k,t}^{\text{opt}} = \arg\min_{s\in\mathbb{M}} \frac{1}{\sigma_w^2}\Big( \|\mathbf{h}_k\|^2 |s|^2 - 2\Re\Big[\mathbf{h}_k^H\Big(\mathbf{z}_{d,t} - \sum_{\substack{k'=1\\k'\neq k}}^{K} \mathbf{h}_{k'}\big\langle x_{d,k't}\big\rangle\Big)s^*\Big]\Big) - \ln p(x_{d,kt} = s). \tag{2.22}$$

From (2.20), we see that $q_{x_{d,kt}}$ is Boltzmann distributed (it is also known as the *softmax* function and is widely used in machine learning applications as an output activation function.) The mean and mean squared values of $x_{d,kt}$ under the distribution $q_{x_{d,kt}}$ can be computed as $\langle x_{d,kt}\rangle = \sum_{m=1}^{M} s_m\, q_{x_{d,kt}}(s_m)$, $\langle |x_{d,kt}|^2\rangle = \sum_{m=1}^{M} |s_m|^2\, q_{x_{d,kt}}(s_m)$.

The distribution $q_{x_{d,kt}}$ depends only on the means of the other latent variables $\{x_{d,k't}\}_{k'\neq k}$ taken with respect to their approximate distributions. Therefore, we initialize the means of the distributions $q_{x_{d,kt}}$, $k = 1, \ldots, K$ and $t = 1, \ldots, \tau_d$ arbitrarily and apply (2.20) iteratively across $k$ and $t$ till the log likelihood in (2.19) converges for all $k$ and $t$. As mentioned earlier, this procedure converges to a stationary point of the original optimization problem of choosing $q$ to maximize the model evidence in (2.13). The pseudocode for the VB procedure for the soft symbol decoding is shown in Algorithm 1.

**Remark 1.** *In an uncoded system, the receiver decodes the transmitted symbol by solving an optimization problem mentioned in* (2.22)*. Given the aprori distribution of the data*

**Algorithm 1:** VB Soft Symbol Decoding with Perfect CSIR

**Input:** $\mathbf{Z}_d, \{\mathbf{h}_1, \ldots, \mathbf{h}_K\}, M, \mathbb{M} = \{s_1, \ldots, s_M\}, \tau_d, \sigma_w$

**Output:** $q_{x_{d,kt}}, \langle x_{d,kt} \rangle \, \forall k \in [K], t \in [\tau_d]$

Initialize $q_{x_{d,kt}}(x_{d,kt}), \langle x_{d,kt} \rangle = 0, \forall k \in [K], t \in [\tau_d]$
**repeat**
  **for** $k = 1$ to $K$, $t = 1$ to $\tau_d$ **do**
    **for** $m = 1$ to $M$ **do**
      Compute $q_{x_{d,kt}}(s_m)$ using (2.20) and (2.21).
    **end for**
    $\langle x_{d,kt} \rangle = \sum_{s \in \mathbb{M}} s \, q_{x_{d,kt}}(s)$.
    $\langle |x_{d,kt}|^2 \rangle = \sum_{s \in \mathbb{M}} |s|^2 \, q_{x_{d,kt}}(s)$.
  **end for**
**until** a stopping condition is met.

*symbols $p(x_{d,kt} = s)$, the above equation can be interpreted as successive interference cancellation followed by maximum aposteriori probability (MAP) decoding of a single user's data symbol. This reduces the exponential complexity of a MAP receiver to linear complexity in the number of UEs. However, in low resolution quantized systems (discussed later), the analog input to the receiver is not observed, and the above approach may lead to suboptimal performance. In this case, VBI provides a systematic approach to updating the soft symbol estimates by incorporating the posterior distribution induced by the quantization.*

In the next subsection, we remove the perfect CSIR assumption and infer the posterior distributions of the channel as well as the data symbols given the unquantized observations.

### 2.5.2 Statistical CSIR and Unquantized Observations

In this subsection, we solve the joint channel estimation and soft symbol decoding problem in an unquantized setup. Therefore, the unquantized received signals

$$\mathbf{Z}_p = [\mathbf{z}_{p,1}, \ldots, \mathbf{z}_{p,\tau_p}] = \mathbf{H}\mathbf{X}_p + \mathbf{W}_p \tag{2.23}$$

$$\mathbf{Z}_d = [\mathbf{z}_{d,1}, \ldots, \mathbf{z}_{d,\tau_d}] = \mathbf{H}\mathbf{X}_d + \mathbf{W}_d \tag{2.24}$$

are the observations. The channel $\mathbf{H}$ and the data symbols $\mathbf{X}_d$ are the latent variables whose posterior distributions need to be inferred. For this problem, a message passing based approach is adopted in [100, 101] in a MIMO OFDM setup where the channel has a diagonal structure, whereas we use an iterative VBI framework which can be used for arbitrary channels. Similar to the perfect CSIR case, we impose a factorized structure on the posterior as follows: $p\left(\mathbf{X}_d, \mathbf{H}|\mathbf{Z}_p, \mathbf{Z}_d, \mathbf{X}_p; \mathbf{R}, \sigma_w^2\right) \approx q\left(\mathbf{X}_d\right) q\left(\mathbf{H}\right)$,[3] where $q\left(\mathbf{H}\right) = \prod_{k=1}^{K} q\left(\mathbf{h}_k\right), q\left(\mathbf{X}_d\right) = \prod_{k=1}^{K} \prod_{t=1}^{\tau_d} q_{x_{d,kt}}\left(x_{d,kt}\right)$.

Using the structure of the model in Fig. 2.2, we factorize the joint distribution of the observations and latent variables as

$$P\left(\mathbf{Z}_p, \mathbf{Z}_d, \mathbf{H}, \mathbf{X}_d|\mathbf{X}_p; \mathbf{R}, \sigma_w^2\right) = P\left(\mathbf{Z}_p|\mathbf{X}_p, \mathbf{H}; \sigma_w^2\right) P\left(\mathbf{Z}_d|\mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) P\left(\mathbf{H}|\mathbf{R}\right) P\left(\mathbf{X}_d\right). \tag{2.25}$$

The likelihood functions of the pilot and data observations $\mathbf{Z}_p$ and $\mathbf{Z}_d$ given the channel $\mathbf{H}$, the pilots $\mathbf{X}_p$, and the data $\mathbf{X}_d$ are

$$P\left(\mathbf{Z}_p|\mathbf{X}_p, \mathbf{H}; \sigma_w^2\right) = \frac{1}{\left(\pi\sigma_w^2\right)^{\tau_p N_r}} \exp\left(-\frac{1}{\sigma_w^2} \sum_{t=1}^{\tau_p} \left\|\mathbf{z}_{p,t} - \mathbf{H}\mathbf{x}_{p,t}\right\|^2\right), \tag{2.26}$$

$$P\left(\mathbf{Z}_d|\mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) = \frac{1}{\left(\pi\sigma_w^2\right)^{\tau_d N_r}} \exp\left(-\frac{1}{\sigma_w^2} \sum_{t=1}^{\tau_d} \left\|\mathbf{z}_{d,t} - \mathbf{H}\mathbf{x}_{d,t}\right\|^2\right), \tag{2.27}$$

[3] We drop the subscripts on $q$ for notational simplicity.

respectively. Also, the prior distribution on the channel $\mathbf{H}$ is

$$p\left(\mathbf{H}|\mathbf{R}\right) = \prod_{k=1}^{K} \frac{1}{\pi^{N_r} \det(\mathbf{R}_k)} \exp\left( - \sum_{k=1}^{K} \mathbf{h}_k^H \mathbf{R}_k^{-1} \mathbf{h}_k \right). \tag{2.28}$$

As before, the data symbols $\mathbf{X}_d$ are drawn i.i.d. uniformly from the $M$-QAM constellation.

Our goal is to compute the approximate posterior distributions $q(\mathbf{X}_d)$ and $q(\mathbf{H})$. As mentioned in the previous section, we optimize the distributions by minimizing the KL divergence between the original and approximate posterior. We present the steps associated in obtaining the approximate posterior distributions $q(\mathbf{X}_d)$ and $q(\mathbf{H})$ below.

#### 2.5.2.1 Computation of $q(\mathbf{h}_k)$

We compute the approximate posterior distribution $q(\mathbf{h}_k)$ of the channel between the $k^{\text{th}}$ user and the BS using the joint distribution in (2.25), as follows:

$$\ln q\left(\mathbf{h}_k\right) \propto \left\langle \ln p\left(\mathbf{Z}_p|\mathbf{X}_p, \mathbf{H}; \sigma_w^2\right) + \ln p\left(\mathbf{Z}_d|\mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) + \ln p\left(\mathbf{H}|\mathbf{R}\right) \right\rangle, \tag{2.29}$$

where $\langle \cdot \rangle$ denotes the expectation of the joint distribution over all the latent variables excluding $\mathbf{h}_k$. In (2.29), we only include the terms involving $\mathbf{h}_k$. Upon simplification, we get (2.30).

$$\begin{aligned} \ln q\left(\mathbf{h}_k\right) \propto -\Bigg[ & \mathbf{h}_k^H \left\{ \frac{\sum_{t=1}^{\tau_p} |x_{p,kt}|^2 + \sum_{t=1}^{\tau_d} \langle |x_{d,kt}|^2 \rangle}{\sigma_w^2} \mathbf{I}_{N_r} + \mathbf{R}_k^{-1} \right\} \mathbf{h}_k \\ & - 2\Re\Bigg\{ \mathbf{h}_k^H \Bigg( \frac{1}{\sigma_w^2} \Bigg( \sum_{t=1}^{\tau_p} \Big( \mathbf{z}_{p,t} - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle x_{p,k't} \Big) x_{p,kt}^* \\ & + \sum_{t=1}^{\tau_d} \Big( \mathbf{z}_{d,t} - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle \langle x_{d,k't} \rangle \Big) \langle x_{d,kt}^* \rangle \Bigg) \Bigg) \Bigg\} \Bigg]. \end{aligned} \tag{2.30}$$

$$\mathbf{\Sigma}_{\mathbf{h}_k} = \left( \frac{\sum_{t=1}^{\tau_p} |x_{p,kt}|^2 + \sum_{t=1}^{\tau_d} \langle |x_{d,kt}|^2 \rangle}{\sigma_w^2} \mathbf{I}_{N_r} + \mathbf{R}_k^{-1} \right)^{-1}, \tag{2.31}$$

$$\langle \mathbf{h}_k \rangle = \frac{1}{\sigma_w^2} \mathbf{\Sigma}_{\mathbf{h}_k} \Bigg( \sum_{t=1}^{\tau_p} \Big( \mathbf{z}_{p,t} - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle x_{p,k't} \Big) x_{p,kt}^* + \sum_{t=1}^{\tau_d} \Big( \mathbf{z}_{d,t} - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle \langle x_{d,k't} \rangle \Big) \langle x_{d,kt}^* \rangle \Bigg). \tag{2.32}$$

Taking the exponential on both sides of (2.30), we see that the structure is that of a complex normal distribution with covariance and mean given by (2.31) and (2.32), respectively.

**Remark 2.** *We provide an interesting interpretation of the channel estimate in* (2.32)*. Consider a single user MIMO channel estimation problem with pilots* $\mathbf{x}_p \in \mathbb{C}^{\tau_p \times 1}$ *transmitted over a duration of* $\tau_p$ *symbols. The received pilot symbols are given by* $\mathbf{Y}_p = \mathbf{h}\mathbf{x}_p^H + \mathbf{W}_p \in \mathbb{C}^{N_r \times \tau_p}$*, where* $\mathbf{h} \in \mathbb{C}^{N_r \times 1}$ *is the channel distributed as* $\mathcal{CN}(\mathbf{0}, \mathbf{R})$ *and* $\mathbf{W}_p \in \mathbb{C}^{N_r \times \tau_p}$ *is the additive noise whose entries are i.i.d.* $\mathcal{CN}(0, \sigma_w^2)$*. Post-multiplying* $\mathbf{Y}_p$ *with* $\mathbf{x}_p$*, we get* $\mathbf{Y}_p \mathbf{x}_p = \mathbf{h}\|\mathbf{x}_p\|_2^2 + \mathbf{W}_p \mathbf{x}_p$*. The MMSE estimate of* $\mathbf{h}$ *can be computed to be*

$$\widehat{\mathbf{h}}^{\mathbf{MMSE}} = \left( \|\mathbf{x}_p\|_2^2 \mathbf{I}_{N_r} + \sigma_w^2 \mathbf{R}^{-1} \right)^{-1} \mathbf{Y}_p \mathbf{x}_p. \tag{2.33}$$

*Now, comparing the solution obtained in* (2.32) *and* (2.31) *with* (2.33)*, we see that the channel estimate* $\langle \mathbf{h}_k \rangle$ *of the* $k^{th}$ *user in* (2.32) *is an MMSE estimate assuming the interference caused by the remaining* $K-1$ *users is cancelled successfully. Therefore, by using a factorized structure on the posterior distribution, and assuming successful interference cancellation, the channel estimate obtained in* (2.32) *matches with the* Bayes' optimal MMSE estimator*.*

#### 2.5.2.2 Computation of $q_{x_{d,kt}}(x_{d,kt})$

In this subsection, we present the steps involved in the computation of the approximate posterior distribution $q_{x_{d,kt}}(x_{d,kt})$ of the $k^{\text{th}}$ user's data symbol transmitted during the $t^{\text{th}}$ symbol interval. Similar to the computation of $q(\mathbf{h}_k)$, we use the joint distribution (2.25)

to compute the expectation in (2.34) with respect to all the latent variables except $x_{d,kt}$.

$$\ln q_{x_{d,kt}}(x_{d,kt}) \propto -\frac{1}{\sigma_w^2}\Big(\big\langle \|\mathbf{h}_k\|^2 \big\rangle |x_{d,kt}|^2 - 2\Re\Big[\big\langle \mathbf{h}_k \big\rangle^H \Big(\mathbf{z}_{d,t} - \sum_{\substack{k'=1 \\ k'\neq k}}^{K} \big\langle \mathbf{h}_{k'} \big\rangle \big\langle x_{d,k't} \big\rangle\Big) x_{d,kt}^*\Big]\Big) + \ln p(x_{d,kt}). \tag{2.34}$$

$$f_{k,t}(s) = -\frac{1}{\sigma_w^2}\Big(\big\langle \|\mathbf{h}_k\|^2 \big\rangle |s|^2 - 2\Re\Big[\big\langle \mathbf{h}_k \big\rangle^H \Big(\mathbf{z}_{d,t} - \sum_{\substack{k'=1 \\ k'\neq k}}^{K} \big\langle \mathbf{h}_{k'} \big\rangle \big\langle x_{d,k't} \big\rangle\Big) s^*\Big]\Big) + \ln p(x_{d,kt} = s). \tag{2.35}$$

We substitute the values of $x_{d,kt}$ from the $M-$QAM constellation in the above equation to get $q_{x_{d,kt}}(x_{d,kt} = s)$ in the same form as (2.20), with $f_{k,t}(s)$ defined in (2.35), where $s$ belongs to a symbol from the $M-$QAM constellation. Similar to the perfect CSIR case, $q_{x_{d,kt}}(x_{d,kt})$ is Boltzmann distributed. We compute the mean and mean squared values of $x_{d,kt}$ in the same manner as in the previous subsection. We present the pseudocode for the above procedure in Algorithm 2.

**Remark 3.** *Both the channel estimate in* (2.32) *and the soft symbol estimator in* (2.35) *depend on all the data symbols. Since the channel remains constant over the entire frame, iterative channel estimation and data detection entails using all the data symbols. Symbol-by-symbol detection is not optimal here, unlike the case where either perfect CSIR is assumed to be available or pilot-only based channel estimates are used for data detection.*

Note that the approximate marginal posterior distributions of the latent variables are dependent on each other. The algorithm runs by randomly initializing the statistics of the factorized distributions of the latent variables, and cycling through to iteratively update the distributions.

**Algorithm 2:** VB Joint Channel Estimation and Soft Symbol Decoding with Statistical CSIR

**Input:** $\mathbf{Z}_p, \mathbf{Z}_d, \{\mathbf{R}_1, \ldots, \mathbf{R}_K\}, \mathbf{X}_p, M, \mathbb{M} = \{s_1, \ldots, s_M\}$, $\tau_p$, $\tau_d$, $\sigma_w$

**Output:** $\{\langle \mathbf{h}_1 \rangle, \ldots, \langle \mathbf{h}_K \rangle\}, q_{x_{d,kt}}, \langle x_{d,kt} \rangle \, \forall k \in [K], t \in [\tau_d]$

Initialize $q_{x_{d,kt}}(x_{d,kt}), \langle x_{d,kt} \rangle = 0, \forall k \in [K], t \in [\tau_d]$
**repeat**
  **for** $k = 1$ to $K$, $t = 1$ to $\tau_d$ **do**
    **for** $m = 1$ to $M$ **do**
      Compute $q_{x_{d,kt}}(s_m)$ using (2.20) and (2.35).
    **end for**
    $\langle x_{d,kt} \rangle = \sum_{s \in \mathbb{M}} s \, q_{x_{d,kt}}(s)$.
    $\langle |x_{d,kt}|^2 \rangle = \sum_{s \in \mathbb{M}} |s|^2 \, q_{x_{d,kt}}(s)$.
  **end for**
  **for** $k = 1$ to $K$ **do**
    Compute $\mathbf{\Sigma}_{\mathbf{h}_k}$ and $\langle \mathbf{h}_k \rangle$ using (2.31) and (2.32), respectively.
  **end for**
**until** a stopping condition is met.

### 2.5.3 Statistical CSIR and Quantized Observations

In this subsection, we infer the marginal posterior distributions of the data symbols and the channel given the *quantized* observations $\mathbf{Y}_p = \mathcal{Q}_b(\mathbf{Z}_p) = \mathcal{Q}_b(\mathbf{H}\mathbf{X}_p + \mathbf{W}_p)$, $\mathbf{Y}_d = \mathcal{Q}_b(\mathbf{Z}_d) = \mathcal{Q}_b(\mathbf{H}\mathbf{X}_d + \mathbf{W}_d)$ and the pilot symbols $\mathbf{X}_p$. The joint distribution of the observations and latent variables is factorized as

$$p\left(\mathbf{Y}_p, \mathbf{Y}_d, \mathbf{H}, \mathbf{X}_d | \mathbf{X}_p; \sigma_w^2, \mathbf{R}\right) = p\left(\mathbf{Y}_p | \mathbf{X}_p, \mathbf{H}; \sigma_w^2\right) p\left(\mathbf{Y}_d | \mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) p\left(\mathbf{H} | \mathbf{R}\right) p\left(\mathbf{X}_d\right).$$

The conditional distribution of the quantized observations $\mathbf{Y}_d$ given $\mathbf{H}, \mathbf{X}_d$ is given by

$$p\left(\mathbf{Y}_d | \mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) = \int_{\mathbf{Z}_d} p\left(\mathbf{Y}_d \,|\, \mathbf{Z}_d\right) p\left(\mathbf{Z}_d | \mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) d\,\mathbf{Z}_d$$

$$= \int_{\mathbf{Z}_d^{(\mathrm{lo})}}^{\mathbf{Z}_d^{(\mathrm{hi})}} \frac{1}{(\pi\sigma_w^2)^{N_r\tau_d}} \exp\left(-\frac{1}{\sigma_w^2}\left\|\mathbf{Z}_d - \mathbf{H}\mathbf{X}_d\right\|_F^2\right) d\,\mathbf{Z}_d, \tag{2.36}$$

where $\mathbf{Z}_d^{(\mathrm{lo})}$ and $\mathbf{Z}_d^{(\mathrm{hi})}$ are the lower and upper thresholds of the quantizer corresponding to the observation $\mathbf{Y}_d$. The conditional distribution of the quantized pilot observations also has a similar structure as (2.36). Now, to obtain a closed form expression for the approximate posterior distributions of the latent variables using VBI, it is necessary to compute the expectation of the logarithm of the difference of two cumulative distribution functions of a complex normal random vector, which is not straightforward. We circumvent this problem by introducing the unquantized pilot and data observations as latent variables. We will see that this leads to a convenient analytical expression for the posterior distribution.

The conditional distributions of the unquantized observations and the channel are as derived in Sec. 2.5.2. The conditional distributions of the quantized observations, given the unquantized received signals, are given by

$$\begin{aligned} p\left(\mathbf{Y}_d|\mathbf{Z}_d\right) &= \mathbb{1}\big(\mathbf{Z}_d \in [\mathbf{Z}_d^{(\mathrm{lo})}, \mathbf{Z}_d^{(\mathrm{hi})}]\big), \\ p\left(\mathbf{Y}_p|\mathbf{Z}_p\right) &= \mathbb{1}\big(\mathbf{Z}_p \in [\mathbf{Z}_p^{(\mathrm{lo})}, \mathbf{Z}_p^{(\mathrm{hi})}]\big), \end{aligned} \tag{2.37}$$

where $\mathbb{1}(\cdot)$ is the indicator function and the arguments of the indicator function $\mathbf{Z}_d^{(\mathrm{lo})}, \mathbf{Z}_d^{(\mathrm{hi})}$ and $\mathbf{Z}_p^{(\mathrm{lo})}, \mathbf{Z}_p^{(\mathrm{hi})}$ are the lower and upper limits of the quantizer corresponding to the observations $\mathbf{Y}_d$ and $\mathbf{Y}_p$, respectively. To elaborate, let us take an example of quantization of a scalar $z$ ($z^{(\mathrm{lo})} \leq z \leq z^{(\mathrm{hi})}$) to a value $y$ that takes values from a discrete set $\mathbb{Q}$ with cardinality $2^b$, where $b$ is the number of bits of the quantizer. Note that this discrete set depends on the choice of the quantizer. For example, a uniform quantizer contains uniformly spaced values. Now, the value to be quantized fully specifies the output of the quantizer, i.e., the probability mass function of the quantizer output is a Kronecker delta function.

We impose a factorized structure on the posterior distribution of the latent variables:

$$p\left(\mathbf{Z}_p, \mathbf{Z}_d, \mathbf{X}_d, \mathbf{H} | \mathbf{Y}_p, \mathbf{Y}_d, \mathbf{X}_p, \mathbf{R}, \sigma_w^2\right) \approx q\left(\mathbf{Z}_p\right) q\left(\mathbf{Z}_d\right) q\left(\mathbf{X}_d\right) q\left(\mathbf{H}\right), \tag{2.38}$$

where

$$q\left(\mathbf{H}\right) = \prod_{k=1}^{K} q\left(\mathbf{h}_k\right), \qquad q\left(\mathbf{X}_d\right) = \prod_{k=1}^{K} \prod_{t=1}^{\tau_d} q_{x_{d,kt}}\left(x_{d,kt}\right), \tag{2.39}$$

$$q\left(\mathbf{Z}_d\right) = \prod_{t=1}^{\tau_d} q\left(\mathbf{z}_{d,t}\right), \qquad q\left(\mathbf{Z}_p\right) = \prod_{t=1}^{\tau_p} q\left(\mathbf{z}_{p,t}\right). \tag{2.40}$$

We compute the approximate marginal posterior distributions of the latent variables in a similar manner as in Sections 2.5.1 and 2.5.2, and outline the differences below.

#### 2.5.3.1 Computation of $q(\mathbf{h}_k)$

We compute the approximate posterior distribution $q(\mathbf{h}_k)$ of the channel between the $k^{\text{th}}$ user and the BS by taking expectation of the natural logarithm of the joint probability distribution with respect to the approximate distributions of all the latent variables excluding $\mathbf{h}_k$. This computation is similar to that in Sec. 2.5.2, except that the unquantized observations $\mathbf{Z}_p$ and $\mathbf{Z}_d$ are latent variables. The final expression is as shown in (2.41).

$$\begin{aligned}\ln q\left(\mathbf{h}_k\right) \propto -\Bigg[ & \mathbf{h}_k^H \left\{ \frac{\sum_{t=1}^{\tau_p} |x_{p,kt}|^2 + \sum_{t=1}^{\tau_d} \langle |x_{d,kt}|^2 \rangle}{\sigma_w^2} \mathbf{I}_{N_r} + \mathbf{R}_k^{-1} \right\} \mathbf{h}_k \\ & - 2\Re\Bigg\{ \mathbf{h}_k^H \Bigg( \frac{1}{\sigma_w^2} \Bigg( \sum_{t=1}^{\tau_p} \Big( \langle \mathbf{z}_{p,t} \rangle - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle x_{p,k't} \Big) x_{p,kt}^* \\ & + \sum_{t=1}^{\tau_d} \Big( \langle \mathbf{z}_{d,t} \rangle - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle \langle x_{d,k't} \rangle \Big) \langle x_{d,kt}^* \rangle \Bigg) \Bigg) \Bigg\} \Bigg]. \end{aligned} \tag{2.41}$$

$$\mathbf{\Sigma}_{\mathbf{h}_k} = \left( \frac{\sum_{t=1}^{\tau_p} |x_{p,kt}|^2 + \sum_{t=1}^{\tau_d} \langle |x_{d,kt}|^2 \rangle}{\sigma_w^2} \mathbf{I}_{N_r} + \mathbf{R}_k^{-1} \right)^{-1}, \tag{2.42}$$

$$\langle \mathbf{h}_k \rangle = \frac{1}{\sigma_w^2} \mathbf{\Sigma}_{\mathbf{h}_k} \Bigg( \sum_{t=1}^{\tau_p} \Big( \langle \mathbf{z}_{p,t} \rangle - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle x_{p,k't} \Big) x_{p,kt}^* + \sum_{t=1}^{\tau_d} \Big( \langle \mathbf{z}_{d,t} \rangle - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \langle \mathbf{h}_{k'} \rangle \langle x_{d,k't} \rangle \Big) \langle x_{d,kt}^* \rangle \Bigg). \tag{2.43}$$

$$f_{k,t}(s) = -\frac{1}{\sigma_w^2} \Big( \big\langle \, \|\mathbf{h}_k\|^2 \, \big\rangle \, |s|^2 - 2\Re \Big[ \big\langle \mathbf{h}_k \big\rangle^H \Big( \big\langle \mathbf{z}_{d,t} \big\rangle - \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \big\langle \mathbf{h}_{k'} \big\rangle \big\langle x_{d,k't} \big\rangle \Big) s^* \Big] \Big) + \ln p(x_{d,kt} = s). \tag{2.44}$$

Taking the exponential on both sides of (2.41), we see that the structure is that of a complex normal distribution with covariance and mean given by (2.42) and (2.43), respectively.

#### 2.5.3.2 Computation of $q_{x_{d,kt}}(x_{d,kt})$

Following a similar approach as in Sections 2.5.1 and 2.5.2, $q_{x_{d,kt}}(x_{d,kt})$, $k = 1, \ldots, K$, $t = 1, \ldots, \tau_d$ is given by the Boltzmann distribution in (2.20), where $s$ is an $M-$QAM symbol, with $f_{k,t}(s)$ defined in (2.44). We compute the mean and mean square values of $q_{x_{d,kt}}$ in a similar manner as in Sec. 2.5.1.

#### 2.5.3.3 Computation of $q(\mathbf{z}_{d,t})$ and $q(\mathbf{z}_{p,t})$

We obtain the distribution $q(\mathbf{z}_{d,t})$ as follows:

$$\begin{aligned} \ln q\left(\mathbf{z}_{d,t}\right) &= \big\langle \ln p\left(\mathbf{Y}_p | \mathbf{Z}_p\right) + \ln p\left(\mathbf{Y}_d | \mathbf{Z}_d\right) + \ln p\left(\mathbf{Z}_p | \mathbf{X}_p, \mathbf{H}; \sigma_w^2\right) + \ln p\left(\mathbf{Z}_d | \mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) \\ &\qquad + \ln p\left(\mathbf{H} | \mathbf{R}\right) + \ln p\left(\mathbf{X}_d\right) \big\rangle \\ &\propto \left\langle \ln \mathbb{1}\big(\mathbf{z}_{d,t} \in [\mathbf{z}_{d,t}^{(\mathrm{lo})}, \mathbf{z}_{d,t}^{(\mathrm{hi})}]\big) - \frac{1}{\sigma_w^2} \left\| \mathbf{z}_{d,t} - \mathbf{H}\mathbf{x}_{d,t} \right\|^2 \right\rangle. \end{aligned}$$

Thus, we see that $q\left(\mathbf{z}_{d,t}\right)$ is a truncated complex normal distribution with mean

$$\left\langle \mathbf{z}_{d,t} \right\rangle = \boldsymbol{\mu}_{\mathbf{z}_{d,t}} + \frac{\phi\left(\frac{\mathbf{z}_{d,t}^{(\mathrm{lo})} - \boldsymbol{\mu}_{\mathbf{z}_{d,t}}}{\sigma_w/\sqrt{2}}\right) - \phi\left(\frac{\mathbf{z}_{d,t}^{(\mathrm{hi})} - \boldsymbol{\mu}_{\mathbf{z}_{d,t}}}{\sigma_w/\sqrt{2}}\right)}{\Phi\left(\frac{\mathbf{z}_{d,t}^{(\mathrm{hi})} - \boldsymbol{\mu}_{\mathbf{z}_{d,t}}}{\sigma_w/\sqrt{2}}\right) - \Phi\left(\frac{\mathbf{z}_{d,t}^{(\mathrm{lo})} - \boldsymbol{\mu}_{\mathbf{z}_{d,t}}}{\sigma_w/\sqrt{2}}\right)} \frac{\sigma_w}{\sqrt{2}}, \tag{2.45}$$

where $\boldsymbol{\mu}_{\mathbf{z}_{d,t}} = \langle \mathbf{H} \rangle \langle \mathbf{x}_{d,t} \rangle$. The division operation is a scalar element-wise division operated individually on real and imaginary components. We compute the approximate posterior distribution $q(\mathbf{z}_{p,t})$ of the unquantized received pilot symbols in the same manner as $q(\mathbf{z}_{d,t})$, and its mean is given by

$$\left\langle \mathbf{z}_{p,t} \right\rangle = \boldsymbol{\mu}_{\mathbf{z}_{p,t}} + \frac{\phi\left(\frac{\mathbf{z}_{p,t}^{(\mathrm{lo})} - \boldsymbol{\mu}_{\mathbf{z}_{p,t}}}{\sigma_w/\sqrt{2}}\right) - \phi\left(\frac{\mathbf{z}_{p,t}^{(\mathrm{hi})} - \boldsymbol{\mu}_{\mathbf{z}_{p,t}}}{\sigma_w/\sqrt{2}}\right)}{\Phi\left(\frac{\mathbf{z}_{p,t}^{(\mathrm{hi})} - \boldsymbol{\mu}_{\mathbf{z}_{p,t}}}{\sigma_w/\sqrt{2}}\right) - \Phi\left(\frac{\mathbf{z}_{p,t}^{(\mathrm{lo})} - \boldsymbol{\mu}_{\mathbf{z}_{p,t}}}{\sigma_w/\sqrt{2}}\right)} \frac{\sigma_w}{\sqrt{2}}, \tag{2.46}$$

where $\boldsymbol{\mu}_{\mathbf{z}_{p,t}} = \langle \mathbf{H} \rangle \mathbf{x}_{p,t}$. Note that computing $q(\mathbf{z}_{p,t})$ and $q(\mathbf{z}_{d,t})$ does not involve an expectation of the logarithm of integrals, as mentioned earlier.

The pseudocode for the VB procedure for the joint channel estimation and soft symbol decoding is shown in Algorithm 3. The VBI algorithm starts by randomly initializing the latent variables and maximizing the ELBO by fixing all but one hidden variable. Once the probabilities $q_{x_{d,kt}}$ are obtained, we choose the symbol with the highest probability as the decoded symbol for each $k \in \{1, \ldots, K\}$ and $t \in \{1, \ldots, \tau_d\}$ in the case of uncoded communication. For coded communication applications, it is straightforward to compute the LLRs of the data bits from $q_{x_{d,kt}}$ and pass them as inputs to the channel decoder [1]. Also, the mean of the marginal posterior of the channel can be used as a channel estimate for SNR computation, CSI feedback etc.

### 2.5.4 No CSIR and Quantized Observations

In this subsection, we extend the VB approach to the case where there is no information about the channel statistics also. This situation may arise in massive machine type communication applications where a large number of low power UEs sporadically wake

**Algorithm 3:** QVB Joint Channel Estimation and Soft Symbol Decoding with Statistical CSIR

**Input:** $\mathbf{Y}_p, \mathbf{Y}_d, \{\mathbf{R}_1, \ldots, \mathbf{R}_K\}, \mathbf{X}_p, M, \mathbb{M} = \{s_1, \ldots, s_M\}$, $\tau_p$, $\tau_d$, $\sigma_w$

**Output:** $\{\langle \mathbf{h}_1 \rangle, \ldots, \langle \mathbf{h}_K \rangle\}, q_{x_{d,kt}}, \langle x_{d,kt} \rangle \, \forall k \in [K], t \in [\tau_d]$

Initialize $\langle \mathbf{Z}_p \rangle, \langle \mathbf{Z}_d \rangle, q_{x_{d,kt}}(x_{d,kt}), \langle x_{d,kt} \rangle = 0, \forall k \in [K], t \in [\tau_d]$
**repeat**
  **for** $k = 1$ to $K$, $t = 1$ to $\tau_d$ **do**
    **for** $m = 1$ to $M$ **do**
      Compute $q_{x_{d,kt}}(s_m)$ using (2.20) and (2.44).
    **end for**
    $\langle x_{d,kt} \rangle = \sum_{s \in \mathbb{M}} s \, q_{x_{d,kt}}(s)$.
    $\langle |x_{d,kt}|^2 \rangle = \sum_{s \in \mathbb{M}} |s|^2 \, q_{x_{d,kt}}(s)$.
  **end for**
  **for** $k = 1$ to $K$ **do**
    Compute $\mathbf{\Sigma}_{\mathbf{h}_k}$ and $\langle \mathbf{h}_k \rangle$ using (2.42) and (2.43), respectively.
  **end for**
  **for** $t = 1$ to $\tau_p$ **do**
    Compute $\langle \mathbf{z}_{p,t} \rangle$ using (2.46).
  **end for**
  **for** $t = 1$ to $\tau_d$ **do**
    Compute $\langle \mathbf{z}_{d,t} \rangle$ using (2.45).
  **end for**
**until** a stopping condition is met.

up and transmit data to the BS or AP. In such scenarios, the assumption of knowledge of channel statistics at the BS or AP may not be appropriate. We assume that the channels between a UE and the different antennas at the BS are uncorrelated and that the LSFCs are the same across all BS antennas. We impose a non-informative conjugate Gamma prior for the inverse of the LSFCs. Let $\alpha_k$ denote the inverse of the LSFC between the

$k^{\text{th}}$ UE and the BS and let $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_K]^T$. The prior on $\alpha_k$ is given by

$$p(\alpha_k; a, r) = \frac{r^a}{\Gamma(a)} \alpha_k^{a-1} \exp(-r\alpha_k), \tag{2.47}$$

where $a$ and $r$ are the parameters of the Gamma distribution. Conditioned on $\alpha_k$, the channel $\mathbf{h}_k$ of the $k^{\text{th}}$ UE is distributed as $\mathcal{CN}(\mathbf{0}, \alpha_k^{-1}\mathbf{I}_{N_r})$. The joint distribution of the observed and the latent variables is given by

$$\begin{aligned} &p\left(\mathbf{Y}_p, \mathbf{Y}_d, \mathbf{Z}_p, \mathbf{Z}_d, \mathbf{H}, \mathbf{X}_d, \boldsymbol{\alpha}; \mathbf{X}_p, \sigma_w^2, a, r\right) \\ &\quad = p\left(\mathbf{Y}_p|\mathbf{Z}_p\right) p\left(\mathbf{Y}_d|\mathbf{Z}_d\right) p\left(\mathbf{Z}_p|\mathbf{X}_p, \mathbf{H}; \sigma_w^2\right) p\left(\mathbf{Z}_d|\mathbf{X}_d, \mathbf{H}; \sigma_w^2\right) p\left(\mathbf{H}|\boldsymbol{\alpha}\right) p\left(\mathbf{X}_d\right) p\left(\boldsymbol{\alpha}; a, r\right). \end{aligned} \tag{2.48}$$

We impose a factorized structure on the posterior distribution:

$$p\left(\mathbf{Z}_p, \mathbf{Z}_d, \mathbf{X}_d, \mathbf{H}, \boldsymbol{\alpha}|\mathbf{Y}_p, \mathbf{Y}_d; \mathbf{X}_p, \sigma_w^2, a, r\right) \approx q\left(\mathbf{Z}_p\right) q\left(\mathbf{Z}_d\right) q\left(\mathbf{X}_d\right) q\left(\mathbf{H}\right) q\left(\boldsymbol{\alpha}\right), \tag{2.49}$$

where

$$q\left(\mathbf{H}\right) = \prod_{k=1}^{K} q\left(\mathbf{h}_k\right), \qquad q\left(\mathbf{X}_d\right) = \prod_{k=1}^{K} \prod_{t=1}^{\tau_d} q\left(x_{d,kt}\right), \tag{2.50}$$

$$q\left(\mathbf{Z}_d\right) = \prod_{t=1}^{\tau_d} q\left(\mathbf{z}_{d,t}\right), \qquad q\left(\mathbf{Z}_p\right) = \prod_{t=1}^{\tau_p} q\left(\mathbf{z}_{p,t}\right), \qquad q\left(\boldsymbol{\alpha}\right) = \prod_{k=1}^{K} q\left(\alpha_k\right). \tag{2.51}$$

The computation of $q(\mathbf{Z}_p)$, $q(\mathbf{Z}_d)$, $q(\mathbf{H})$ and $q(\mathbf{X}_d)$ is the same as in Sec. 2.5.3. We obtain $q(\boldsymbol{\alpha})$ by taking the expectation of the logarithm of the joint distribution in (2.48) with respect to all the latent variables except $\boldsymbol{\alpha}$ to get

$$\ln q(\alpha_k) \propto (a + N_r - 1) \ln \alpha_k - \alpha_k (r + \langle \|\mathbf{h}_k\|^2 \rangle). \tag{2.52}$$

Taking the exponential on both sides, we observe that $q(\alpha_k)$ is Gamma distributed with

**Algorithm 4:** QVB Joint Channel Estimation and Soft Symbol Decoding with No CSIR

**Input:** $\mathbf{Y}_p, \mathbf{Y}_d, \mathbf{X}_p, M, \mathbb{M} = \{s_1, \ldots, s_M\}, \tau_p, \tau_d, \sigma_w, a, r, N_r$

**Output:** $\{\langle \mathbf{h}_1 \rangle, \ldots, \langle \mathbf{h}_K \rangle\}, q_{x_{d,kt}}, \langle x_{d,kt} \rangle \, \forall k \in [K], t \in [\tau_d]$

1: Initialize $\langle \mathbf{Z}_p \rangle, \langle \mathbf{Z}_d \rangle, q_{x_{d,kt}}(x_{d,kt}), \langle x_{d,kt} \rangle = 0, \forall k \in [K], t \in [\tau_d], \{\langle \alpha_1 \rangle, \ldots, \langle \alpha_K \rangle\}$

2: **repeat**

3: Computation of $q_{x_{d,kt}}, \langle x_{d,kt} \rangle, \langle |x_{d,kt}|^2 \rangle$: Follow steps 3 to 8 of Algorithm 3.

4: **for** $k = 1$ to $K$ **do**

5: Compute $\mathbf{\Sigma}_{\mathbf{h}_k} = \left( \frac{\sum_{t=1}^{\tau_p} |x_{p,kt}|^2 + \sum_{t=1}^{\tau_d} \langle |x_{d,kt}|^2 \rangle}{\sigma_w^2} + \langle \alpha_k \rangle \right)^{-1} \mathbf{I}_{N_r}$.

6: Compute $\langle \mathbf{h}_k \rangle$ using (2.43).

7: Compute $\langle \|\mathbf{h}_k\|^2 \rangle = \text{Trace}(\mathbf{\Sigma}_{\mathbf{h}_k}) + \|\langle \mathbf{h}_k \rangle\|^2$ and $\langle \alpha_k \rangle = \frac{a+N_r}{r+\langle \|\mathbf{h}_k\|^2 \rangle}$.

8: **end for**

9: Computation of $\langle \mathbf{z}_{p,t} \rangle, \langle \mathbf{z}_{d,t} \rangle \, \forall t$: Follow steps 13 to 18 of Algorithm 3.

10: **until** a stopping condition is met.

shape and rate parameters $(a + N_r)$ and $(r + \langle \|\mathbf{h}_k\|^2 \rangle)$, respectively. The mean of $q(\alpha_k)$ is

$$\langle \alpha_k \rangle = \frac{a + N_r}{r + \langle \|\mathbf{h}_k\|^2 \rangle}. \tag{2.53}$$

If we denote the LSFC of the $k^{\text{th}}$ UE by $\beta_k$, then $\langle \beta_k \rangle = 1/\langle \alpha_k \rangle = (r + \langle \|\mathbf{h}_k\|^2 \rangle)/] \, (a + N_r)$. We set $a = 0$ and $r = 10^{-4}$ in our simulations. The pseudocode for joint channel estimation and soft symbol decoding with no CSIR is similar to Algorithm 3; we highlight the changes in Algorithm 4. Note that the QVB with no CSIR algorithm does not have any matrix inverse operations, which makes it computationally attractive. Also, the estimation of $\langle \alpha_k \rangle$ involves only $K$ scalar divisions and does not add much complexity to the procedure.

### 2.5.5 VB-BP Receiver

In order to improve the BER of a communication system, a typical engineering approach is to employ an iterative receiver architecture where the posterior beliefs of the output bits of the channel decoder (belief propagation (BP) decoder in our case) are fed back to the detector block. However, in commercial wireless systems, the stringent time requirements to complete the detection and channel decoding task within the transmission interval needs fast detection and decoding. Our VB approach makes it convenient to implement such a receiver without additional complexity. Instead of running the VB algorithm till convergence, we execute only one iteration of Algorithms 2, 3 or 4 and feed the LLRs obtained to the channel decoder. The BP decoder also runs only one iteration and outputs the aposteriori LLRs that are used to compute the extrinsic information.[4] We use this to adapt the data prior probability distribution in the next VB iteration. This approach is also guaranteed to converge as the VB converges to a stationary point from any initialization, and BP also converges if the associated factor graph is cycle-free. We illustrate this using an LDPC code with a cycle-free sparse parity check matrix. We find that the approach converges fast and also leads to a performance improvement of around 0.5 dB compared to performing soft symbol estimation and channel decoding one after the other. Hence, we use this receiver architecture in our simulations.

### 2.5.6 Computational Complexity

In this subsection, we analyze the per-iteration computational complexity of the variational Bayesian algorithms. We provide the complexity of the VB algorithm with statistical CSIR and quantized observations, as the algorithms for the perfect CSIR and unquantized observations are special cases with lower complexity. Table 2.1 shows the order ($\mathcal{O}$) of the per-iteration computational complexity of the steps involved in one iteration of the VB algorithm. The per-iteration complexity of the overall algorithm scales cubically with the number of rx antennas, quadratically with the number of users, and

[4]The extrinsic information is obtained by subtracting the LLRs output by the VB receiver from the LLRs output by the channel decoder.

Table 2.1:

Per-iteration Complexity of the QVB Algorithm 3

| **Matrix** | **Order Complexity** |
|---|---|
| $\langle \mathbf{X}_d \rangle$ | $MK^2N_r\tau_d$ |
| $\{\mathbf{\Sigma}_{\mathbf{h}_k}\}_{k=1}^K$ | $K(N_r^3 + \tau_d)$ |
| $\langle \mathbf{H} \rangle$ | $KN_r^2 + K^2N_r(\tau_p + \tau_d)$ |
| $\langle \mathbf{Z}_p \rangle$ | $KN_r\tau_p$ |
| $\langle \mathbf{Z}_d \rangle$ | $KN_r\tau_d$ |

linearly with the constellation size and number of pilot and data symbols. In particular, the complexity is linear in the number of data symbols $\tau_d$, unlike maximum likelihood approaches where the complexity grows exponentially with $\tau_d$. The total time taken by VB algorithm is low due to its fast and guaranteed convergence to a local optimum. In our experiments, we find that the VB algorithm typically converges within 16 iterations, and the improvement from running further iterations is negligible.

## 2.6 Simulation Results

In this section, we evaluate the normalized mean square error (NMSE) in channel estimation and the data bit error rate (BER) of the VBI algorithms in an uplink massive MIMO wireless communication system with low-resolution ADCs at the BS. We use an LDPC channel code from 3GPP 5G NR specifications [103]. We use the parity check matrix from LDPC base graph 0 with a lifting size $Z_c$ set to 8 and set index 0, which results in 176 message bits and 544 coded bits per block. We set $N_r = \{100, 200\}, K = \{25, 50\}$. We vary $\tau_d$ from 100 to 450 and set the ADC resolution to 3 bits. We also evaluate the performance with pilot power boosting in which pilots are transmitted at a slightly higher power level to improve the channel estimation (see Fig. 2.11). Throughout this section,

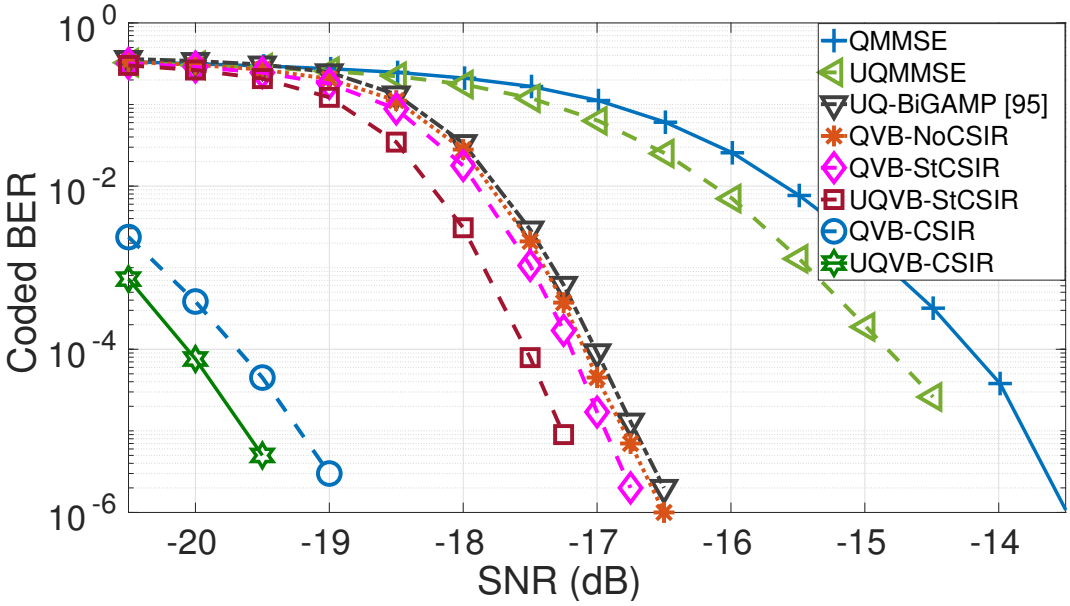


Figure 2.3: Coded BER vs. SNR (dB), with $N_r = 200$, $K = 50$, $\tau_d = 450$, $\tau_p = 50$, and 3 bits quantization.

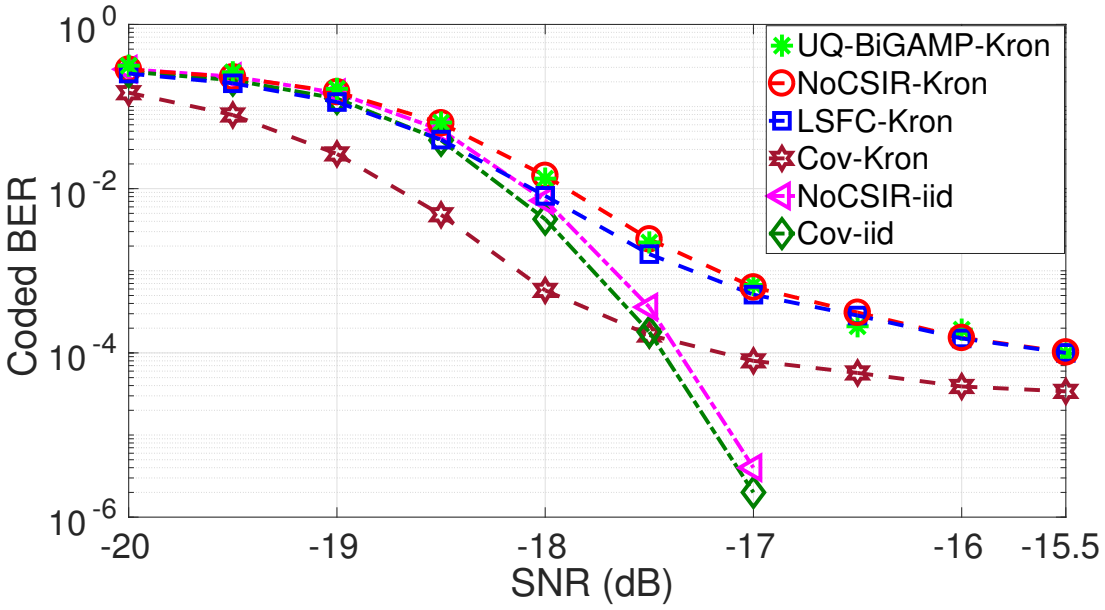


Figure 2.4: Coded BER vs. SNR (dB), with $N_r = 200$, $K = 50$, $\tau_d = 450$, $\tau_p = 50$, and 3 bits quantization.

the data symbols are drawn i.i.d. from a 4-QAM constellation with unit energy. With higher order constellations, the performance is similar, with an expected shift in the SNR

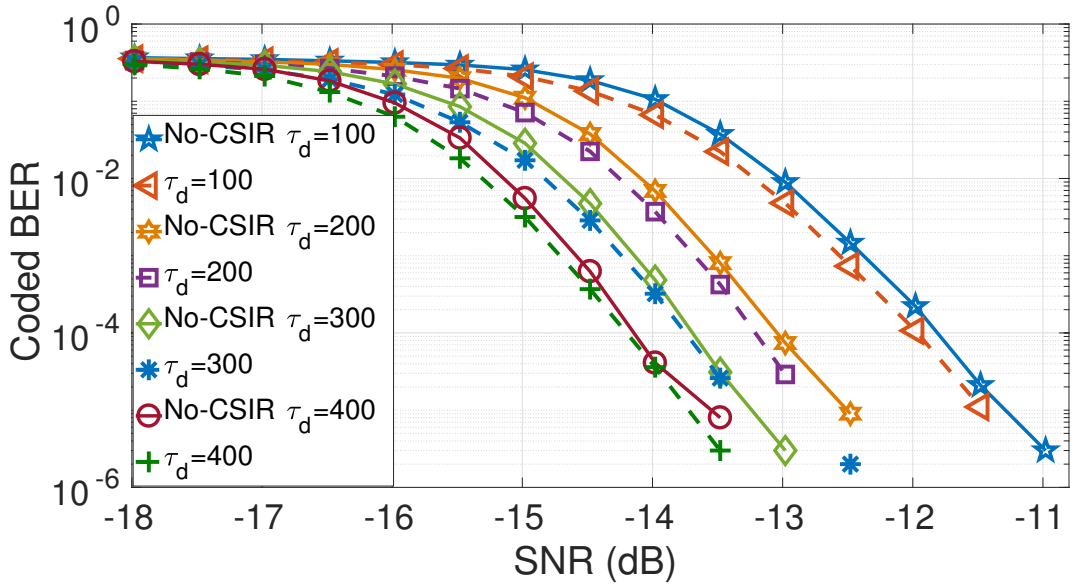


Figure 2.5: Coded BER vs. SNR (dB), with $N_r = 100$, $K = 25$, $\tau_p = 25$, 3 bits quantization; comparing QVB with statistical CSIR and no CSIR when the channel correlation matrix is set to $\mathbf{I}_{N_r}$.

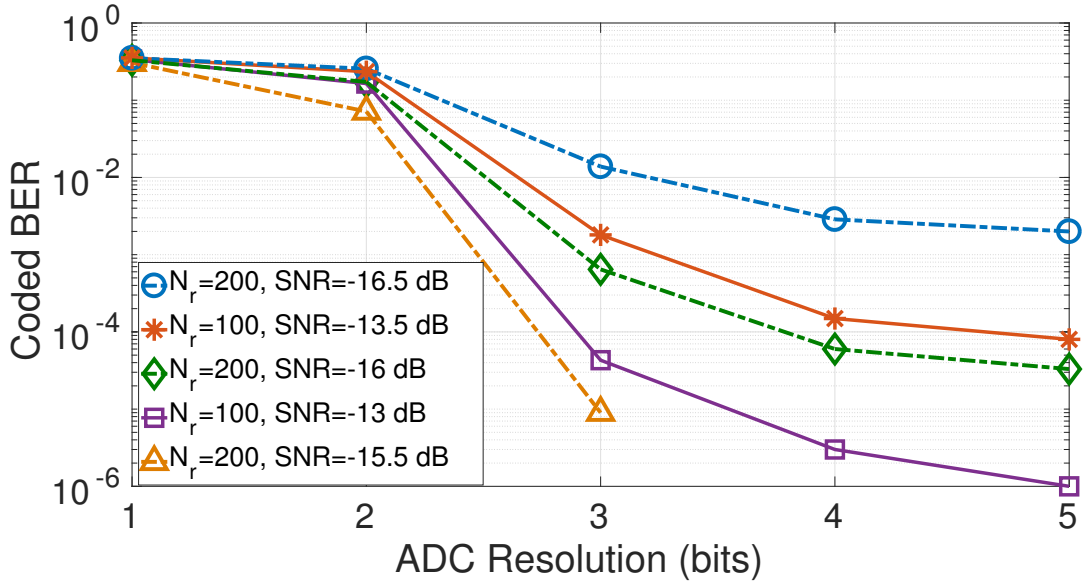


Figure 2.6: Coded BER vs. ADC resolution (bits), with $K = 25$, $\tau_d = 100$, and $\mathbf{R}_k = \mathbf{I}_{N_r}$ $\forall k$.

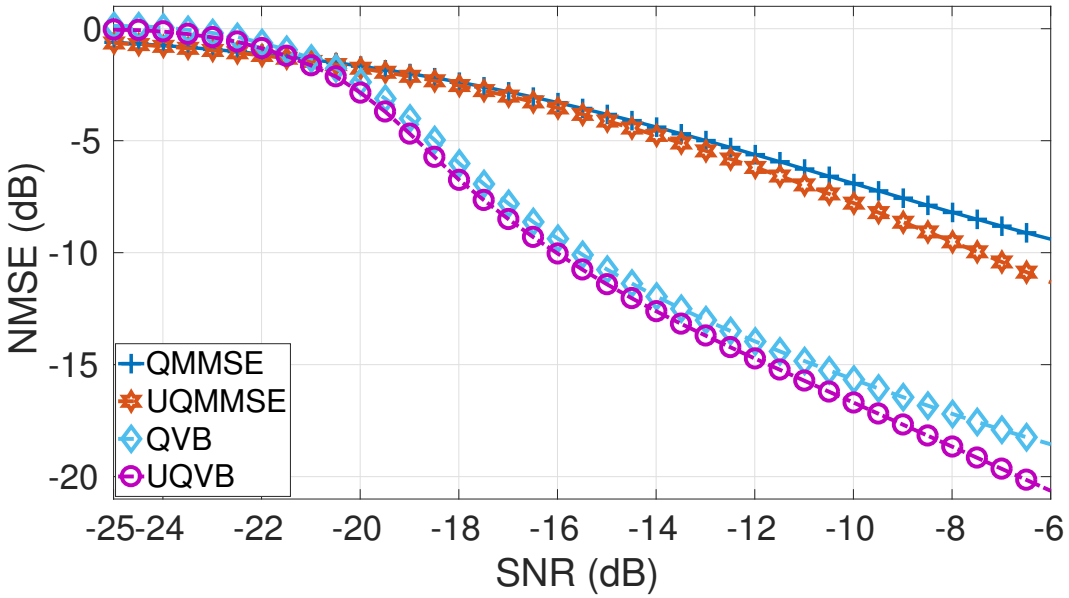


Figure 2.7: NMSE vs. SNR (dB) for $N_r = 200$, $K = 50$, $\tau_d = 450$, $\tau_p = 50$, 3 bits quantization.

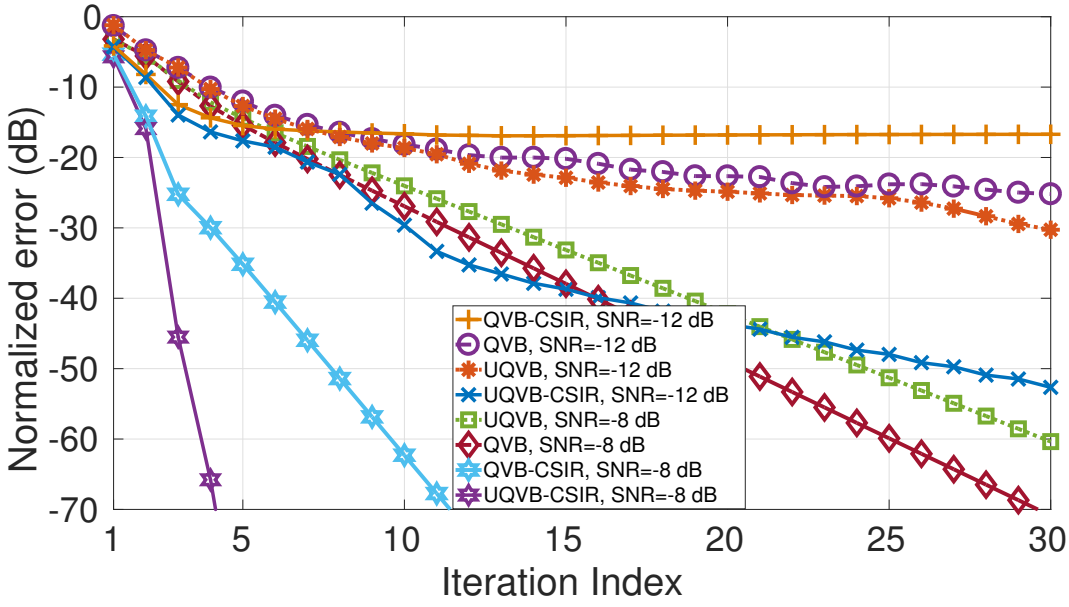


Figure 2.8: Convergence of VB algorithm. $N_r = 100$, $K = 25$, $\tau_d = 200$, 3 bits quantization.

required to achieve a given BER. Except in Fig. 2.8, we fix the maximum number of iterations to 16, as the BER improvement beyond 16 iterations is marginal.

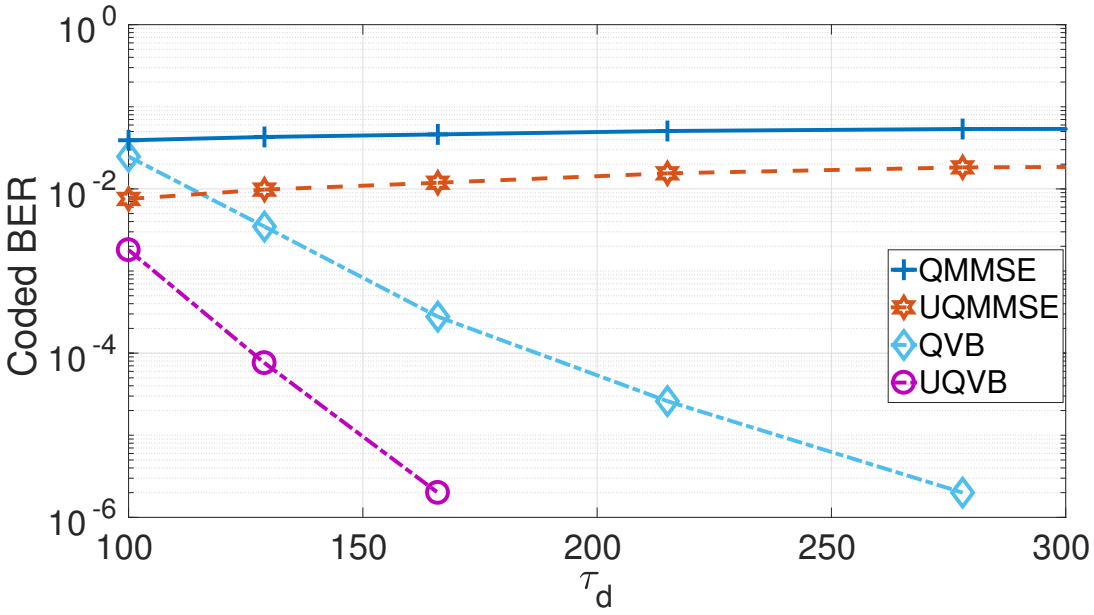


Figure 2.9: Coded BER vs. $\tau_d$ for $N_r = 100$, $K = 25$, SNR=$-13.5$ dB, $\tau_p = 25$, and 3 bits quantization.

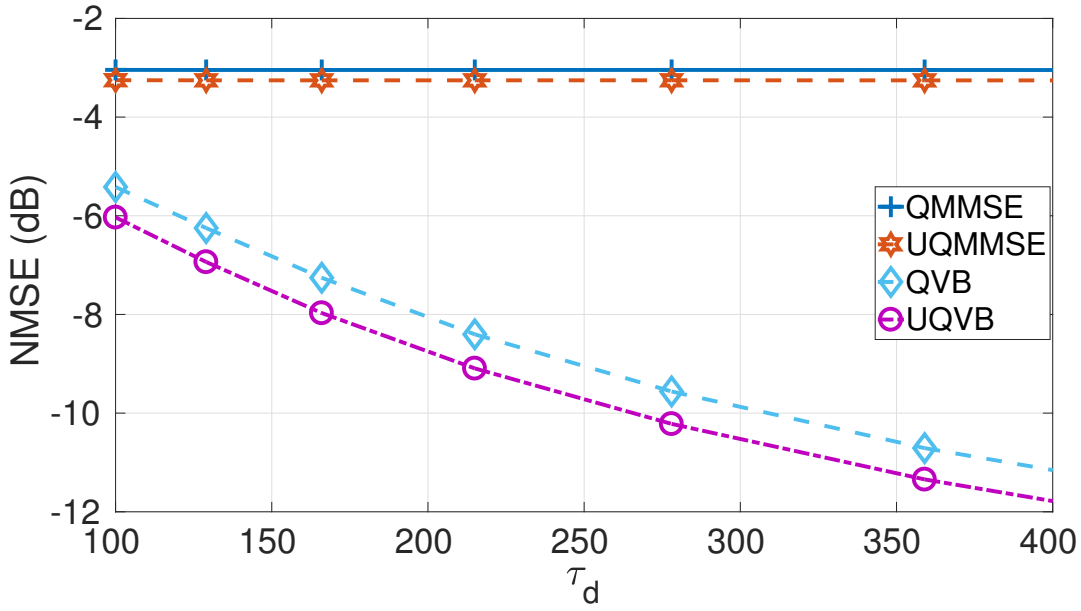


Figure 2.10: NMSE vs. $\tau_d$ for $N_r = 100$, $K = 25$, SNR=$-13.5$ dB, $\tau_p = 25$, and 3 bits quantization.

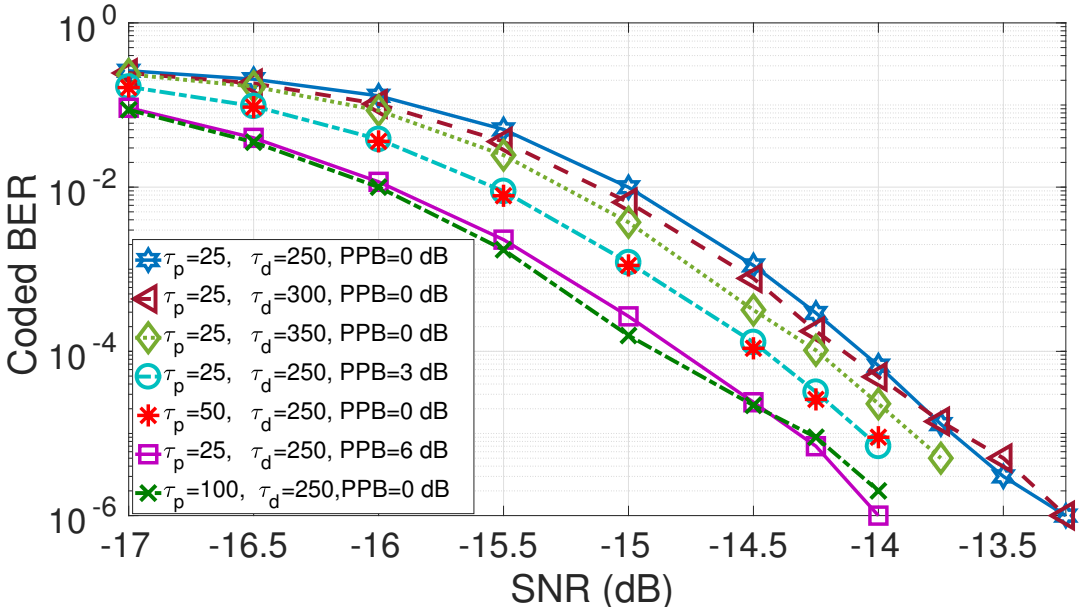


Figure 2.11: Coded BER vs. SNR (dB), with $N_r = 100$, $K = 25$, 3 bits quantization, pilot power boosting and increased $\tau_p$.

The channel coefficient between the $k^{\text{th}}$ UE and the $n^{\text{th}}$ rx antenna at the BS is denoted by $h_{nk}$, and $\mathbf{h}_k \triangleq [h_{1k}, h_{2k}, \ldots, h_{N_r k}]^T$. We assume that the LSFCs between the $k^{\text{th}}$ UE and each antenna at the BS are the same due to the close spacing between the BS antennas. We also assume that the transmit antennas at each UE are spatially uncorrelated. We model the spatial correlation between the rx antennas at the BS using a Kronecker spatial fading correlation model [104–106]. The channel vector $\mathbf{h}_k$ is expressed as $\mathbf{h}_k = \mathbf{R}_k^{\frac{1}{2}} \mathbf{h}_k^{\text{iid}}$, where $\mathbf{h}_k^{\text{iid}}$ is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$, and $\mathbf{R}_k \in \mathbb{C}^{N_r \times N_r}$ is the spatial correlation matrix of the $k^{\text{th}}$ UE's channel. The $(m,n)^{\text{th}}$ element of $\mathbf{R}_k$ is given in (2.54),

$$[\boldsymbol{R}_k]_{mn} = \frac{\beta_k}{\sqrt{v_k}} \exp\left( -\frac{1}{2v_k} \left[ a_k^2 \cos^2(\theta_k^{\text{az}}) - 2jc_k \cos(\theta_k^{\text{az}}) + \nu_k^{\text{az}} c_k^2 \sin^2(\theta_k^{\text{az}}) \right] \right), \tag{2.54}$$

where $\beta_k$ is the LSFC of the $k^{\text{th}}$ UE, $a_k = \frac{2\pi d}{\lambda} \sqrt{\nu_k^{\text{el}}} (n-m) \cos\left(\theta_k^{\text{el}}\right)$, $v_k = 1 + \nu_k^{\text{az}} a^2 \sin^2(\theta_k^{\text{az}})$ and $c_k = \frac{2\pi d}{\lambda} (n-m) \sin\left(\theta_k^{\text{el}}\right)$. Here, $\lambda$ is the carrier wavelength, $d$ is the antenna spacing; $\theta_k^{\text{az}}$ and $\theta_k^{\text{el}}$ are the means of horizontal angle of arrival (AoA) and vertical AoA, respectively; $\nu_k^{\text{az}}$ and $\nu_k^{\text{el}}$ are the variances of horizontal AoA and vertical AoA, respectively.

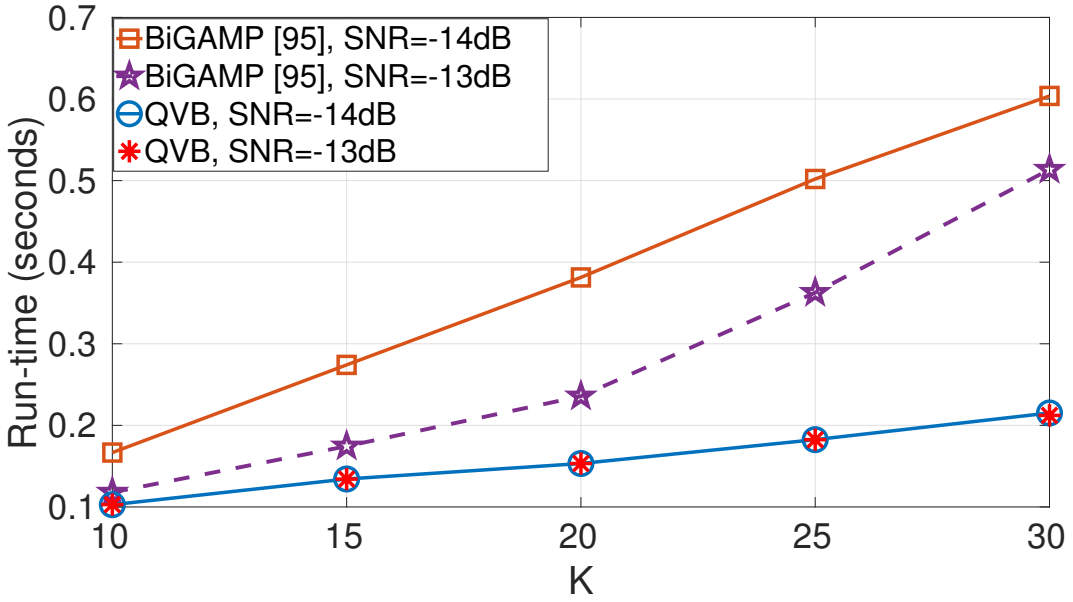


Figure 2.12: Run-time ($s$) vs. $K$ for $N_r = 100$, $\tau_d = 250$, 3 bits quantization.

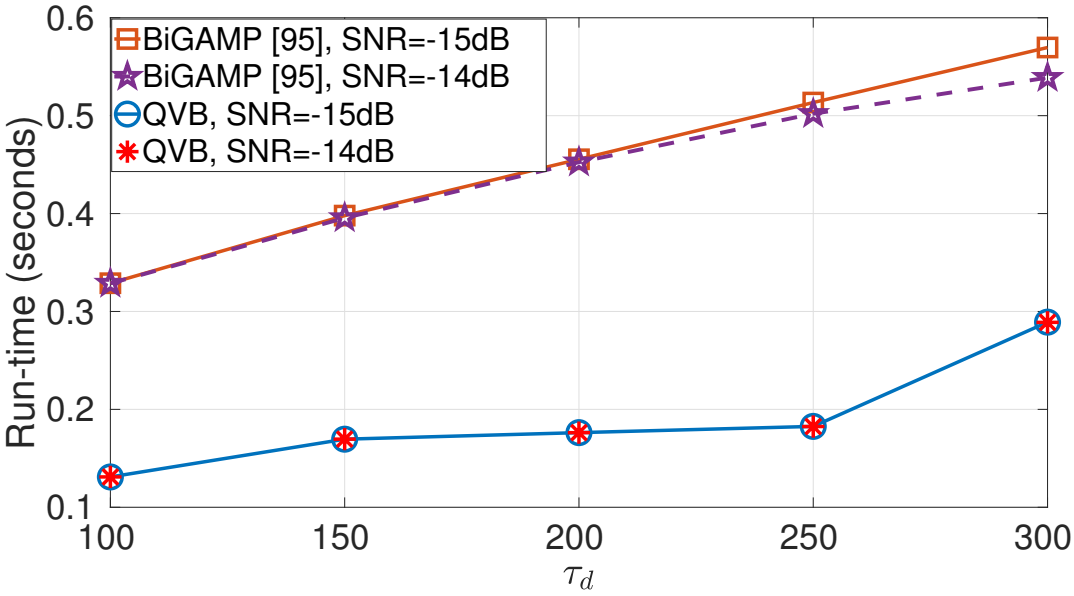


Figure 2.13: Run-time ($s$) vs. $\tau_d$ for $N_r = 100$, $K = 25$, 3 bits quantization.

We consider a uniform linear antenna array with the spacing between its elements set to $\lambda$. We set the mean and standard deviation of the horizontal and vertical AoA to be uniformly distributed between $-\pi/3$ and $\pi/3$ radians, and 0 and $\pi/6$ radians, respectively. The UEs adopt path loss inversion based transmit power control that compensates for the LSFCs,[5] i.e., the diagonal entries of the channel covariance matrix are close to 1 [107, 108]. Without power control, if the LSFCs of the UEs' channels are very different, the UEs that have small LSFCs will suffer high quantization noise. This leads to low signal to quantization noise ratio and therefore poor BER. Taking into account possibly imperfect power control, we assume that the diagonal entries of the channel covariance matrix are uniformly distributed between 1 and 1.2. For the uncorrelated case, we set the channel covariance to be $\mathbf{I}_{N_r}$.

We benchmark the coded BER and NMSE performance of the quantized VB algorithm with that of a genie channel aided unquantized VB algorithm, an unquantized BiGAMP based joint channel and data estimator [95] and a unquantized linear MMSE soft-decoder [1]. We set the maximum number of iterations for the BiGAMP receiver to 500. The unquantized VB algorithm with perfect CSIR serves as a lower bound for the BER of the quantized algorithm. We note that there are variants of MMSE decoder such as MMSE-SIC, but their computational complexities are much higher due to the matrix inversions involved in every iteration of SIC [98]. Hence, we do not compare against these methods.

Figure 2.3 shows the *coded* BER when SNR (dB) is varied for $N_r = 200$, $K = 50$, $\tau_d = 450$ and 3 bits quantization when the channel covariance matrix is set to $\mathbf{I}_{N_r}$. We compare the performance of the VB algorithm with perfect CSIR (labeled UQVB-CSIR, QVB-CSIR for the unquantized and quantized cases, respectively) with the performance with statistical CSIR (labeled UQVB-StCSIR, QVB-StCSIR in the unquantized and quantized cases, respectively) and no CSIR case (labeled QVB-NoCSIR).

There are three groups of curves. The best performance is achieved by the genie-aided receivers that have perfect CSIR (QVB-CSIR and UQVB-CSIR). With 3-bit quantization in the ADCs, the gap between QVB-CSIR and UQVB-CSIR is less than 0.4 dB. The next

[5]The LFSCs can be estimated at the UEs, for example, using the primary synchronization signals that are periodically transmitted by the BS.

set of curves correspond to the VB algorithm with statistical CSIR and no CSIR. There is almost no performance gap between the quantized VB algorithm with and without statistical CSIR. Further, as in the perfect CSIR case, the loss due to 3-bit quantization compared to the unquantized VB algorithm is less than 0.4 dB. We also note that the performance gap between the first and second set of curves is around 2.5 dB, in both the quantized and unquantized cases. This illustrates that it is important to account for the effect of channel estimation errors, in order to realistically estimate the performance. The performance gap can be reduced by employing pilot power boosting or longer pilot sequences. UQVB-StCSIR outperforms unquantized BiGAMP (labeled UQ-BiGAMP) by more than 0.5 dB. Moreover, with only 3 bits quantization, both QVB-StCSIR and QVB-NoCSIR marginally outperform UQ-BiGAMP. Finally, the worst performance is achieved by the MMSE based receivers [1], with the gap between the quantized MMSE receiver and the quantized VB algorithm with no CSIR being nearly 3 dB at a BER of $10^{-4}$. The significantly better performance achieved by the VB algorithms is clear from the plot. In addition, the gap between the quantized MMSE and the unquantized MMSE receivers shows that linear receivers can result in suboptimal performance, even if the channel state is made available to the receiver. For the quantized and unquantized MMSE detectors, the channels are estimated using the quantized and unquantized pilot received signals, respectively.

Figure 2.4 shows the BER vs. SNR performance of the VB algorithm with correlated channels with and without CSIR (curves labeled Kron). We also show the performance under i.i.d. channels (curves labeled iid). We observe that, in a spatially correlated scenario, the algorithm that has the knowledge of the channel covariance matrices (Cov-Kron) performs around 1 dB better at a BER of $10^{-3}$ than the algorithms that do not have the knowledge of the channel covariance matrices (NoCSIR-Kron and LSFC-Kron). Note that the VB algorithm with NoCSIR assumes i.i.d. channels, which results in a degraded performance under spatially correlated scenarios. This shows the importance of utilizing the correlation information when designing receiver schemes. We observe that the performance of the VB algorithm when the BS has the knowledge of only the LSFCs

of the UEs' channels (LSFC-Kron) is poor compared to Cov-Kronecker. Thus, utilizing only the knowledge of LSFCs is not sufficient, and the full correlation information is necessary to obtain better performance. Also, under an i.i.d. channel scenario, VB with NoCSIR (NoCSIR-iid) performs very close to VB with complete knowledge of channel covariance matrices (Cov-iid). This shows that VB with NoCSIR is able to estimate the LSFCs accurately with very little additional computational complexity. We see an error floor in the spatially correlated case at high SNRs. This is because, the channel AoAs are spread within a narrow range (standard deviation of $\pi/6$), which results in non-negligible multiuser interference at high SNRs. We also observe that, in the spatially correlated scenario, the unquantized BiGAMP based receiver (labeled UQ-BiGAMP-Kron) performs about 1 dB worse than Cov-Kron with 3 bit ADCs, and marginally worse than LSFC-Kron with 3 bit ADCs. The BiGAMP receiver assumes i.i.d. channels but has the knowledge of the LSFCs. Finally, the NoCSIR-Kron curve, which is obtained without knowledge of the channel covariance matrices or the LSFCs, but 3 bits quantization, performs the same as UQ-BiGAMP-Kron.

Figure 2.5 compares the BER of the QVB algorithms with and without the knowledge of statistical CSIR for various values of $\tau_d$, for uncorrelated channels ($\mathbf{R}_k = \mathbf{I}_{N_r}$). The performance of QVB with no CSIR is only marginally worse than that of QVB with statistical CSIR. Also, we see a roughly 10-fold improvement BER when the number of data symbols is doubled. This shows that the VB algorithm is able to effectively use the data symbols to improve the channel estimates. We recall that the computational complexity of the receiver for uncorrelated channels is lower than the correlated channel case, as the channel covariance matrix is diagonal and all the matrix operations can be computed using scalar computations.

Figure 2.6 shows the BER vs. ADC resolution (in bits) for various rx antenna and SNR configurations. We set $\tau_d = 100$, $K = 25$ and the spatial correlation matrices to $\mathbf{I}_{N_r}$. The BER improves as the ADC resolution increases, as expected, but the slope of the BER curve decreases and becomes almost 0 beyond $4-$bits resolution. This illustrates that low-resolution ADCs are relevant in wireless communication systems, especially in

massive MIMO systems where the number of antennas are large compared to the number of users. For example, in order to achieve a BER of $10^{-3}$, UEs have to expend twice the transmit power if the BS is equipped with 100 antennas and 3 bit ADCs compared to a system with 200 antennas at the BS and 5 bit ADCs. Thus, a 3 dB higher transmit power at the UEs can lead to significant power savings at the BS. On the other hand, if a UE does not have the power budget to increase its transmit power, using additional antennas at the BS can provide the required rx antenna gain. Such tradeoff analyses can be used by a system designer to configure the system parameters based on the bit budget and power constraints at the BS and UEs.

We now turn to the channel estimation performance of the VB algorithms. Figure 2.7 shows the NMSE of channel estimation as a function of the SNR (dB), for $N_r = 200$, $K = 50$, $\tau_d = 450$ and 3 bits quantization. The VBI algorithms *with quantized observations* provide around 8 dB improvement at an NMSE of $-10$ dB compared to the MMSE estimation based on *unquantized observations.* We also see that the NMSE of the unquantized VB (UQVB) is almost the same as the 3-bit quantized VB (QVB) algorithm. This is because the VBI algorithms refine the channel estimates based on the posterior beliefs of the data symbols. This feature can be directly translated to a reduction in the training overhead required in massive MIMO systems with low-resolution ADCs, and thereby improve the achievable spectral efficiency.

Figure 2.8 shows the convergence behavior of the VB algorithms, with $N_r = 100$, $K = 25$, $\tau_d = 200$ for both unquantized and 3 bits quantization cases. The convergence behaviors for both the unknown CSIR and genie aided case (with legend suffixed with `CSIR`) are shown. We use the means of the estimated data symbols and channel to compute the normalized error for the perfect CSIR and the unknown CSIR cases, respectively. We see that the VBI algorithms converge to a normalized error below $-20$ dB within about 20 iterations. This illustrates that the developed VB algorithms are of polynomial complexity with fast convergence.

Figures 2.9 and 2.10 show the coded BER and NMSE performance of the VB algorithms, respectively, with $N_r = 100$, $K = 25$, SNR $= -13.5$ dB and 3 bits quantization,

as a function of the data duration $\tau_d$. As $\tau_d$ increases, the BER and NMSE of the VBI algorithms improves, unlike the MMSE receivers. Again, this is because the VBI algorithm uses the posterior beliefs of the data symbols to refine its channel estimates, which in turn improves the quality of the posterior beliefs of the data symbols. Therefore, the performance can be dramatically improved by jointly decoding a larger number of data symbols (up to the coherence time of the channel), leading to a reduced training overhead even in the presence of low-resolution ADCs.

Figure 2.11 compares the BER across different algorithms when the pilot transmit power is boosted (PPB) by a certain amount above that of the data transmit power, and with $\tau_p$ set to $K$, $2K$ and $4K$. We see that, as the pilot duration increases, the BS is able to estimate the channels better, resulting in improved performance. A similar performance improvement occurs with PPB. For example, PPB of 3 dB results in nearly the same performance as the case when the pilot duration is doubled. We also show the BER when $\tau_d = \{250, 300, 350\}$. If the coherence interval is large, the VBI approach provides a system designer the option to avoid PPB or increasing $\tau_p$ while still meeting the QoS requirements.

Figures 2.12 and 2.13 compare the average run times of the QVB algorithm based on quantized observations with that of unquantized BiGAMP procedure for various values of the number of users $K$ and the data duration $\tau_d$, respectively. The simulations were executed using MATLAB R2020b in an Intel core i7, 3 GHz $\times$ 8 CPU with 64 GB RAM running an Ubuntu 18.04 LTS 64 bits operating system. We use the normalized mean squared difference in the channel estimate between two successive iterations as the convergence metric, and set it to $10^{-5}$. We see that the total run time taken by quantized VB algorithm is far less than that of unquantized BiGAMP, even though the per iteration complexity of BiGAMP scales linearly with system dimensions. This shows that our quantized VB-BP based joint channel estimation and soft symbol decoding not only performs better than unquantized BiGAMP, but is also faster.

## 2.7 Conclusions

In this chapter, we considered joint channel estimation and soft symbol decoding in a single carrier uplink massive multiple input multiple output (MIMO) receiver with low-resolution ADCs. We presented a novel, low-complexity VB procedure that directly outputs the posterior beliefs of the data symbols. The channel estimates obtained can potentially be used for signal to interference noise (SINR) computation and link adaptation. We provided a flexible approach to integrate the VB receiver with a BP channel decoder via extrinsic information feedback. We evaluated the coded data BER and the NMSE in the channel estimates obtained by our algorithm using Monte Carlo simulations and benchmarked it against the state-of-the-art receivers.

# Chapter 3

# Massive MIMO-OFDM Systems with Low Resolution ADCs: Cramér-Rao Lower Bound, Sparse Channel Estimation, and Soft Symbol Decoding

## 3.1 Introduction

Recent research in wireless communications has investigated the use of a massive number of antennas at the base station (BS) to increase the network capacity and data rates [77]. While the benefits of massive multiple input multiple output (MIMO) communications are now very well understood, they come at the expense of high power consumption and hardware cost, which needs to be addressed to make it commercially viable. One potential solution is to employ low-resolution analog-to-digital converters (ADCs) in the receivers [8, 82, 85]. The power consumption of an ADC increases exponentially with its bit-width. Hence, in massive MIMO systems with tens or hundreds of antennas and one RF chain per antenna, employing low-resolution ADCs can result in dramatic power savings [79,80].

Further, low-resolution ADCs relax the stringent linearity range requirements on the RF circuitry, which in turn reduces the hardware cost [109]. However, they also bring new challenges in the design of receivers, as advanced signal processing techniques need to be used to counter the large quantization noise introduced by them. This chapter investigates several key aspects of receiver design, and presents novel receiver architectures in the context of multiuser massive MIMO orthogonal frequency division multiplexing (OFDM) communication with low-resolution ADCs.

Three main challenges arise in the use of low-resolution ADCs in multi-user MIMO-OFDM systems. First, the non-linearities introduced by coarse quantizers lead to sub-optimal performance of conventional receivers such as (regularized) zero-forcing (ZF/RZF) and minimum mean square error (MMSE) detectors [68]. In a conventional OFDM receiver, we remove the cyclic prefix (CP), decouple the subcarriers using a discrete Fourier transform (DFT), and perform frequency domain equalization on a per-subcarrier basis. However, in low-resolution ADC based systems, we obtain the complex baseband time-domain samples after being coarsely quantized by the ADC, and it is not possible to decouple the subcarriers by a DFT operation, resulting in inter-carrier interference (ICI). Due to this, conventional receivers may perform poorly when low-resolution ADCs are employed.

Second, the pilot signals transmitted by the user equipment (UE) for channel estimation at the base station (BS) are also received through the low-resolution ADCs. This necessitates the use of long pilot sequences for accurate channel estimation, leading to a loss in spectral efficiency [69, 70, 82].

Third, a channel encoder and decoder are integral parts of any commercial wireless communication system, and are used to correct for errors introduced by the channel. The channel decoders require the bit *log-likelihood ratios (LLRs)*, rather than hard bit-decisions, to provide good performance. The bit LLRs are a function of the posterior beliefs (probabilities) of the data symbols. Therefore, the aim of the receiver is not only to detect the data symbols, but also to obtain their posterior beliefs (also known as *soft symbols*), based on the quantized observations obtained from the low-resolution ADCs.

### 3.1.1 Prior Art

We now briefly review the existing literature on channel estimation and data detection in low-resolution ADC based multi-user MIMO systems, before presenting the key contributions in this chapter. Channel estimation in massive MIMO systems was considered in [51, 82, 85], while [89, 93, 94, 99, 110–113] develop data detection methods in massive MIMO single carrier (SC) and multi-carrier systems. Joint/iterative channel estimation and data detection was considered in [95, 96, 109, 114–116]. In [109], the authors develop a single iteration mutiuser MIMO-OFDM channel estimator using convex optimization techniques, and a data detector using a suboptimal soft-output MMSE algorithm. A bilinear generalized approximate message passing (BiGAMP) algorithm to solve the joint channel estimation and data detection problem is developed in [95,96]. The authors in [95] also analyze the performance of the BiGAMP algorithm using the replica method. Recently, a variational Bayesian (VB) channel estimation and data detection algorithm was developed in [116], in the context of a single-user single input single output (SU-SISO) OFDM system. While [116] restricts to a single OFDM symbol, we consider a more general multiple pilot and data symbols model in a multi-user massive MIMO-OFDM system.

An angular domain joint sparse channel estimation and data detection algorithm using the sparse Bayesian learning (SBL) framework in a hybrid millimeter wave communication system was proposed in [114]. The idea here is to utilize the decoded data symbols as virtual pilots for channel estimation. The receiver starts by forming an initial estimate of the channel using the pilot symbols, which is used to detect the data symbols. Then, in subsequent iterations, the detected data symbols are used as virtual pilots to refine the channel estimates and re-estimate the data symbols. This process is repeated until a suitable convergence condition is satisfied [60]. In [115], the authors adopt a supervised learning framework to solve the joint channel estimation and data decoding problem in a single carrier massive MIMO system. These approaches usually require careful parameter tuning for fast convergence and accurate data detection. Moreover, several heuristics are required to transform the detected data symbols into soft outputs which are required

for the subsequent channel decoding. Furthermore, none of the above mentioned papers directly address all three challenges mentioned in the preceding paragraphs.

In this chapter, we develop an iterative delay-domain *sparse channel estimation* and *soft symbol decoding* algorithm for a massive MIMO-OFDM system with low-resolution ADCs. As a first step, we develop a benchmark to evaluate the performance of any sparse channel estimator with measurements acquired using low-resolution ADCs. To this end, we consider a general quantized compressed sensing problem, and derive different types of Cramér-Rao lower bounds (CRLBs) on the mean squared error (MSE) performance of an estimator [117, 118]. We impose a two-stage hierarchical circularly symmetric complex Gaussian prior on the estimand (in our case, the channel) parameterized by a diagonal precision matrix. We further impose a non-informative conjugate Gamma hyperprior on the diagonal elements of the precision matrix. This results in a Student's $t$ marginalized prior on the estimand, which is heavy-tailed and hence promotes sparse solutions.

It is worth mentioning that CRLBs for the compressed sensing problem with unquantized measurements have been derived in [119]. In [120, 121], the CRLB on the MSE of an estimator with 1-bit measurements is derived under a non-sparse setting. While [120] derives the CRLB in a deterministic setup, [121] obtains the Bayesian CRLB. To the best of our knowledge, different types of CRLBs for the estimation of jointly compressible vectors [122] from multi-bit quantized noisy underdetermined measurements does not exist in the literature. We develop a CRLB for this case in Sec. 3.2. It turns out that the expectations required to obtain the Bayesian information matrix (BIM) are computationally intractable, and, consequently, the CRLB cannot be obtained in closed form. We therefore resort to numerical methods for evaluating the bound. While our CRLB for the quantized compressed sensing problem is of independent interest, we empirically illustrate its utility in the context of sparse massive MIMO-OFDM channel estimation by comparing it with the MSE performance of our algorithm.

Next, we use a statistical inference framework to compute the posterior distributions of the UEs' channels and data symbols given the quantized received pilot and data observations. We adopt a minorization-maximization based procedure called variational Bayes

(VB) inference, which is a principled approach for developing low-complexity algorithms for high-dimensional inference problems with guaranteed convergence from any initialization. The key novelty lies in how we construct the underlying probabilistic graphical models and how we identify and group the latent variables. The latent variables can also be used to compute side information such as the signal to noise ratio (SNR), which can, in turn, be used for link adaptation.

### 3.1.2 Main Contributions

Our main contributions in this chapter are as follows:

- We derive the Bayesian CRLB for the MSE incurred by an estimator for recovering jointly compressible vectors from quantized compressed sensing measurements. Specifically, we impose a hierarchical circularly symmetric complex Gaussian prior on the estimand, parameterized by a diagonal precision matrix. The precision matrix is in turn hyper-parameterized by a Gamma distribution. Although the CRLB is not available in closed-form, it can be evaluated by numerical methods.

- We consider both deterministic and random cases for the precision matrix to obtain two different CRLBs on the MSE of jointly compressible vectors. Also, in the case of exactly-sparse signals, we derive a support-aware CRLB, which assumes the knowledge of the support set, to compute the bound. We analytically show that our derived CRLB subsumes both the unquantized and 1 bit CRLBs as special cases.

- We exploit the lag/delay domain sparsity of the channels to formulate channel estimation in a massive MIMO-OFDM system as a quantized MMV sparse signal recovery problem. We develop a VB algorithm to infer the posterior distributions of the channels. We benchmark the MSE performance of the VB channel estimator with the derived Bayesian CRLB, and empirically show that our estimator meets the CRLB.

- We then present a quantized VB soft symbol decoding algorithm that uses the estimated channels to obtain the posterior beliefs of the data symbols. We use

these posterior statistics to generate virtual pilots, and develop a data-aided channel estimation procedure to refine the initial channel estimates. Based on this, we develop an iterative algorithm that alternately runs the soft symbol decoder and data-aided channel estimator steps. Finally, we generate the bit LLRs from the posterior symbol probabilities, and input them to the channel decoder.

- We also develop a variant of the iterative channel estimation and data decoding algorithm, which utilizes the aposteriori bit LLRs output from the channel decoder to adapt the prior used by the data detector. The resulting combined channel estimator, data detector and channel decoder further improves the system performance.

We evaluate the normalized MSE (NMSE) and coded bit-error-rate (BER) performance of the VB algorithms, and bechmark it against the state-of-the-art BiGAMP based joint channel estimator and data detector [95] and the conventional soft MMSE detector. Further, we study the impact of the system parameters on the performance of our algorithm, and provide several interesting insights.

One of the main takeaways from our work is that VB is a powerful and flexible technique for designing receivers in massive MIMO-OFDM systems, particularly when the BS employs low-resolution ADCs. This is because the subcarriers are no longer orthogonal after the quantization step. Due to this, conventional subcarrier-by-subcarrier data detection performs poorly (See Fig. 3.11). Also, our choice of latent variables and approximate posterior distributions is crucial for obtaining analytically and computationally tractable solutions. Another key takeaway is that the assumption of perfect CSI at the receiver significantly overestimates the system performance, which we illustrate through empirical studies in Sec. 3.8. Therefore, it is important to account for channel estimation errors while designing receivers, especially when both received pilots and data are coarsely quantized.

### Notation

We denote matrices, vectors and scalars by boldface upper case, boldface lower case, and lowercase letters, respectively. $\mathbf{A}^T$, $\mathbf{A}^H$ and $|\mathbf{A}|$ denote the transpose, conjugate transpose,

and determinant of $\mathbf{A}$, respectively. $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. diag($\mathbf{x}$) returns a diagonal matrix with the entries of $\mathbf{x}$ on the diagonal. $\mathbb{E}$ and $\langle \cdot \rangle$ both denote the expectation operation. $f(x) \triangleq \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$ and $F(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{t^2}{2}) \mathrm{d}t$ denote the probability density and cumulative distribution functions of a standard normal random variable evaluated at $x$. $\Gamma(a) \triangleq \int_0^\infty t^{a-1} \exp(-t)\, \mathrm{d}t$ denotes the Gamma function evaluated at $a > 0$. $\mathbf{I}_M$, $\mathbf{0}_M$ and $\mathbf{1}_M$ denote an $M \times M$ identity matrix, $M \times M$ zero matrix and all-ones vector of size $M \times 1$, respectively. $\Re$ and $\Im$ are the real and imaginary part operators, respectively.

## 3.2 Quantized Compressed Sensing and Bayesian Cramér-Rao Lower Bound

We consider the estimation of high-dimensional jointly compressible vectors $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_T] \in \mathbb{C}^{N \times T}$ from quantized low-dimensional measurements $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_T] \in \mathbb{C}^{M \times T}$, where $M < N$. The measurements are obtained as

$$\mathbf{Y} = \mathcal{Q}_b \left( \mathbf{\Phi X} + \mathbf{W} \right), \tag{3.1}$$

where $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$ is a known measurement matrix, and $\mathbf{W} \in \mathbb{C}^{M \times T}$ is the additive noise matrix whose entries are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables with mean 0 and variance $\sigma_w^2$. $\mathcal{Q}_b(\cdot)$ denotes an element-wise scalar $b$-bit quantizer of both real and imaginary components of its argument. We assume a common support structure on the columns of $\mathbf{X}$. In many applications, the signals are not exactly sparse, i.e., many entries may not be exactly equal to zero. An example is the effective wireless channel with the non-ideal transmit and receive filters. Therefore, we consider compressible signals [122] instead, where there are only a few entries with high magnitude and the remaining entries have very low magnitude. Here, by common support structure, we mean that the indices of the large magnitude entries are the same in each column of $\mathbf{X}$.

A $b$-bit quantizer on a real valued input $z$ is defined as $\mathcal{Q}_b(z) = L_i,\ z \in [\delta_i, \delta_{i+1})$, $i = 0, 1, \ldots, B-1$, where $B = 2^b$ is the number of quantization levels, $-\infty = \delta_0 < \delta_1 < \cdots < \delta_B = \infty$ are the quantization thresholds, and $L_0, L_1, \ldots, L_{B-1}$ are the quantizer outputs. The quantizer is deterministic and the quantization thresholds of $\mathcal{Q}_b(\cdot)$ are known. We now derive the Bayesian CRLB on the MSE of any estimator of $\mathbf{X}$.

To develop the CRLB, we impose a two-stage hierarchical prior on $\mathbf{X}$ [123]. That is, $\mathbf{x}_\ell \sim \mathcal{CN}(\mathbf{x}_\ell; \mathbf{0}, \mathbf{P}^{-1})\, \forall\, \ell$, where $\mathbf{P}$ is a diagonal precision matrix containing the hyperparameters $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]^T$. As mentioned earlier, we assume a non-informative conjugate Gamma hyperprior on $\alpha_n$, $\forall n$ with shape and rate parameters $a$ and $r$, respectively:

$$p(\mathbf{X} \,|\, \mathbf{P}) = \prod_{\ell=1}^{T} \frac{|\mathbf{P}|}{\pi^N} \exp(-\mathbf{x}_\ell^H \mathbf{P} \mathbf{x}_\ell), \tag{3.2}$$

$$p(\boldsymbol{\alpha}; a, r) = \prod_{n=1}^{N} \frac{r^a}{\Gamma(a)} \alpha_n^{a-1} \exp(-r\alpha_n), \tag{3.3}$$

where $|\mathbf{P}|$ denotes the determinant of $\mathbf{P}$ and $\Gamma(a)$ denotes the Gamma function.

Now, we compute the BIM for the above model. For this, we need the joint probability distribution $p(\mathbf{Y}, \mathbf{X}, \mathbf{P}; \boldsymbol{\Phi}, \sigma_w^2)$. Denoting the unquantized measurements by $\mathbf{Z}$, we write (3.1) as $\mathbf{Y} = \mathcal{Q}_b(\mathbf{Z})$, where $\mathbf{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_T] \in \mathbb{C}^{M \times T}$. It is convenient to transform the system from the complex field to the real field as follows:

$$\widetilde{\boldsymbol{\Phi}} = \begin{bmatrix} \Re(\boldsymbol{\Phi}) & -\Im(\boldsymbol{\Phi}) \\ \Im(\boldsymbol{\Phi}) & \Re(\boldsymbol{\Phi}) \end{bmatrix},\ \widetilde{\mathbf{x}}_\ell = \begin{bmatrix} \Re(\mathbf{x}_\ell) \\ \Im(\mathbf{x}_\ell) \end{bmatrix},\ \widetilde{\mathbf{y}}_\ell = \begin{bmatrix} \Re(\mathbf{y}_\ell) \\ \Im(\mathbf{y}_\ell) \end{bmatrix},$$
$$\widetilde{\mathbf{z}}_\ell = \begin{bmatrix} \Re(\mathbf{z}_\ell) \\ \Im(\mathbf{z}_\ell) \end{bmatrix},\ \widetilde{\mathbf{w}}_\ell = \begin{bmatrix} \Re(\mathbf{w}_\ell) \\ \Im(\mathbf{w}_\ell) \end{bmatrix},\ \ell = 1, \ldots, T, \tag{3.4}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary part operators, respectively. Let us denote $\widetilde{\sigma}_w^2 = \frac{\sigma_w^2}{2}$, $\widetilde{N} = 2N$, $\widetilde{M} = 2M$. In (3.4), $\mathbf{w}_\ell$ is the $\ell^{\text{th}}$ column of $\mathbf{W}$. Now, the system model becomes

$$\widetilde{\mathbf{Y}} = \mathcal{Q}_b(\widetilde{\mathbf{Z}}) = \mathcal{Q}_b(\widetilde{\boldsymbol{\Phi}} \widetilde{\mathbf{X}} + \widetilde{\mathbf{W}}). \tag{3.5}$$

Let us denote the precision matrix for the real field by $\widetilde{\mathbf{P}}$ ($\triangleq \text{diag}([2\boldsymbol{\alpha}^T, 2\boldsymbol{\alpha}^T]^T)$), where $\text{diag}(\cdot)$ returns a diagonal matrix. Now, the prior becomes

$$p(\widetilde{\mathbf{X}} \,|\, \widetilde{\mathbf{P}}) = \prod_{\ell=1}^{T} \frac{|\widetilde{\mathbf{P}}|^{\frac{1}{2}}}{(2\pi)^{\frac{\widetilde{N}}{2}}} \exp\left(-\frac{\widetilde{\mathbf{x}}_\ell^T \widetilde{\mathbf{P}} \widetilde{\mathbf{x}}_\ell}{2}\right). \tag{3.6}$$

Since the columns of $\widetilde{\mathbf{X}}$ are independent of each other, the BIM has a block diagonal structure with the off diagonal blocks as all zero matrices. With this prior, we present the expression for the BIM in the following theorem.

**Theorem 1.** *The $\ell^{th}$ diagonal block of the BIM required to compute the CRLB for the MSE of a Bayesian sparse signal estimator using quantized compressive measurements is given by* (3.7)*, where the expectation $\mathbb{E}[\cdot]$ is w.r.t. the joint probability distribution $p(\widetilde{\mathbf{Y}}, \widetilde{\mathbf{X}}, \widetilde{\mathbf{P}}; \widetilde{\mathbf{\Phi}}, \widetilde{\sigma}_w^2, a, r)$,*

$$\begin{aligned}\widetilde{\mathbf{M}}_\ell(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2) &= \mathbb{E}\left[-\frac{\partial^2}{\partial \widetilde{\mathbf{x}}_\ell \partial \widetilde{\mathbf{x}}_\ell^T} \log p(\widetilde{\mathbf{Y}}, \widetilde{\mathbf{X}}, \widetilde{\mathbf{P}}; \widetilde{\mathbf{\Phi}}, \widetilde{\sigma}_w^2, a, r)\right] \\ &= \widetilde{\mathbf{\Phi}}^T \, diag\left(\frac{1}{\widetilde{\sigma}_w^2}\mathbb{E}\left[\frac{\widetilde{\eta}_{m\ell}^{(hi)} f(\widetilde{\eta}_{m\ell}^{(hi)}) - \widetilde{\eta}_{m\ell}^{(lo)} f(\widetilde{\eta}_{m\ell}^{(lo)})}{F(\widetilde{\eta}_{m\ell}^{(hi)}) - F(\widetilde{\eta}_{m\ell}^{(lo)})} + \left(\frac{f(\widetilde{\eta}_{m\ell}^{(hi)}) - f(\widetilde{\eta}_{m\ell}^{(lo)})}{F(\widetilde{\eta}_{m\ell}^{(hi)}) - F(\widetilde{\eta}_{m\ell}^{(lo)})}\right)^2\right]\right)_{m=1}^{\widetilde{M}} \widetilde{\mathbf{\Phi}} + \mathbb{E}[\widetilde{\mathbf{P}}].\end{aligned} \tag{3.7}$$

$$\widetilde{\eta}_{m\ell}^{(hi)} \triangleq \frac{\widetilde{z}_{m\ell}^{(hi)} - \sum_{n=1}^{\widetilde{N}} \widetilde{\Phi}_{mn} \widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}, \tag{3.8}$$

$$\widetilde{\eta}_{m\ell}^{(lo)} \triangleq \frac{\widetilde{z}_{m\ell}^{(lo)} - \sum_{n=1}^{\widetilde{N}} \widetilde{\Phi}_{mn} \widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}, \tag{3.9}$$

*where $\ell \in \{1, \ldots, T\}$, $\widetilde{z}_{m\ell}^{(lo)}$ and $\widetilde{z}_{m\ell}^{(hi)}$ are the lower and upper quantization thresholds corresponding to the $(m,\ell)^{th}$ entry of $\widetilde{\mathbf{Y}}$, respectively. $\widetilde{\Phi}_{mn}$ and $\widetilde{x}_{n\ell}$ denote the $(m,n)^{th}$ and $(n,\ell)^{th}$ entries of $\widetilde{\mathbf{\Phi}}$ and $\widetilde{\mathbf{X}}$, respectively. $f(\cdot)$ and $F(\cdot)$ denote the probability density function (PDF) and cumulative distribution functions (CDF) of a standard normal random variable, respectively.*

*Proof.* The result follows from direct computation of the BIM, and is detailed in Section

3.10.1. □

We provide a step by step procedure to compute the Bayesian CRLB using (3.7). Given a realization of $\mathbf{X}$ generated according to a prior distribution $p(\mathbf{X}|\mathbf{P})$, we denote the instance specific BIM at iteration $\tau$ as $\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2)$ which is given by (3.10). Here, we use the law of iterated expectation, and the range of summation over $\widetilde{y}_{m\ell}$ is the number of discrete quantization levels.

$$\begin{aligned}\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2) \\ = \widetilde{\mathbf{\Phi}}^T \operatorname{diag}\left(\frac{1}{\widetilde{\sigma}_w^2}\sum_{\widetilde{y}_{m\ell}=L_0}^{L_{B-1}}\left(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - \widetilde{\eta}_{m\ell}^{(\mathrm{lo})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})}) + \frac{\left(f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})\right)^2}{F(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - F(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})}\right)\right)_{m=1}^{\widetilde{M}} \widetilde{\mathbf{\Phi}} \\ + \mathbb{E}[\widetilde{\mathbf{P}}]. \qquad (3.10)\end{aligned}$$

We provide the procedure to compute the instance specific BIM for the multi-bit quantized compressed sensing case in Algorithm 5.

Next, we use the chain rule to convert the real valued BIM to a complex valued BIM as

$$\begin{aligned}\mathbf{M}_{\ell}^{(\tau)}(\mathbf{\Phi}, a, r, \sigma_w^2) = \frac{1}{4}\left(\left[\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2)\right]_{\Re\Re} + \left[\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2)\right]_{\Im\Im}\right) \\ + \frac{j}{4}\left(\left[\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2)\right]_{\Re\Im} - \left[\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2)\right]_{\Im\Re}\right),\end{aligned}$$

where $\mathbf{M}_{\ell}^{(\tau)}(\mathbf{\Phi}, a, r, \sigma_w^2)$ is the $\ell^{\text{th}}$ diagonal block of the complex BIM. Finally, the instance specific CRLB is

$$\mathrm{CRLB}^{(\tau)}(\mathbf{\Phi}, a, r, \sigma_w^2) = \operatorname{blkdiag}\left[\left(\left[\mathbf{M}_{\ell}^{(\tau)}(\mathbf{\Phi}, a, r, \sigma_w^2)\right]^{-1}\right)_{\ell=1}^{T}\right], \qquad (3.11)$$

where blkdiag$(\cdot)$ returns a block diagonal matrix. We use the inverse property of block diagonal matrices to obtain (3.11), which reduces the complexity in computing the CRLB. We vary $\tau$ from 1 to $\tau_{max}$, compute the instance specific CRLB using independent random

**Algorithm 5:** Computation of $\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2)$

**Input:** $\widetilde{\mathbf{\Phi}}$, $\widetilde{\mathbf{X}}$, $\{\widetilde{z}_{m\ell}^{(\mathrm{lo})}, \widetilde{z}_{m\ell}^{(\mathrm{hi})}\}_{m=1}^{\widetilde{M}}$, $a$, $r$, $\widetilde{\sigma}_w$.

**Output:** $\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2)$.

1: Initialize $\mathbf{J} = \mathbf{0}_{\widetilde{M}}$.

2: **for** $m = 1$ to $\widetilde{M}$ **do**

3: **for** $\widetilde{y}_{m\ell} = \{L_0, \ldots, L_{B-1}\}$ **do**

4: Compute $\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}$ and $\widetilde{\eta}_{m\ell}^{(\mathrm{lo})}$ using (3.8) and (3.9).

5: Compute

$$\mathbf{J}(m,m) = \mathbf{J}(m,m) + \widetilde{\eta}_{m\ell}^{(\mathrm{hi})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - \widetilde{\eta}_{m\ell}^{(\mathrm{lo})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})}) + \frac{\left(f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})\right)^2}{F(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - F(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})}$$

6: **end for**

7: $\mathbf{J}(m,m) = \widetilde{\sigma}_w^{-2}\mathbf{J}(m,m)$.

8: **end for**

9: Compute $\widetilde{\mathbf{M}}_{\ell}^{(\tau)}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2) = \widetilde{\mathbf{\Phi}}^T\mathbf{J}\widetilde{\mathbf{\Phi}} + \frac{a}{r}\mathbf{I}_{\widetilde{N}}$.

realizations of $\mathbf{X}$, and then compute the average Bayesian CRLB as

$$\mathrm{CRLB}(\mathbf{\Phi}, a, r, \sigma_w^2) = \frac{1}{\tau_{max}} \sum_{\tau=1}^{\tau_{max}} \mathrm{CRLB}^{(\tau)}(\mathbf{\Phi}, a, r, \sigma_w^2). \tag{3.12}$$

The MSE of an estimator is lower bounded by the trace of the CRLB in (3.12). In the derivation above, we consider that the precision matrix $\mathbf{P}$ is random, which leads to a Bayesian bound. We can also consider the case where the precision matrix is deterministic. In this context, we contrast three types of bounds:

1. Support-aware Bayesian CRLB: Precision matrix is random, but the support set is known.

2. Hybrid CRLB: Random $\mathbf{X}$ parameterized by a deterministic $\mathbf{P}$.

3. Bayesian CRLB: Random $\mathbf{X}$ parameterized by a random $\mathbf{P}$ with a conjugate hyperprior.

For the support-aware Bayesian CRLB, the BIM is computed as follows: For the columns in the support set, the diagonal entries of $\mathbb{E}[\mathbf{P}]$ in the complex BIM are equal to $\frac{a}{r}$, and the remaining columns are removed from the measurement matrix, to compute the CRLB. We note that the support-aware Bayesian CRLB provides a lower bound on the MSE in the estimation of exact sparse vectors.

In the case of hybrid CRLB, the expectation term $\mathbb{E}[\widetilde{\mathbf{P}}]$ in (3.7) is replaced by the deterministic and known $\widetilde{\mathbf{P}}$. We generate a compressible signal using a generative model with a circularly symmetric complex normal prior parameterized by the known precision matrix, and average the CRLB over multiple realizations as in (3.12).

For the Bayesian CRLB, we use the mean of the Gamma hyperprior for $\mathbb{E}[\mathbf{P}] = \frac{a}{r}\mathbf{I}_N$ in the complex case. Note that this does not require the realization of the precision parameters; it only depends on the shape and rate parameters.

Next, we consider the two special cases, namely, 1-bit and infinite-bit quantization of the noisy compressive measurements, and specialize the derived CRLB to these two cases. It is easy to see that the second term $\mathbb{E}[\widetilde{\mathbf{P}}]$ in the BIM given in (3.7) depends only on the hyperparameters and not on the quantizer. So, we only discuss the first term in the sequel. The BIMs for the unquantized and 1-bit cases are obtained by careful algebraic manipulation of the multi-bit BIM, and we provide the details below.

### 3.2.1 Infinite-bit Quantized Compressed Sensing

The following Lemma is useful for obtaining the BIM in the infinite-bit quantized compressed sensing case.

**Lemma 1.** *For $\eta$, $\delta \in \mathbb{R}$,*

$$\lim_{\delta \to 0} \frac{(\eta+\delta)\, f(\eta+\delta) - \eta f(\eta)}{F(\eta+\delta) - F(\eta)} = 1 - \eta^2, \tag{3.13}$$

*and*

$$\lim_{\delta\to 0}\left(\frac{f(\eta+\delta)-f(\eta)}{F(\eta+\delta)-F(\eta)}\right)^2=\eta^2, \tag{3.14}$$

*where $f(\cdot)$ and $F(\cdot)$ are as defined earlier.* ■

*Proof.* Proof is available in the Sec. 3.10.2. □

As the number of bits $b$ increases, the difference between the quantization intervals $\widetilde{z}_{m\ell}^{(\mathrm{lo})}$ and $\widetilde{z}_{m\ell}^{(\mathrm{hi})}$ decreases, and tends to zero as $b\to\infty$. Therefore, if we apply Lemma 1 to each term inside the expectation in (3.7), it becomes unity, which results in the $\ell^{\text{th}}$ diagonal block of the BIM as

$$\widetilde{\mathbf{M}}_\ell(\widetilde{\mathbf{\Phi}},a,r,\widetilde{\sigma}_w^2)=\frac{\widetilde{\mathbf{\Phi}}^T\widetilde{\mathbf{\Phi}}}{\widetilde{\sigma}_w^2}+\mathbb{E}[\widetilde{\mathbf{P}}]. \tag{3.15}$$

Next, we compute the BIM for the 1-bit quantization case.

### 3.2.2 1-bit Quantized Compressed Sensing

For the 1-bit quantized compressed sensing case, we consider the output of the quantizer as the sign of its input. Thus, if we denote the $(m,\ell)^{\text{th}}$ entry of $\widetilde{\mathbf{Y}}$ and $\widetilde{\mathbf{Z}}$ as $\widetilde{y}_{m\ell}$ and $\widetilde{z}_{m\ell}$, respectively, then $\widetilde{y}_{m\ell}=+1$ if $\widetilde{z}_{m\ell}\geq 0$, and $\widetilde{y}_{m\ell}=-1$ otherwise. We simplify the BIM in the Lemma below.

**Lemma 2.** *The $\ell^{th}$ diagonal block of the BIM required to compute the CRLB for the MSE of a Bayesian sparse signal estimator using* 1*-bit compressive measurements is given by*

$$\widetilde{\mathbf{M}}_\ell(\widetilde{\mathbf{\Phi}},a,r,\widetilde{\sigma}_w^2)=\widetilde{\mathbf{\Phi}}^T diag\left(\frac{1}{\widetilde{\sigma}_w^2}\mathbb{E}\left[\frac{\widetilde{\xi}_{m\ell}f(\widetilde{\xi}_{m\ell})}{F(\widetilde{\xi}_{m\ell})}+\left(\frac{f(\widetilde{\xi}_{m\ell})}{F(\widetilde{\xi}_{m\ell})}\right)^2\right]\right)_{m=1}^{\widetilde{M}}\widetilde{\mathbf{\Phi}}+\mathbb{E}[\widetilde{\mathbf{P}}], \tag{3.16}$$

*where $\mathbb{E}[\cdot]$ is w.r.t. $p(\widetilde{\mathbf{Y}},\widetilde{\mathbf{X}},\widetilde{\mathbf{P}};\widetilde{\mathbf{\Phi}},\widetilde{\sigma}_w^2,a,r)$. Here,*

$$\widetilde{\xi}_{m\ell}\triangleq\frac{\widetilde{y}_{m\ell}\sum_{n=1}^{\widetilde{N}}\widetilde{\Phi}_{mn}\widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}, \tag{3.17}$$

*$\ell \in \{1, \ldots, T\}$, and the other notations are as in Theorem 1.* ■

*Proof.* Proof is available in the Sec. 3.10.3. □

It is worth mentioning that the BIMs for the unquantized compressed sensing in [119] and 1-bit compressed sensing in [120] (Fisher information matrix in [120]), are special cases of the BIM for multi-bit compressed sensing. The BIM for the 1-bit compressed sensing case in (3.16) can be further simplified (along the same lines as in [120]) to get

$$\widetilde{\mathbf{M}}_{\ell}(\widetilde{\mathbf{\Phi}}, a, r, \widetilde{\sigma}_w^2) = \widetilde{\mathbf{\Phi}}^T \mathrm{diag}\left(\frac{1}{\widetilde{\sigma}_w^2}\mathbb{E}\left[\frac{(f(\widetilde{\nu}_{m\ell}))^2}{F(\widetilde{\nu}_{m\ell})\left(1 - F(\widetilde{\nu}_{m\ell})\right)}\right]\right)_{m=1}^{\widetilde{M}} \widetilde{\mathbf{\Phi}} + \mathbb{E}[\widetilde{\mathbf{P}}] \tag{3.18}$$

where

$$\widetilde{\nu}_{m\ell} = \frac{\sum_{n=1}^{\widetilde{N}} \widetilde{\Phi}_{mn}\widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}, \tag{3.19}$$

and the first expectation is w.r.t. $p(\widetilde{\mathbf{X}}, \widetilde{\mathbf{P}})$.

We now turn to the massive MIMO-OFDM wireless communication system, and present the system model for the channel estimation and soft symbol decoding problems.

## 3.3 Description of Massive MIMO-OFDM System and Problem Statements

We consider the uplink (UL) of a single cell massive MIMO-OFDM system with $N_r$ antennas at the base station (BS) and $K$ single antenna user equipments (UEs), where $N_r \geq K$. Fig. 3.1 shows a block diagram of the system model. Each UE encodes and interleaves its information bits, and maps them to constellation symbols. The symbols are then loaded onto the subcarriers and OFDM modulated using an inverse discrete Fourier transform (IDFT). After passing the OFDM modulated data symbols through a parallel to serial converter, a cyclic prefix (CP) is added, RF up-converted to the passband, and transmitted over a frequency-selective wireless channel to the BS. At the BS, the received

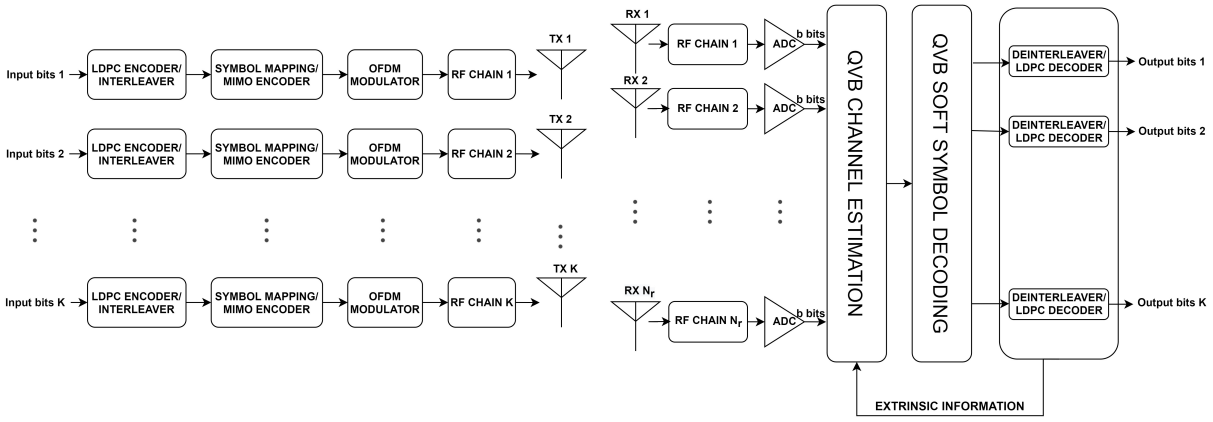


Figure 3.1: Massive MIMO-OFDM wireless communication system model.

RF signal is down-converted to baseband, the CP is removed, the signal is sampled, and quantized using $b$-bit ADCs to obtain the complex baseband signal.

Each UE transmits $\tau_p$ pilot OFDM symbols followed by $\tau_d$ data OFDM symbols. We assume that the coherence interval of the channel is at least $\tau_p + \tau_d$ OFDM symbols. In conventional OFDM systems, pilots are embedded in the same OFDM symbol along with the data, at specific subcarriers. However, due to the quantization errors introduced by the low-resolution ADCs, this results in severe inter-carrier interference between the pilot and data subcarriers, which cannot be canceled to obtain channel estimates. So, in this work, we consider a model where the pilot OFDM symbols are distinct from the data OFDM symbols. We denote the number of subcarriers by $N_c$. The unquantized received pilot and data signals at the $n_r^{\text{th}}$ receive antenna in the $n^{\text{th}}$ symbol interval ($0 \leq n \leq N_c - 1$) within the $t^{\text{th}}$ pilot and data OFDM symbol durations, respectively, are

$$\begin{aligned} z_{n_r}^{(p)}[t][n] &= \sum_{k=1}^{K} \sum_{\ell=0}^{L-1} h_{n_r,k}[\ell] \overline{x}_k^{(p)}[t][n-\ell] + w_{n_r}^{(p)}[t][n], \\ z_{n_r}^{(d)}[t][n] &= \sum_{k=1}^{K} \sum_{\ell=0}^{L-1} h_{n_r,k}[\ell] \overline{x}_k^{(d)}[t][n-\ell] + w_{n_r}^{(d)}[t][n], \end{aligned} \tag{3.20}$$

where $t \in \{1, \ldots, \tau_p\}$ and $t \in \{\tau_p + 1 \ldots \tau_p + \tau_d\}$ for the pilot and data phases, respectively, $h_{n_r,k}[\ell]$ is the complex channel gain of the $\ell^{\text{th}}$ delay tap of the channel between the $k^{\text{th}}$ UE and the $n_r^{\text{th}}$ receive antenna at the BS, distributed as $\mathcal{CN}(h_{n_r,k}[\ell]; 0, \beta_{k\ell})$,

where $\beta_{k\ell}$ is the large scale fading coefficient (LSFC), $L$ is the total number of delay taps of the frequency selective channel, $\overline{\mathbf{x}}_k^{(p)}[t] = [\overline{x}_k^{(p)}[t][0], \ldots, \overline{x}_k^{(p)}[t][N_c - 1]]$ and $\overline{\mathbf{x}}_k^{(d)}[t] = [\overline{x}_k^{(d)}[t][0], \ldots, \overline{x}_k^{(d)}[t][N_c - 1]]$ are the pilot and data symbols transmitted in the time domain by the $k^{\text{th}}$ UE in the $t^{\text{th}}$ OFDM symbol, respectively, and $w_{n_r}^{(p)}[t][n]$ and $w_{n_r}^{(d)}[t][n]$ are the complex additive white Gaussian noise during the pilot and data phases, respectively, with mean 0 and variance $\sigma_w^2$.

### 3.3.1 Delay-Domain Sparse Channel Estimation Model

In this subsection, we reformulate the received signal model above to utilize the lag-domain sparsity for channel estimation. We denote the channel sparsity, i.e., the maximum number of nonzero delay taps in the channel, by $L_{\text{sp}}$, where $L_{\text{sp}} \ll L$. We vectorize the unquantized received pilot signal in the $t^{\text{th}}$ OFDM symbol at the $n_r^{\text{th}}$ receive antenna as

$$\begin{aligned}\mathbf{z}_{n_r}^{(p)}[t] &= \begin{bmatrix} z_{n_r}^{(p)}[t][0] & z_{n_r}^{(p)}[t][1] & \ldots & z_{n_r}^{(p)}[t][N_c-1] \end{bmatrix}^T \\ &= \sum_{k=1}^{K} \overline{\mathbf{X}}_k^{(p)}[t] \overline{\mathbf{h}}_{n_r,k} + \mathbf{w}_{n_r}^{(p)}[t] \in \mathbb{C}^{N_c \times 1}, \end{aligned} \tag{3.21}$$

where $\overline{\mathbf{h}}_{n_r,k} = [\mathbf{h}_{n_r,k}^T, \mathbf{0}_{N_c-L}^T]^T \in \mathbb{C}^{N_c \times 1}$, $\mathbf{h}_{n_r,k} = [h_{n_r,k}[0], \ldots, h_{n_r,k}[L-1]]^T \in \mathbb{C}^{L \times 1}$ is the $k^{\text{th}}$ UE's frequency selective channel, and $\overline{\mathbf{X}}_k^{(p)}[t] \in \mathbb{C}^{N_c \times N_c}$ is a circulant matrix with its first column as $\overline{\mathbf{x}}_k^{(p)}[t]$. Using the fact that any circulant matrix is diagonalized by a unitary DFT matrix with the frequency domain coefficients as the eigenvalues, we rewrite (3.21) as

$$\begin{aligned}\mathbf{z}_{n_r}^{(p)}[t] &= \sum_{k=1}^{K} \mathbf{F}_{N_c}^H \mathbf{X}_k^{(p)}[t] \mathbf{F}_{N_c,L} \mathbf{h}_{n_r,k} + \mathbf{w}_{n_r}^{(p)}[t] \\ &= (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H) \mathbf{X}^{(p)}[t] \left(\mathbf{I}_K \otimes \mathbf{F}_{N_c,L}\right) \mathbf{h}_{n_r} + \mathbf{w}_{n_r}^{(p)}[t], \end{aligned} \tag{3.22}$$

where $\otimes$ denotes the matrix Kronecker product operator, $\mathbf{X}_k^{(p)}[t] = \mathbf{F}_{N_c} \overline{\mathbf{X}}_k^{(p)}[t] \mathbf{F}_{N_c}^H$ is a diagonal matrix with its entries as the pilots loaded on the subcarriers, $\mathbf{X}^{(p)}[t] = \texttt{diag}(\mathbf{X}_1^{(p)}[t], \ldots, \mathbf{X}_K^{(p)}[t]) \in \mathbb{C}^{KN_c \times KN_c}$, $\mathbf{h}_{n_r} = [\mathbf{h}_{n_r,1}^T, \ldots, \mathbf{h}_{n_r,K}^T]^T \in \mathbb{C}^{KL \times 1}$ is the vectorized lag domain frequency selective channel between all the users and the $n_r^{\text{th}}$ BS antenna.

$\mathbf{F}_{N_c} \in \mathbb{C}^{N_c \times N_c}$ and $\mathbf{F}_{N_c,L} \in \mathbb{C}^{N_c \times L}$ are the DFT and the $L$ column truncated DFT matrices, respectively. Note that $\mathbf{h}_{n_r}$ is a sparse vector with sparsity $KL_{\text{sp}}$. We stack the received vector of all the $N_r$ antennas and $\tau_p$ pilot symbols to obtain the unquantized received pilot matrix as shown in (3.23),

$$\begin{aligned}
\mathbf{Z}^{(p)} &= \begin{bmatrix} \mathbf{z}_1^{(p)}[1] & \dots & \mathbf{z}_{N_r}^{(p)}[1] \\ \vdots & \ddots & \vdots \\ \mathbf{z}_1^{(p)}[\tau_p] & \dots & \mathbf{z}_{N_r}^{(p)}[\tau_p] \end{bmatrix} \\
&= \begin{bmatrix} (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\mathbf{X}^{(p)}[1]\,(\mathbf{I}_K \otimes \mathbf{F}_{N_c,L}) \\ \vdots \\ (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\mathbf{X}^{(p)}[\tau_p]\,(\mathbf{I}_K \otimes \mathbf{F}_{N_c,L}) \end{bmatrix} \mathbf{H} + \mathbf{W}^{(p)} \\
&\triangleq \mathbf{\Phi}^{(p)}\mathbf{H} + \mathbf{W}^{(p)}. \qquad (3.23)
\end{aligned}$$

where $\mathbf{Z}^{(p)} \in \mathbb{C}^{\tau_p N_c \times N_r}$, $\mathbf{\Phi}^{(p)} \in \mathbb{C}^{\tau_p N_c \times KL}$, $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_{N_r}] \in \mathbb{C}^{KL \times N_r}$ is a row sparse channel matrix, and $\mathbf{W}^{(p)}$ is the additive noise matrix.

Now, we quantize the received signal using low-resolution ADCs. A $b$-bit quantizer on a real valued input $z$ is defined as $\mathcal{Q}_b(z) = L_i,\ z \in [\delta_i, \delta_{i+1}),\ i = 0, 1, \dots, B-1$, where $B = 2^b$ is the number of quantization levels, $\delta_0 < \delta_1 < \cdots < \delta_B$ are the quantization thresholds, and $L_0, L_1, \dots, L_{B-1}$ are the quantizer outputs. In this chapter, for simplicity and concreteness, we consider a uniform quantizer, where $\delta_l = (-B/2+l)\Delta, l = 1, \dots, B-1$, $\Delta$ is the quantization step size, and $L_l = (\delta_l + \delta_{l+1})/2, l = 0, \dots, B-1$. We set the dynamic range of the real and imaginary parts of the quantizer using the expected received signal power, $P_R$, as $\delta_0 = -2.5\sqrt{P_R/2}$, $\delta_B = 2.5\sqrt{P_R/2}$.[1] Our choice of $\delta_0$ and $\delta_B$ is motivated by the fact that the absolute value of a Gaussian distributed zero mean real-valued random variable with variance $P_R/2$ exceeds $2.5\sqrt{P_R/2}$ with probability less than 0.01, i.e., the quantizer gets overloaded with low probability. We quantize the received

[1]In practice, we quantize any value below $\delta_0$ to $L_0$, and any value above $\delta_B$ to $L_{B-1}$. Also, in practical systems, an automatic gain control unit is used to ensure that the power in the analog baseband signal is approximately equal to a predefined value, $P_R$, before quantization.

pilots in (3.23) using the $b$-bit ADCs to obtain the quantized received pilots as

$$\mathbf{Y}^{(p)} = \mathcal{Q}_b(\mathbf{Z}^{(p)}) = \mathcal{Q}_b(\mathbf{\Phi}^{(p)}\mathbf{H} + \mathbf{W}^{(p)}) \in \mathbb{C}^{\tau_p N_c \times N_r}. \tag{3.24}$$

Our first goal is to estimate $\mathbf{H}$ given $\mathbf{Y}^{(p)}$ and $\mathbf{\Phi}^{(p)}$ in (3.24). Note that, if $\tau_p N_c < KL$, (3.24) represents an underdetermined system of linear equations. In order to exploit the lag-domain sparsity in the channel, as in [124], we use a two stage hierarchical prior on $\mathbf{H}$ i.e., $\forall\, i,\ \mathbf{h}_i \sim \mathcal{CN}(\mathbf{h}_i; \mathbf{0}, \mathbf{P}^{-1})$, where the precision matrix $\mathbf{P}$ is diagonal and contains the hyperparameters $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_{KL}]^T$ as its diagonal elements. Further, we impose a Gamma hyperprior on $\boldsymbol{\alpha}$. This results in a Student's $t$ prior on $\mathbf{h}_i$, which is known to promote sparse channel estimates [123]. After estimating $\mathbf{H}$, our goal is to decode the data symbols. Next, we describe the signal model in the data transmission phase.

### 3.3.2 MIMO-OFDM Data Detection Model

We vectorize the unquantized received data at the $n_r^{\text{th}}$ receive antenna during the $t^{\text{th}}$ OFDM symbol in (3.20) as

$$\begin{aligned}\mathbf{z}_{n_r}^{(d)}[t] &= \begin{bmatrix} z_{n_r}^{(d)}[t][0] & z_{n_r}^{(d)}[t][1] & \ldots & z_{n_r}^{(d)}[t][N_c-1] \end{bmatrix}^T \\ &= \sum_{k=1}^{K} \mathbf{H}_{n_r,k}^{\text{time}} \mathbf{F}_{N_c}^H \mathbf{x}_k^{(d)}[t] + \mathbf{w}_{n_r}^{(d)}[t], \end{aligned} \tag{3.25}$$

where $t \in \{\tau_p+1, \ldots, \tau_p+\tau_d\}$ and $\mathbf{x}_k^{(d)}[t] = \begin{bmatrix} x_k^{(d)}[t][0] & \ldots & x_k^{(d)}[t][N_c-1] \end{bmatrix}^T = \mathbf{F}_{N_c}\overline{\mathbf{x}}_k^{(d)}[t] \in \mathbb{C}^{N_c \times 1}$ is the $M$-QAM modulated data symbols loaded on the subcarriers, where $\overline{\mathbf{x}}_k^{(d)}[t] = \begin{bmatrix} \overline{x}_k^{(d)}[t][0] & \ldots & \overline{x}_k^{(d)}[t][N_c-1] \end{bmatrix}^T \in \mathbb{C}^{N_c \times 1}$ is the time domain transmitted signal of the $k^{\text{th}}$ user. Also, $\mathbf{H}_{n_r,k}^{\text{time}} \in \mathbb{C}^{N_c \times N_c}$ is a circulant matrix with the first column as $\overline{\mathbf{h}}_{n_r,k}$ (from (3.21)). Using the diagonalizability property of a circulant matrix, we represent (3.25) as shown in (3.26),

$$\mathbf{z}_{n_r}^{(d)}[t] = \sum_{k=1}^{K} \mathbf{F}_{N_c}^H \mathbf{H}_{n_r,k}^{\text{freq}} \mathbf{x}_k^{(d)}[t] + \mathbf{w}_{n_r}^{(d)}[t]$$

$$= (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H) \begin{bmatrix} \mathbf{H}_{n_r,1}^{\text{freq}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{n_r,2}^{\text{freq}} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{H}_{n_r,K}^{\text{freq}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1^{(d)}[t] \\ \mathbf{x}_2^{(d)}[t] \\ \vdots \\ \mathbf{x}_K^{(d)}[t] \end{bmatrix} + \mathbf{w}_{n_r}^{(d)}[t]$$

$$= (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\, \mathbf{H}_{n_r}^{\text{freq}} \mathbf{x}^{(d)}[t] + \mathbf{w}_{n_r}^{(d)}[t]. \tag{3.26}$$

where $\mathbf{H}_{n_r,k}^{\text{freq}} = \mathbf{F}_{N_c} \mathbf{H}_{n_r,k}^{\text{time}} \mathbf{F}_{N_c}^H \in \mathbb{C}^{N_c \times N_c}$ is diagonal, containing the frequency domain representation of $\overline{\mathbf{h}}_{n_r,k}$, $\mathbf{x}^{(d)}[t] \triangleq \begin{bmatrix} \mathbf{x}_1^{(d)^T}[t] & \dots & \mathbf{x}_K^{(d)^T}[t] \end{bmatrix}^T$, $\mathbf{1}_K$ is the $K \times 1$ all ones vector, and $\mathbf{0}$ is an $N_c \times N_c$ all zero matrix. Now, we vectorize and stack the signal received over the $N_r$ receive antennas and $\tau_d$ OFDM data symbols to obtain $\mathbf{Z}^{(d)}$ as shown in (3.27),

$$\begin{aligned} \mathbf{Z}^{(d)} &= \begin{bmatrix} \mathbf{z}_1^{(d)}[\tau_p+1] & \dots & \mathbf{z}_1^{(d)}[\tau_p+\tau_d] \\ \vdots & \ddots & \vdots \\ \mathbf{z}_{N_r}^{(d)}[\tau_p+1] & \dots & \mathbf{z}_{N_r}^{(d)}[\tau_p+\tau_d] \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\mathbf{H}_1^{\text{freq}} \\ (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\mathbf{H}_2^{\text{freq}} \\ \vdots \quad \ddots \quad \vdots \\ (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\mathbf{H}_{N_r}^{\text{freq}} \end{bmatrix} \begin{bmatrix} \mathbf{x}^{(d)}[\tau_p+1] & \dots & \mathbf{x}^{(d)}[\tau_p+\tau_d] \end{bmatrix} + \mathbf{W}^{(d)} \\ &= \mathbf{D}\mathbf{X}^{(d)} + \mathbf{W}^{(d)}, \end{aligned} \tag{3.27}$$

where $\mathbf{D} \in \mathbb{C}^{N_r N_c \times K N_c}$ is the measurement matrix for data detection, $\mathbf{X}^{(d)} \in \mathbb{C}^{K N_c \times \tau_d}$ is the transmit data matrix, and $\mathbf{W}^{(d)}$ is the additive white Gaussian noise matrix during the data phase. Now, we quantize the received signal (3.27) using the $b$-bit ADCs to obtain

$$\mathbf{Y}^{(d)} = \mathcal{Q}_b(\mathbf{Z}^{(d)}) = \mathcal{Q}_b\left(\mathbf{D}\mathbf{X}^{(d)} + \mathbf{W}^{(d)}\right). \tag{3.28}$$

Our goal in this part is to decode the data symbols $\mathbf{X}^{(d)}$ given $\mathbf{Y}^{(d)}$ and $\mathbf{D}$. Next, we explain the data-aided channel estimation model that is used to refine the channel estimates.

### 3.3.3 Virtual Pilots-Aided MIMO-OFDM Channel Estimation

From section 3.3.1, we write the unquantized pilot received signal as $\mathbf{Z}^{(p)} = \mathbf{\Phi}^{(p)}\mathbf{H} + \mathbf{W}^{(p)}$. Similar to the pilot reception phase, if we consider the decoded data as known virtual pilot symbols, then we can write the received data signal as $\mathbf{Z}^{(d)}[t] = (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\left\langle \mathbf{X}^{(d)}[t] \right\rangle (\mathbf{I}_K \otimes \mathbf{F}_{N_c,L})\,\mathbf{H} + \mathbf{W}^{(d)}[t]$, where $t = \{\tau_p + 1, \ldots, \tau_p + \tau_d\}$, $\left\langle \mathbf{X}^{(d)}[t] \right\rangle = \texttt{diag}(\left\langle \mathbf{X}_1^{(d)}[t] \right\rangle, \ldots, \left\langle \mathbf{X}_K^{(d)}[t] \right\rangle) \in \mathbb{C}^{KN_c \times KN_c}$, and $\left\langle \mathbf{X}_k^{(d)}[t] \right\rangle = \texttt{diag}(\left\langle \mathbf{x}_k^{(d)}[t] \right\rangle) \in \mathbb{C}^{N_c \times N_c}$. Here, $\left\langle \mathbf{x}_k^{(d)}[t] \right\rangle$ are the posterior means of the decoded data symbols of the $k^{\text{th}}$ user during the $t^{\text{th}}$ OFDM symbol. We stack $\mathbf{Z}^{(p)}$ and $\mathbf{Z}^{(d)}[t]$ to obtain an expression for the unquantized received signal over one coherence interval as shown in (3.29),

$$\begin{aligned}\mathbf{Z} = \begin{bmatrix} \mathbf{Z}^{(p)} \\ \mathbf{Z}^{(d)}[\tau_p + 1] \\ \vdots \\ \mathbf{Z}^{(d)}[\tau_p + \tau_d] \end{bmatrix} &= \begin{bmatrix} \mathbf{\Phi}^{(p)} \\ (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\left\langle \mathbf{X}^{(d)}[\tau_p + 1] \right\rangle (\mathbf{I}_K \otimes \mathbf{F}_{N_c,L}) \\ \vdots \quad \ddots \quad \vdots \\ (\mathbf{1}_K^T \otimes \mathbf{F}_{N_c}^H)\left\langle \mathbf{X}^{(d)}[\tau_p + \tau_d] \right\rangle (\mathbf{I}_K \otimes \mathbf{F}_{N_c,L}) \end{bmatrix} \mathbf{H} + \mathbf{W} \\ &= \begin{bmatrix} \mathbf{\Phi}^{(p)} \\ \mathbf{\Phi}^{(d)} \end{bmatrix} \mathbf{H} + \mathbf{W} \triangleq \mathbf{\Phi}\mathbf{H} + \mathbf{W}. \end{aligned} \tag{3.29}$$

where $\mathbf{\Phi} \in \mathbb{C}^{(\tau_p+\tau_d)N_c \times KL}$ is the augmented measurement matrix and $\mathbf{W} \in \mathbb{C}^{(\tau_p+\tau_d)N_c \times N_r}$ is the additive white Gaussian noise matrix. The $b$-bit quantized received signal after the ADCs then reads

$$\mathbf{Y} = \mathcal{Q}_b(\mathbf{Z}) = \mathcal{Q}_b(\mathbf{\Phi}\mathbf{H} + \mathbf{W}) \in \mathbb{C}^{(\tau_p+\tau_d)N_c \times N_r}. \tag{3.30}$$

Our goal here is to estimate $\mathbf{H}$ given $\mathbf{Y}$ and $\mathbf{\Phi}$. Once we obtain an estimate of $\mathbf{H}$, we use it to obtain $\mathbf{D}$ as mentioned in (3.28), which in turn is used to refine the posterior beliefs of the $M-$QAM modulated data symbols in the next iteration of data decoding.

In the subsequent sections, we present our solutions to the above channel estimation and data detection problems.

## 3.4 Channel Estimation And Data Detection As Statistical Inference

We formulate the received system as probabilistic graphical models, and infer the posterior distributions of the channel and data symbols given the quantized pilot and data observations. We represent these Bayesian network graphical models in Figures 3.2, 3.3 and 3.4. As our goal is to obtain the posterior beliefs or LLRs of the data symbols that will be input to a channel decoder, a statistical inference framework is a suitable approach to solve our problems. We use shaded circles, transparent circles, and squares to represent the observations, latent variables, and deterministic variables. In our channel estimation and data detection problem, the quantized received pilot and data signals are the observations, channel and data symbols are the latent variables, pilot symbols and noise variance are the deterministic variables.

As mentioned earlier, the computational intractability of joint channel estimation and data detection problem necessitates us to adopt an iterative algorithm. So, we use separate Bayesian network models for the channel estimation and data decoding problems. We explain the intractability issue mathematically here. The posterior distribution of channel $\mathbf{H}$ and data $\{\mathbf{x}_1^{(d)}[\tau_p+1],\ldots,\mathbf{x}_K^{(d)}[\tau_p+\tau_d]\}$ given the observations $\mathbf{Y}^{(p)}$, $\mathbf{Y}^{(d)}$ and pilots $\mathbf{X}^{(p)}[1],\ldots,\mathbf{X}^{(p)}[\tau_p]$ is given by (3.31), where $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(d)}$ are the marginal likelihoods as shown in (3.32) and (3.33), respectively.[2]

$$
\begin{aligned}
&p\left(\mathbf{H},\mathbf{x}_1^{(d)}[\tau_p+1],\ldots,\mathbf{x}_K^{(d)}[\tau_p+\tau_d]\,|\,\mathbf{Y}^{(p)},\mathbf{Y}^{(d)};\mathbf{X}^{(p)}[1],\ldots,\mathbf{X}^{(p)}[\tau_p]\right)\\
&=\frac{p(\mathbf{Y}^{(p)}\,|\,\mathbf{H};\mathbf{X}^{(p)}[1],\ldots,\mathbf{X}^{(p)}[\tau_p])\,p(\mathbf{Y}^{(d)}\,|\,\mathbf{H},\mathbf{x}_1^{(d)}[\tau_p+1],\ldots,\mathbf{x}_K^{(d)}[\tau_p+\tau_d])}{p(\mathbf{Y}^{(p)};\mathbf{X}^{(p)}[1],\ldots,\mathbf{X}^{(p)}[\tau_p])\,p(\mathbf{Y}^{(d)})}\\
&\qquad\times p(\mathbf{H})\prod_{k=1}^{K}\prod_{t=\tau_p+1}^{\tau_p+\tau_d}p(\mathbf{x}_k^{(d)}[t]).
\end{aligned}
\tag{3.31}
$$

[2] $\mathbf{X}^{(d)}$ comes from a discrete $M$-QAM constellation, but we use integrals here for convenience. In the actual derivation, the integrals are replaced by summations.

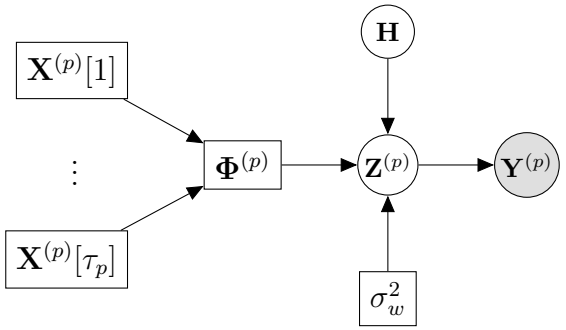


Figure 3.2: Bayesian network model for the channel estimation problem (3.24).

$$p(\mathbf{Y}^{(p)}) = \int p(\mathbf{Y}^{(p)} \,|\, \mathbf{H}; \mathbf{X}^{(p)}[1], \ldots, \mathbf{X}^{(p)}[\tau_p])\, p(\mathbf{H})\, \mathrm{d}\mathbf{H}, \tag{3.32}$$

$$p(\mathbf{Y}^{(d)}) = \int p(\mathbf{Y}^{(d)} \,|\, \mathbf{H}, \mathbf{x}_1^{(d)}[\tau_p+1], \ldots, \mathbf{x}_K^{(d)}[\tau_p+\tau_d])\, p(\mathbf{H}) \prod_{k=1}^{K} \prod_{t=\tau_p+1}^{\tau_p+\tau_d} p(\mathbf{x}_k^{(d)}[t])\, \mathrm{d}\mathbf{x}_k^{(d)}[t]\, \mathrm{d}\mathbf{H}. \tag{3.33}$$

Exact computation of the posterior distributions using the above is computationally intractable, as it requires solving high dimensional integrals over $\mathbf{H}$, $\mathbf{x}_1^{(d)}[\tau_p+1], \ldots, \mathbf{x}_K^{(d)}[\tau_p+\tau_d]$ to obtain the partition functions $P(\mathbf{Y}^{(p)})$ and $P(\mathbf{Y}^{(d)})$. Moreover, we estimate the UEs' channels in their lag domain, and use their frequency domain representation for data detection, which complicates the joint channel estimation and data detection problems further. These difficulties motivate the need to employ approximate inference techniques to solve the channel estimation and data detection problems.

## 3.5 Quantized VB Channel Estimation

Our goal is to infer the posterior distributions of the channels and the LLRs of the data symbols, given the quantized pilot and data observations. To this end, we adopt a statistical inference approach, where we formulate the received pilot and data signals using probabilistic graphical models. As mentioned in Sec. 3.4, exact computation of the posterior distributions is computationally intractable, as it requires solving high dimensional integrals over $\mathbf{H}$, $\mathbf{x}_1^{(d)}[\tau_p+1], \ldots, \mathbf{x}_K^{(d)}[\tau_p+\tau_d]$ to obtain the partition functions $P(\mathbf{Y}^{(p)})$ and $P(\mathbf{Y}^{(d)})$. This motivates the need for alternative approximate inference techniques,

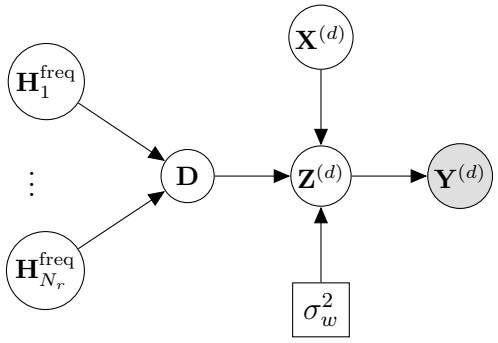


Figure 3.3: Bayesian network model for the data detection problem (3.28).

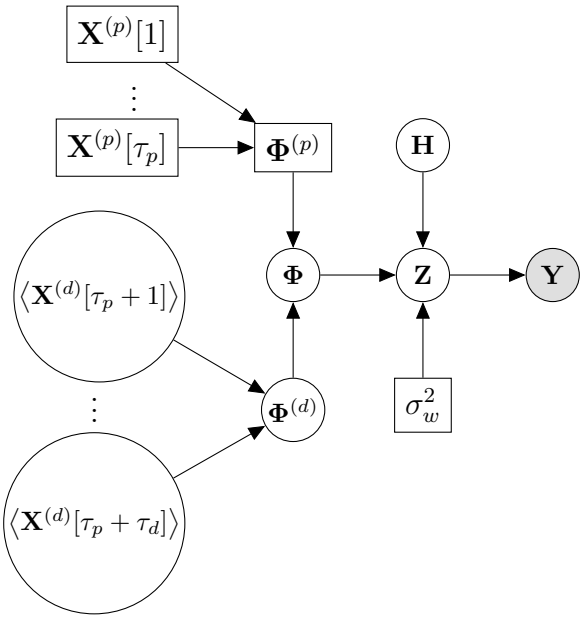


Figure 3.4: Bayesian network graphical model for the data-aided channel estimation problem in (3.30).

where we replace the exact posterior distribution with a tractable distribution that is close to the original in a particular distance measure, and is also easy to compute. As we will show, this leads to computationally tractable algorithms for the problem at hand. An excellent introduction to approximate inference can be found in [102].

We present Bayesian network graphical models for the channel estimation, data detection and data-aided channel estimation problems in Figures 3.2, 3.3, and 3.4, respectively. We use shaded circles, transparent circles, and squares to represent the observations, latent variables, and deterministic variables. In our channel estimation and data detection

problems, the quantized received pilot and data signals are the observations, and the channel and data symbols are the latent variables. Also, the pilot symbols and noise variance are deterministic and known. In the following paragraphs, we briefly describe variational Bayesian (VB) inference, which is the approximate inference technique adopted in this work to solve the pilot aided channel estimation problem given in (3.24).

VB is an iterative procedure to compute approximate posterior distributions of the latent variables given the observations. In this, we first write the logarithm of the observations $\mathbf{Y}^{(p)}$ as the sum of two terms and lower-bound it as

$$\ln p(\mathbf{Y}^{(p)}) = \mathcal{L}(q) + \mathrm{KL}(q \,\|\, p) \geq \mathcal{L}(q), \tag{3.34}$$

where $\mathcal{L}(q) \triangleq \int q(\mathbf{H}) \ln \left\{ \frac{p(\mathbf{Y}^{(p)}, \mathbf{H}; \mathbf{\Phi}^{(p)}, \sigma_w^2)}{q(\mathbf{H})} \right\} \mathrm{d}\mathbf{H}$ and
$\mathrm{KL}(q \| p) \triangleq -\int q(\mathbf{H}) \ln \left\{ \frac{p(\mathbf{H}|\mathbf{Y}^{(p)}, \mathbf{\Phi}^{(p)}, \sigma_w^2)}{q(\mathbf{H})} \right\} \mathrm{d}\mathbf{H} \geq 0$ are the evidence lower bound (ELBO) and non-negative Kullback Leibler (KL) divergence terms, respectively. Here, $q(\mathbf{H})$ is an approximate posterior distribution which is arbitrary, and can be approximated and optimized. Note that $q(\mathbf{H})$ depends on the observations $\mathbf{Y}^{(p)}$, but we do not explicitly include it in the notation for brevity. In the above, maximizing the ELBO $\mathcal{L}(q)$ would render an approximate distribution $q$ that is close to the original model evidence. We formally state the ELBO maximization problem as

$$q_{\mathrm{opt}} = \underset{q \in \mathcal{P}}{\arg\max} \;\; \mathcal{L}(q) = \underset{q \in \mathcal{P}}{\arg\min} \;\; \mathrm{KL}(q \,\|\, p), \tag{3.35}$$

where $\mathcal{P}$ is the space of probability distributions. The maximum of $\mathcal{L}(q)$ occurs when $q(\mathbf{H}) = p(\mathbf{H}|\mathbf{Y}^{(p)}, \mathbf{\Phi}^{(p)}, \sigma_w^2)$, but computing it is intractable. Therefore, we impose a factorized structure on each column of $\mathbf{H}$ i.e., $q(\mathbf{H}) = \prod_{i=1}^{N_r} q_i(\mathbf{h}_i)$. Substituting this in the ELBO, and simplifying it by fixing one of the factors, say $q_j(\mathbf{h}_j)$, we get

$$\mathcal{L}(q) = -\mathrm{KL}\left(q_j \| \widetilde{p}(\mathbf{Y}^{(p)}, \mathbf{h}_j; \mathbf{\Phi}^{(p)}, \sigma_w^2)\right) + \text{constant}, \tag{3.36}$$

where the constant terms do not depend on $q_j(\mathbf{h}_j)$, and $\widetilde{p}(\mathbf{Y}^{(p)}, \mathbf{h}_j; \mathbf{\Phi}^{(p)}, \sigma_w^2)$ is defined using $\ln \widetilde{p}(\mathbf{Y}^{(p)}, \mathbf{h}_j; \mathbf{\Phi}^{(p)}, \sigma_w^2) \triangleq \mathbb{E}_{i \neq j} \left[ \ln p(\mathbf{Y}^{(p)}, \mathbf{H}; \mathbf{\Phi}^{(p)}, \sigma_w^2) \right] + \text{constant}$, where the notation $\mathbb{E}_{i \neq j}[.]$ denotes the expectation with respect to the distributions $q_1(\mathbf{h}_1), \ldots, q_{N_r}(\mathbf{h}_{N_r})$ except $q_j(\mathbf{h}_j)$. Now, $\mathcal{L}(q)$ is maximized when the KL divergence term in (3.36) is minimized, which happens when $q_j(\mathbf{h}_j) = \widetilde{p}(\mathbf{Y}^{(p)}, \mathbf{h}_j; \mathbf{\Phi}^{(p)}, \sigma_w^2)$. Therefore, the optimal marginal distribution is

$$q_j(\mathbf{h}_j) = \text{const} \times \exp\left( \mathbb{E}_{i \neq j} \left[ \ln p(\mathbf{Y}^{(p)}, \mathbf{H}; \mathbf{\Phi}^{(p)}, \sigma_w^2) \right] \right), \tag{3.37}$$

where the constant is chosen such that $q_j$ becomes a probability distribution. Thus, VB is an iterative algorithm that falls in the category of minorization-maximization (MM), which solves a maximization problem by iteratively obtaining a lower bound on the objective function as in (3.34), and maximizing it. It is known that MM based optimization converges to a stationary point of the original optimization problem from any initialization [64].

Another important factor in our sparse channel estimation problem is the choice of prior on $\mathbf{H}$. In order to exploit the lag-domain sparsity in the channel, as in [124], we use a two stage hierarchical prior on $\mathbf{H}$ i.e., $\forall\, i,\ \mathbf{h}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{P}^{-1})$, where the precision matrix $\mathbf{P}$ is diagonal and contains the hyperparameters $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_{KL}]^T$ as its diagonal elements. Further, we impose a Gamma hyperprior on $\boldsymbol{\alpha}$, which results in a Student-$t$ prior on $\mathbf{h}_i$, which is known to promote sparse channel estimates [123].

We note that the marginal distribution in (3.37) is still hard to compute, as $p(\mathbf{Y}^{(p)}, \mathbf{H}; \mathbf{\Phi}^{(p)}, \sigma_w^2)$ contains terms involving the difference of the CDF of complex Gaussian random vectors, which does not lead to analytical expressions for the approximate posterior distribution. Hence, we add $\mathbf{Z}^{(p)}$ also as a latent variable. This leads to a closed form solution as described below.

We use the Bayesian network in Fig. 3.2 to express the logarithm of the joint probability distribution of the observations and latent variables as

$$\ln\, p(\mathbf{Y}^{(p)}, \mathbf{Z}^{(p)}, \mathbf{H}, \boldsymbol{\alpha}; \mathbf{\Phi}^{(p)}, \sigma_w^2, a, b) = \ln\, p(\mathbf{Y}^{(p)} \,|\, \mathbf{Z}^{(p)}) + \ln\, p(\mathbf{Z}^{(p)} \,|\, \mathbf{H}; \mathbf{\Phi}^{(p)}, \sigma_w^2)$$

$$+ \ln p(\mathbf{H}|\mathbf{P}) + \ln p(\boldsymbol{\alpha}; a, b), \quad (3.38)$$

where the prior distributions of $\mathbf{H}$ and $\boldsymbol{\alpha}$ are

$$p(\mathbf{H}|\mathbf{P}) = \prod_{n=1}^{N_r} \frac{|\mathbf{P}|}{\pi^{KL}} \exp\left(-\mathbf{h}_n^H \mathbf{P} \mathbf{h}_n\right), \quad (3.39)$$

$$p(\boldsymbol{\alpha}; a, b) = \prod_{k=1}^{KL} \frac{b^a}{\Gamma(a)} \alpha_k^{a-1} \exp\left(-b\alpha_k\right), \quad (3.40)$$

respectively. We set $a$ and $b$ to small values (say, $10^{-4}$) such that the hyperprior $p(\boldsymbol{\alpha}; a, b)$ is non-informative; and $\Gamma(a) \triangleq \int_0^\infty t^{a-1} \exp(-t)\, \mathrm{d}t$. We approximate the posterior distribution $p(\mathbf{Z}^{(p)}, \mathbf{H}, \boldsymbol{\alpha} \,|\, \mathbf{Y}^{(p)}; \boldsymbol{\Phi}^{(p)}, \sigma_w^2, a, b)$ of the latent variables as the factorized distribution:

$$\begin{aligned} p(\mathbf{Z}^{(p)}, \mathbf{H}, \boldsymbol{\alpha} \,|\, \mathbf{Y}^{(p)}; \boldsymbol{\Phi}^{(p)}, \sigma_w^2, a, b) &\approx q_{\mathbf{H}}(\mathbf{H}) q_{\mathbf{Z}}(\mathbf{Z}^{(p)}) q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \\ &= \prod_{n=1}^{N_r} q_{\mathbf{h}_n}(\mathbf{h}_n) \prod_{n=1}^{N_r} q_{\mathbf{z}_n}(\mathbf{z}_n^{(p)}) \prod_{k=1}^{KL} q_{\alpha_k}(\alpha_k), \end{aligned} \quad (3.41)$$

where we define $\mathbf{Z}^{(p)} \triangleq [\mathbf{z}_1^{(p)}, \ldots, \mathbf{z}_{N_r}^{(p)}]$ and $\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_{N_r}]$. Next, we express the conditional probability distributions of the observations and latent variables that are needed to compute the posterior distributions under the factorized structure as

$$\begin{aligned} p(\mathbf{Y}^{(p)} \,|\, \mathbf{Z}^{(p)}) &= \prod_{t=1}^{\tau_p N_c} \prod_{n=1}^{N_r} \mathbb{1}\left(\Re(z_{tn}^{(p)}) \in \left(\Re(z_{tn}^{(\mathrm{lo})}), \Re(z_{tn}^{(\mathrm{hi})})\right)\right) \mathbb{1}\left(\Im(z_{tn}^{(p)}) \in \left(\Im(z_{tn}^{(\mathrm{lo})}), \Im(z_{tn}^{(\mathrm{hi})})\right)\right) \\ &\triangleq \prod_{n=1}^{N_r} \mathbb{1}\left(\mathbf{z}_n^{(p)} \in \left(\mathbf{z}_n^{(\mathrm{lo})}, \mathbf{z}_n^{(\mathrm{hi})}\right)\right), \end{aligned} \quad (3.42)$$

$$p(\mathbf{Z}^{(p)} \,|\, \mathbf{H}; \boldsymbol{\Phi}^{(p)}, \sigma_w^2) = \prod_{n=1}^{N_r} \frac{1}{(\pi\sigma_w^2)^{\tau_p N_c}} \exp\left(-\frac{1}{\sigma_w^2} \|\mathbf{z}_n^{(p)} - \boldsymbol{\Phi}^{(p)} \mathbf{h}_n\|^2\right), \quad (3.43)$$

where $z_{tn}^{(p)}$ is the $(t,n)^{\text{th}}$ entry of $\mathbf{Z}^{(p)}$, $\mathbb{1}(\cdot)$ is the indicator function, $z_{tn}^{(\mathrm{lo})}$ and $z_{tn}^{(\mathrm{hi})}$ are the lower and upper quantization thresholds corresponding to the $(t,n)^{\text{th}}$ entry of $\mathbf{Y}^{(p)}$, respectively. The approximate posterior distributions of the latent variables are computed

by finding the expectations of the log of the joint distribution (3.38) with respect to the latent variables, and are provided in closed form in the following three Lemmas. The proofs for the Lemmas follow from (3.37), and are provided in the Sec. 3.10.

**Lemma 3** (Computation of $q_{\mathbf{H}}(\mathbf{H})$)**.** *The posterior distribution $q_{\mathbf{H}}(\mathbf{H})$ is complex normal with the covariance matrix of each of its columns and mean given by*

$$\mathbf{\Sigma}_{\mathbf{H}} = \left( \frac{1}{\sigma_w^2} \mathbf{\Phi}^{(p)^H} \mathbf{\Phi}^{(p)} + \langle \mathbf{P} \rangle \right)^{-1}, \tag{3.44}$$

$$\langle \mathbf{H} \rangle = \frac{1}{\sigma_w^2} \mathbf{\Sigma}_{\mathbf{H}} \mathbf{\Phi}^{(p)^H} \left\langle \mathbf{Z}^{(p)} \right\rangle, \tag{3.45}$$

*respectively. Here, $\langle \mathbf{P} \rangle = diag(\langle \boldsymbol{\alpha} \rangle)$, and $\left\langle \mathbf{Z}^{(p)} \right\rangle$ and $\langle \boldsymbol{\alpha} \rangle$ are the means of the approximate posteriors $q_{\mathbf{Z}}(\mathbf{Z}^{(p)})$ and $q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$, respectively.* ■

**Lemma 4** (Computation of $q_{\mathbf{Z}}(\mathbf{Z}^{(p)})$)**.** *The posterior distribution $q_{\mathbf{Z}}(\mathbf{Z}^{(p)})$ is truncated complex normal with mean $\left\langle \mathbf{Z}^{(p)} \right\rangle$ given by*

$$\left\langle \mathbf{Z}^{(p)} \right\rangle = \mathbf{\Phi}^{(p)} \langle \mathbf{H} \rangle + \frac{\sigma_w}{\sqrt{2}} \frac{f\left( \frac{\mathbf{Z}^{(lo)} - \mathbf{\Phi}^{(p)} \langle \mathbf{H} \rangle}{\sigma_w / \sqrt{2}} \right) - f\left( \frac{\mathbf{Z}^{(hi)} - \mathbf{\Phi}^{(p)} \langle \mathbf{H} \rangle}{\sigma_w / \sqrt{2}} \right)}{F\left( \frac{\mathbf{Z}^{(hi)} - \mathbf{\Phi}^{(p)} \langle \mathbf{H} \rangle}{\sigma_w / \sqrt{2}} \right) - F\left( \frac{\mathbf{Z}_n^{(lo)} - \mathbf{\Phi}^{(p)} \langle \mathbf{H} \rangle}{\sigma_w / \sqrt{2}} \right)}, \tag{3.46}$$

*where $\mathbf{Z}^{(lo)}$ and $\mathbf{Z}^{(hi)}$ are the lower and upper quantization levels corresponding to the observation $\mathbf{Y}^{(p)}$, respectively, and $\langle \mathbf{H} \rangle$ is the posterior mean of $q_{\mathbf{H}}(\mathbf{H})$. Also, $f(\cdot)$ and $F(\cdot)$ are the PDF and CDF of a standard normal random variable, respectively, computed element-wise on the real and imaginary parts of the argument. The division operation in* (3.46) *is also performed element-wise.* ■

**Lemma 5** (Computation of $q_{\alpha_k}(\alpha_k)$, $k = 1, \ldots, KL$)**.** *The posterior distribution $q_{\alpha_k}(\alpha_k)$ follows a Gamma distribution with shape and rate parameters given by*

$$\widetilde{a}_k = a + N_r \quad \textit{and} \quad \widetilde{b}_k = b + \sum_{n=1}^{N_r} \langle |h_{kn}|^2 \rangle, \tag{3.47}$$

**Algorithm 6:** Quantized VB Channel Estimation

**Input:** $\mathbf{Y}^{(p)}$, $\mathbf{\Phi}^{(p)}$, $\tau_p$, $\sigma_w$, $N_r$, $K$, $L$, $N_c$

**Output:** $\langle \mathbf{H} \rangle$

1: Initialize $\langle \mathbf{Z}^{(p)} \rangle$, $\langle \boldsymbol{\alpha} \rangle$, $a$, $b$.
2: **repeat**
3: $\quad \langle \mathbf{P} \rangle = \mathtt{diag}(\langle \boldsymbol{\alpha} \rangle)$
4: $\quad$ Compute $\mathbf{\Sigma_H}$ using (3.44).
5: $\quad$ Compute $\langle \mathbf{H} \rangle$ using (3.45).
6: $\quad$ Compute $\left\langle \mathbf{Z}^{(p)} \right\rangle$ using (3.46).
7: $\quad$ Compute $\langle \alpha_k \rangle$ using (3.48), $k = 1, \ldots, KL$.
8: **until** stopping condition is met

*respectively. Its mean is given by*

$$\langle \alpha_k \rangle = \frac{a + N_r}{b + \sum_{n=1}^{N_r} \langle |h_{kn}|^2 \rangle}, \tag{3.48}$$

*where* $h_{kn}$ *is the* $(k,n)^{th}$ *element of* $\mathbf{H}$*, and* $\langle |h_{kn}|^2 \rangle = |\langle h_{kn} \rangle|^2 + \mathbf{\Sigma_H}[k,k]$. ■

Note that we have included the subscript $k$ in $\widetilde{a}_k$ for consistency of notation, even though it is independent of $k$. From (3.44), (3.45), (3.46), and (3.48), we see that the statistics of the approximate posterior distributions $q_{\mathbf{H}}(\mathbf{H})$, $q_{\mathbf{Z}}(\mathbf{Z}^{(p)})$, and $q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$ depend on each other. The VB algorithm proceeds iteratively by randomly initializing the posteriors and alternately computing each of the approximate posterior distributions until a suitable convergence condition is satisfied. Once the algorithm converges, we use the posterior mean from (3.45) as the final channel estimate. Then, we compute the DFT of the lag domain channel estimates, and use them for data decoding. We present VB channel estimation procedure in Algorithm 6.

## 3.6 Quantized VB Soft Symbol Decoding

In this section, we develop a VB algorithm for soft symbol decoding in MIMO-OFDM systems using the system model in (3.28), reproduced here for convenience:

$$\mathbf{Y}^{(d)} = \mathcal{Q}_b(\mathbf{Z}^{(d)}) = \mathcal{Q}_b\left(\mathbf{D}\mathbf{X}^{(d)} + \mathbf{W}^{(d)}\right), \tag{3.49}$$

where $\mathbf{Y}^{(d)} = [\mathbf{y}^{(d)}[\tau_p+1], \ldots, \mathbf{y}^{(d)}[\tau_p+\tau_d]] \in \mathbb{C}^{N_r N_c \times \tau_d}$, $\mathbf{Z}^{(d)} = [\mathbf{z}^{(d)}[\tau_p+1], \ldots, \mathbf{z}^{(d)}[\tau_p+\tau_d]] \in \mathbb{C}^{N_r N_c \times \tau_d}$, $\mathbf{X}^{(d)} = [\mathbf{x}^{(d)}[\tau_p+1], \ldots, \mathbf{x}^{(d)}[\tau_p+\tau_d]] \in \mathbb{C}^{KN_c \times \tau_d}$, $\mathbf{x}^{(d)}[t] = [x_{1t}^{(d)}, \ldots, x_{KN_c,t}^{(d)}]^T$, $\mathbf{D} \in \mathbb{C}^{N_r N_c \times KN_c}$, $\mathbf{W}^{(d)} \in \mathbb{C}^{N_r N_c \times \tau_d}$. We represent the corresponding Bayesian network in Fig. 3.3.

Similar to Sec. 3.5, we consider the unquantized received data signal as a latent variable, and express the logarithm of the joint probability distribution of the observations and the latent variables as

$$\ln\, p\left(\mathbf{Y}^{(d)}, \mathbf{Z}^{(d)}, \mathbf{X}^{(d)} | \mathbf{D}, \sigma_w^2\right) = \ln p(\mathbf{Y}^{(d)} | \mathbf{Z}^{(d)}) + \ln p\left(\mathbf{Z}^{(d)} | \mathbf{X}^{(d)}, \mathbf{D}, \sigma_w^2\right) + \ln p(\mathbf{X}^{(d)}). \tag{3.50}$$

We factorize the posterior distribution of $\mathbf{Z}^{(d)}$ and $\mathbf{X}^{(d)}$ as

$$p\left(\mathbf{Z}^{(d)}, \mathbf{X}^{(d)} \,|\, \mathbf{Y}^{(d)}, \mathbf{D}, \sigma_w^2\right) \approx q_{\mathbf{Z}}\left(\mathbf{Z}^{(d)}\right) \prod_{t=\tau_p+1}^{\tau_p+\tau_d} \prod_{k=1}^{KN_c} q_{x_{kt}}\left(x_{kt}^{(d)}\right), \tag{3.51}$$

where $\mathbf{Z}^{(d)} = [\mathbf{z}_{\tau_p+1}^{(d)}, \ldots, \mathbf{z}_{\tau_p+\tau_d}^{(d)}]$, and $x_{kt}^{(d)}$ is the $k^{\text{th}}$ component of $\mathbf{x}^{(d)}[t]$. We write the conditional probability distributions in (3.50) as follows:

$$p(\mathbf{Y}^{(d)} \,|\, \mathbf{Z}^{(d)}) = \mathbb{1}\left(\mathbf{Z}^{(d)} \in (\mathbf{Z}^{(\text{lo})}, \mathbf{Z}^{(\text{hi})})\right), \tag{3.52}$$

$$p(\mathbf{Z}^{(d)} \,|\, \mathbf{X}^{(d)}; \mathbf{D}, \sigma_w^2) = \prod_{t=\tau_p+1}^{\tau_p+\tau_d} \frac{1}{(\pi\sigma_w^2)^{N_r N_c}} \exp\left(-\frac{1}{\sigma_w^2} \left\| \mathbf{z}^{(d)}[t] - \mathbf{D}\mathbf{x}^{(d)}[t] \right\|_2^2\right), \tag{3.53}$$

where $\mathbb{1}(\cdot)$ is the indicator function, $\mathbf{Z}^{(\text{lo})}, \mathbf{Z}^{(\text{hi})}$ are the entry-wise lower and upper quantization intervals of the real and imaginary components of $\mathbf{Y}^{(d)}$. We present the approximate posterior distributions that maximize the ELBO in the following two Lemmas. The computation of the approximate posterior distribution $q_{\mathbf{Z}}\left(\mathbf{Z}^{(d)}\right)$ is similar to the computation of the posterior distribution of the unquantized pilot received signal in Lemma 4. Therefore, we omit the proof of Lemma 7 to avoid repetition. The proof of Lemma 6 is available in the Sec. 3.10.

**Lemma 6** (Computation of $q_{x_{kt}}(x_{kt}^{(d)})$)**.** *The posterior* $q_{x_{kt}}\left(x_{kt}^{(d)}\right)$ *follows a Boltzmann distribution with the probability mass function*

$$q_{x_{kt}}\left(x_{kt}^{(d)} = s_m\right) = \frac{\exp\left(f_{kt}(s_m)\right)}{\sum_{s' \in \mathbb{M}} \exp\left(f_{kt}(s')\right)} \tag{3.54}$$

*for* $m = 1, \ldots, M$*, where* $k \in \{1, \ldots, KN_c\}$*,* $t \in \{\tau_p + 1, \ldots, \tau_p + \tau_d\}$*,* $\mathbb{M} = \{s_1, \ldots, s_M\}$ *is the signal constellation set of cardinality* $M$*, and* $f_{kt}(s)$ *is shown in* (3.55)*,*

$$f_{kt}(s) = -\frac{1}{\sigma_w^2}\Big( \|\mathbf{D}_{:,k}\|^2 |s|^2 - 2\Re\Big[\mathbf{D}_{:,k}^H\Big( \left\langle \mathbf{z}^{(d)}[t] \right\rangle - \sum_{\substack{k'=1 \\ k' \neq k}}^{KN_c} \mathbf{D}_{:,k'} \left\langle x_{k't}^{(d)} \right\rangle \Big) s^*\Big]\Big) + \ln p(x_{kt}^{(d)} = s), \tag{3.55}$$

*where* $\Re$ *and* $*$ *denote the real part and complex conjugate operators, respectively,* $\mathbf{D}_{:,k}$ *is the* $k^{th}$ *column of* $\mathbf{D}$*,* $\left\langle \mathbf{z}^{(d)}[t] \right\rangle$ *and* $\left\langle x_{k't}^{(d)} \right\rangle$ *are the posterior means of* $q_{\mathbf{z}_t}(\mathbf{z}^{(d)}[t])$ *and* $q_{x_{k't}}(x_{k't}^{(d)})$*, respectively.* ■

We compute the mean and mean square value of $q_{x_{kt}}(x_{d,kt})$ as follows:

$$\left\langle x_{kt}^{(d)} \right\rangle = \sum_{s \in \mathbb{M}} s\, q_{x_{kt}}(s), \quad \left\langle |x_{kt}^{(d)}|^2 \right\rangle = \sum_{s \in \mathbb{M}} |s|^2\, q_{x_{kt}}(s).$$

**Lemma 7** (Computation of $q_{\mathbf{Z}}\left(\mathbf{Z}^{(d)}\right)$)**.** *The posterior distribution* $q_{\mathbf{Z}}\left(\mathbf{Z}^{(d)}\right)$ *is truncated*

**Algorithm 7:** Quantized VB Soft Symbol Decoding

**Input:** $\mathbf{Y}^{(d)}, \mathbf{D}$, $\mathbb{M} = \{s_1, \ldots, s_M\}$, $\tau_p$, $\tau_d$, $\sigma_w$, $K$, $N_c$.

**Output:** $q_{\mathbf{X}}(\mathbf{X}^{(d)})$, $\left\langle \mathbf{X}^{(d)} \right\rangle$

1: Initialize $\left\langle \mathbf{Z}^{(d)} \right\rangle$, $\langle \mathbf{X}^{(d)} \rangle = \mathbf{0}_{KN_c \times \tau_d}$
2: **repeat**
3:  **for** $k = 1$ to $KN_c$ **do**
4:  **for** $t = \tau_p + 1$ to $\tau_p + \tau_d$ **do**
5:  Compute $q_{x_{kt}}\left(x_{kt}^{(d)} = s\right)$ using (3.54) $\forall s \in \mathbb{M}$.
6:  Compute $\langle x_{kt}^{(d)} \rangle = \sum_{s \in \mathbb{M}} s\, q_{x_{kt}}(x_{kt}^{(d)} = s)$.
7:  **end for**
8:  **end for**
9:  Compute $\left\langle \mathbf{Z}^{(d)} \right\rangle$ using (3.56).
10: **until** stopping condition is met

*complex normal, with mean*

$$\left\langle \mathbf{Z}^{(d)} \right\rangle = \mathbf{D} \left\langle \mathbf{X}^{(d)} \right\rangle + \frac{\sigma_w}{\sqrt{2}} \frac{f\left(\frac{\mathbf{Z}^{(lo)} - \mathbf{D}\langle \mathbf{X}^{(d)} \rangle}{\sigma_w/\sqrt{2}}\right) - f\left(\frac{\mathbf{Z}^{(hi)} - \mathbf{D}\langle \mathbf{X}^{(d)} \rangle}{\sigma_w/\sqrt{2}}\right)}{F\left(\frac{\mathbf{Z}^{(hi)} - \mathbf{D}\langle \mathbf{X}^{(d)} \rangle}{\sigma_w/\sqrt{2}}\right) - F\left(\frac{\mathbf{Z}^{(lo)} - \mathbf{D}\langle \mathbf{X}^{(d)} \rangle}{\sigma_w/\sqrt{2}}\right)}, \tag{3.56}$$

*where* $\mathbf{Z}^{(lo)}$ *and* $\mathbf{Z}^{(hi)}$ *are defined in* (3.52), $\left\langle \mathbf{X}^{(d)} \right\rangle$ *contains the posterior means of* $q_{x_{kt}}(x_{kt}^{(d)})$ $\forall k, t$ *as its entries,* $f(\cdot)$, $F(\cdot)$, *and the division operation are as defined in Lemma 4.* ■

As mentioned in Section 3.5, the VB algorithm starts by randomly initializing the latent variables, and iteratively computes the posterior distributions of data symbols. We use the posterior distributions of the data symbols in (3.54) to calculate the bit LLRs. We present the quantized VB soft symbol decoding procedure in Algorithm 7. Next, we describe the data-aided channel estimation procedure.

## 3.7 Iterative Quantized VB Channel Estimation and Soft Symbol Decoding

In this section, we merge the channel estimation and soft symbol decoding into an iterative algorithm that improves on the system performance obtained by only executing Algorithms 6 and 7. We utilize the data-aided channel estimation system model to refine the channel estimates in an iterative fashion. Recall our system model (3.30) from section 3.3.3: $\mathbf{Y} = \mathcal{Q}_b(\mathbf{Z}) = \mathcal{Q}_b(\mathbf{\Phi H} + \mathbf{W})$. We start with the pilot based channel estimation Algorithm 6 followed by the soft symbol decoding Algorithm 7. Now, we utilize the posterior means of the decoded data symbols to form a new measurement matrix $\mathbf{\Phi}$ that is input to the channel estimation block. In a VB procedure, we obtain the posterior distribution of a given latent variable by computing the expectation of the joint probability distribution w.r.t. the posterior distributions of all the other latent variables. This in turn means that its posterior distribution depends only on the posterior statistics of the other latent variables. Moreover, in the context of soft symbol decoding, the posterior statistics of the data symbol enter the equivalent measurement equation through their posterior means (see (3.55) and (3.56)). Therefore, we use the posterior means of the data symbols to construct a new measurement matrix for iterative channel estimation.

The data-aided channel estimation procedure follows the same steps as in algorithm 6 except that its inputs $\mathbf{Y}^{(p)}$ and $\mathbf{\Phi}^{(p)}$ are replaced by $\mathbf{Y}$ and $\mathbf{\Phi}$, respectively. The derivation follows the same procedure as in Sec. 3.5; we provide the final expressions of the posterior statistics of the latent variables below.

$$\mathbf{\Sigma_H} = \left(\frac{1}{\sigma_w^2}\mathbf{\Phi}^H\mathbf{\Phi} + \langle \mathbf{P} \rangle\right)^{-1}, \quad \langle \mathbf{H} \rangle = \frac{1}{\sigma_w^2}\mathbf{\Sigma_H}\mathbf{\Phi}^H \langle \mathbf{Z} \rangle, \tag{3.57}$$

$$\langle \alpha_k \rangle = \frac{a + N_r}{b + \sum_{n=1}^{N_r} \langle |h_{kn}|^2 \rangle}, \qquad k = 1, \ldots, KL, \tag{3.58}$$

$$\langle \mathbf{Z} \rangle = \mathbf{\Phi} \langle \mathbf{H} \rangle + \frac{\sigma_w}{\sqrt{2}} \frac{f\left(\frac{\mathbf{Z}^{(\mathrm{lo})} - \mathbf{\Phi}\langle \mathbf{H} \rangle}{\sigma_w/\sqrt{2}}\right) - f\left(\frac{\mathbf{Z}^{(\mathrm{hi})} - \mathbf{\Phi}\langle \mathbf{H} \rangle}{\sigma_w/\sqrt{2}}\right)}{F\left(\frac{\mathbf{Z}^{(\mathrm{hi})} - \mathbf{\Phi}\langle \mathbf{H} \rangle}{\sigma_w/\sqrt{2}}\right) - F\left(\frac{\mathbf{Z}^{(\mathrm{lo})} - \mathbf{\Phi}\langle \mathbf{H} \rangle}{\sigma_w/\sqrt{2}}\right)}, \tag{3.59}$$

where $\mathbf{Z}^{(\mathrm{lo})}$ and $\mathbf{Z}^{(\mathrm{hi})}$ are the lower and upper quantization thresholds corresponding to

$\mathbf{Y}$. The other notations are as defined in Lemma 4. We repeat this process of channel estimation and data decoding for a fixed number of iterations. Finally, we use the posterior distribution of the transmit symbols to obtain the bit LLRs, which are deinterleaved and input to the channel decoder. We present the iterative VB channel estimation and soft symbol decoding in Algorithm 8.

Now, we present a variant of the iterative channel estimation and soft symbol decoding Algorithm 8 that marginally improves the system performance. In this, we merge the quantized VB algorithms with the channel decoder using a data prior adaptation strategy. We utilize the aposteriori bit LLRs from the channel decoder to adapt the data prior that is input to the soft symbol detector in the next iteration. We interleave the posterior bit LLRs output by the channel decoder, and generate the extrinsic information to compute the symbol LLRs, which are then used to adapt the data prior. Instead of using a uniform prior on the data symbols, we bias the data detector by a non-uniform data prior. In every outer iteration of the iterative channel estimator and soft symbol decoder, we increase the probability mass on the data symbol output by the extrinsic information progressively by a judiciously chosen step size. At lower SNRs, such prior adaptation may lead to error propagation effects, but at SNRs of interest, this leads to performance improvement. We show the final block diagram for one outer iteration of the iterative channel estimation and soft symbol decoding algorithm with data prior adaptation in Fig. 3.5. We use this algorithm in our simulations, unless specified otherwise.

### 3.7.1 Computational Complexity

In this subsection, we analyze the per-iteration computational complexity of the VB algorithm. Specifically, we analyze the scaling of the number of floating point operations (flops) with the system dimensions. Table 3.1 shows the order ($\mathcal{O}$) of the per-iteration computational complexity of the steps involved in one iteration of Algorithm 8. The per-iteration complexity scales cubically with the number of subcarriers, as the square of the number of users and the channel length, and linearly with the number of receive antennas, constellation size, and number of pilot and data symbols. In particular, the complexity

**Algorithm 8:** Iterative Quantized VB Channel Estimation and Soft Symbol Decoding

**Input:** $\mathbf{Y}^{(p)}$, $\mathbf{Y}^{(d)}$, $\mathbf{X}^{(p)}$, $\mathbb{M} = \{s_1, \ldots, s_M\}$, $M$, $\tau_p$, $\tau_d$, $\sigma_w$, $N_r$, $K$, $L$, $N_c$, MAX_ITER.

**Output:** $\langle \mathbf{H} \rangle, q_{\mathbf{X}}(\mathbf{X}^{(d)}), \langle \mathbf{X}^{(d)} \rangle$

Initialize $\langle \mathbf{H} \rangle$, $\langle \boldsymbol{\alpha} \rangle$, $\langle \mathbf{X}^{(d)} \rangle = \mathbf{0}_{KN_c \times \tau_d}$, $a$, $b$
**Initial Channel Estimation**: Run Algorithm 6.
**repeat**
  **Soft Symbol Decoding**:
  **repeat**
    Compute $\mathbf{D}$ using (3.27).
    **for** $k = 1$ to $KN_c$ **do**
      **for** $t = \tau_p + 1$ to $\tau_p + \tau_d$ **do**
        Compute $q_{x_{kt}}\left(x_{kt}^{(d)} = s\right)$ using (3.54) $\forall s \in \mathbb{M}$.
        Compute $\langle x_{kt}^{(d)} \rangle = \sum_{s \in \mathbb{M}} s\, q_{x_{kt}}(x_{kt}^{(d)} = s)$.
      **end for**
    **end for**
    Compute $\left\langle \mathbf{Z}^{(d)} \right\rangle$ using (3.56).
  **until** stopping condition is met
  **Data-Aided Channel Estimation**:
  Compute $\boldsymbol{\Phi}$ using (3.29), $\tilde{a}_k = a + N_r$, $k = 1, \ldots, KL$.
  Initialize $\langle \mathbf{Z} \rangle$, $\langle \boldsymbol{\alpha} \rangle$.
  **repeat**
    $\langle \mathbf{P} \rangle = \texttt{diag}(\langle \boldsymbol{\alpha} \rangle)$
    Compute $\boldsymbol{\Sigma}_{\mathbf{H}}$ and $\langle \mathbf{H} \rangle$ using (3.57).
    Compute $\langle \alpha_k \rangle$ using (3.58), $k = 1, \ldots, KL$.
    Compute $\langle \mathbf{Z} \rangle$ using (3.59).
  **until** stopping condition is met
**until** MAX_ITER times

Table 3.1:

Per-iteration Complexity of the QVB Algorithm 8

| **Matrix** | **Order Complexity** |
|---|---|
| $\langle \mathbf{X}^{(d)} \rangle$ | $MK^2N_c^3N_r\tau_d$ |
| $\mathbf{\Sigma_H}$ | $K^2L^2N_c(\tau_p+\tau_d)$ |
| $\langle \mathbf{H} \rangle$ | $KLN_c(KL+N_r)(\tau_p+\tau_d)$ |
| $\langle \mathbf{Z}^{(d)} \rangle$ | $KN_c^2N_r\tau_d$ |
| $\langle \mathbf{Z} \rangle$ | $KLN_cN_r(\tau_p+\tau_d)$ |

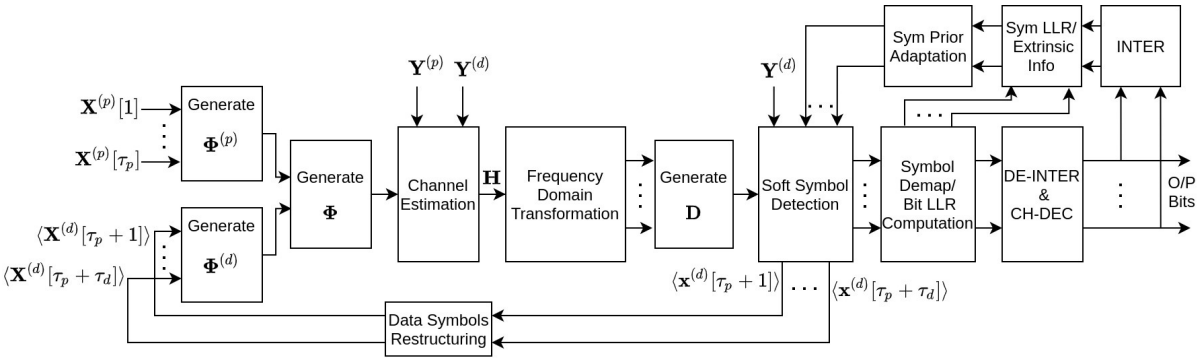


Figure 3.5: Quantized VB Iterative Channel Estimation and Soft Symbol Decoding with data prior adaptation.

is linear in the number of data symbols $\tau_d$, unlike maximum likelihood approaches where the complexity grows exponentially with $\tau_d$. The complexity of the conventional MIMO-OFDM MMSE channel estimation algorithm scales cubically with the number of users and the channel length, whereas the per-iteration complexity of BiGAMP scales linearly with the pilot and data symbols, subcarriers, number of users and channel length, and returns hard decisions of the data symbols. Note that, in order to speed up computations, we can precompute the PDF and CDF of a standard normal variable and store them in a lookup table.

## 3.8 Simulation Results

In the following subsections, we first evaluate the Bayesian and hybrid CRLBs developed in Sec. 3.2, and then study the NMSE and coded BER performance of the iterative VB channel estimation and soft symbol decoding algorithm.

### 3.8.1 Cramér-Rao Lower Bounds

In this subsection, we evaluate the Bayesian and hybrid CRLBs on the NMSE of any estimator of a compressible signal using quantized measurements. We compute the CRLBs for the recovery of joint sparse vectors from compressive measurements acquired using a random underdetermined measurement matrix $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$, whose entries are i.i.d. and complex normal distributed with mean 0 and variance 1. We generate each column of a compressible signal $\mathbf{X} \in \mathbb{C}^{N \times L}$ by sampling from a complex normal distribution with mean $\mathbf{0}$ and precision matrix $\mathbf{P} = \text{diag}(\boldsymbol{\alpha})$. The precision matrix is generated using a Gamma distribution with shape and rate parameters $a$ and $r$, respectively. We show the decay profile of the sorted magnitudes of compressible signals of length $N = 512$, generated using different shape and rate parameters, in Fig. 3.6. The rapid decay of the coefficients shows that the signals are compressible. We fix $N = 512$, $T = 20$, and set the rate parameter to $10^{-8}$ for all the simulations in this subsection. We define the SNR as $1/\sigma_w^2$. In this case, since the support set of the jointly compressible signals contains all the indices, the support-aware CRLB and Bayesian CRLB coincide. Hence, we evaluate only the Bayesian and hybrid CRLBs.

Figure 3.7 shows the NMSE of the VB algorithm for the unquantized (labeled "UQ") and 3 bits quantization (labeled "3 bits") cases, and the Bayesian (labeled "BCRLB") and hybrid CRLBs (labeled "HCRLB"), as a function of the shape parameter of the Gamma hyperprior of the precision matrix. We set the number of measurements to 250 and the SNR to 40 dB. In the case of hybrid CRLB, we know the generative model of the compressible signal, which provides extra information. Hence, the hybrid CRLB is a tight lower bound on the performance of the VB algorithm. The Bayesian CRLB only uses the parameters for the Gamma hyperprior and is therefore looser. Also, the Bayesian

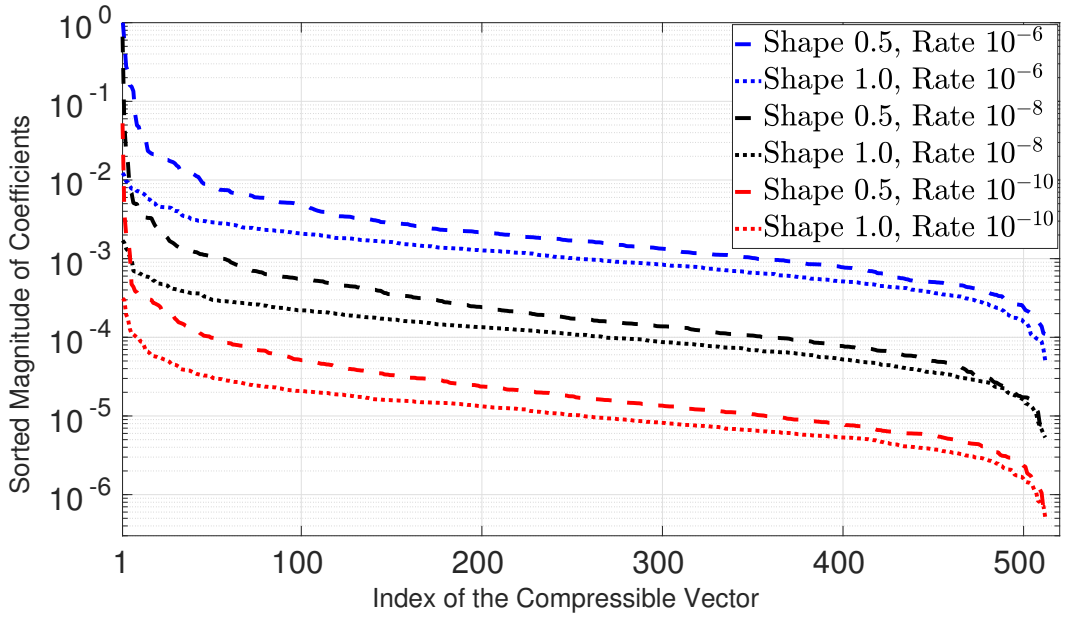


Figure 3.6: Magnitude decay profile of the sorted magnitudes of i.i.d. samples drawn from a complex normal distribution parameterized by a Gamma distributed precision matrix.

CRLB for the unquantized and 3 bits almost overlap on each other, whereas there is a small gap between the two in the case of hybrid CRLB. As the shape parameter increases, the peak value of the compressible signal decreases, and hence the compressibility of the signal decreases. Due to this, the NMSE and the CRLBs both increase with the shape parameter.

Figure 3.8 shows the NMSE and hybrid CRLB as a function of the number of measurements for the unquantized and $\{1, 2, 3\}$ bits quantized cases. We do not include the Bayesian CRLB in this figure to avoid clutter. The shape parameter is set to 0.55 and the SNR to 30 dB. We observe that the gap between the NMSE and CRLB for the 1-bit quantization is higher compared to the $\{2, 3\}$ bits and the unquantized cases. This shows that having only sign measurements leads to severe performance loss due to the large quantization noise, which results in higher NMSE. We also see that the gap between the unquantized case and the 3 bits quantization is very small, which empirically shows that a 3 bit quantizer provides a good trade-off between the performance and system complexity. We thus set the ADC resolution to 3 bits in all the further simulations, unless specified

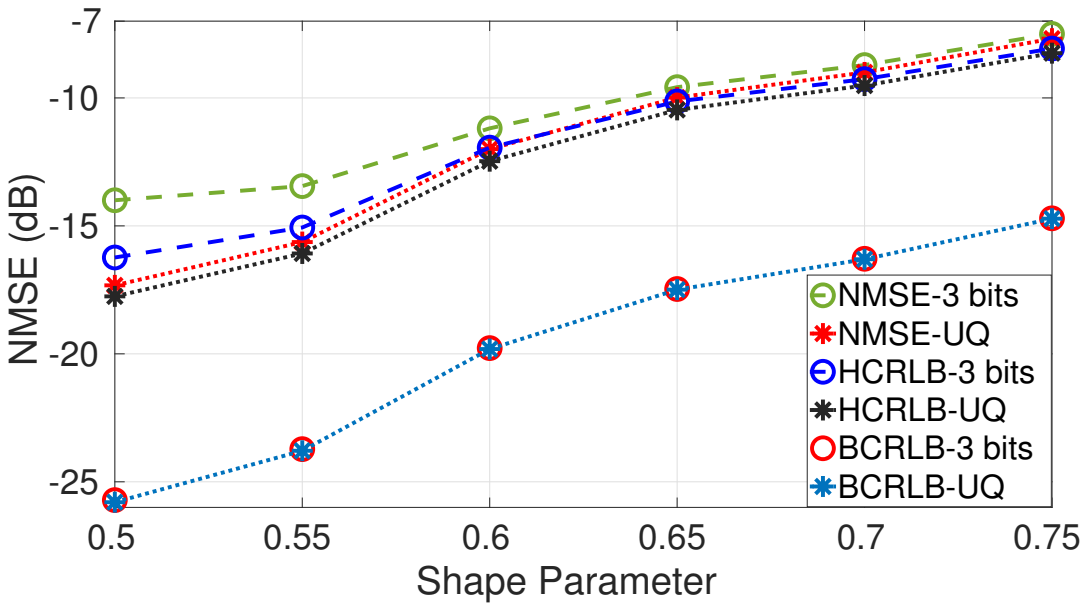


Figure 3.7: NMSE (dB) vs. the shape parameter of the Gamma hyperprior of the precision matrix, with $M = 250$, $N = 512$, $T = 20$, SNR $= 40$ dB.

otherwise.

### 3.8.2 QVB Channel Estimation and Soft Symbol Decoding

In this subsection, we evaluate the NMSE and coded BER performance of the iterative VB channel estimation and soft symbol decoding algorithm. We also study the impact of the various system parameters on the support-aware Bayesian CRLB derived in Section 3.2. The data bits are generated i.i.d. from a uniform distribution. Each UE's data bits are encoded with an LDPC channel code from 3GPP 5G NR specifications [103]. We use the parity check matrix from LDPC base graph 0 with a lifting size $Z_c$ set to 8 and set index 0, which results in 176 message bits and 544 coded bits per block. The coded bits are interleaved by a random interleaver which is known to both the UE and the BS, mapped to 4-QAM constellation of unit energy, OFDM modulated, and transmitted over frequency-selective wireless channels.[3] We assume that the LSFCs between the $k^{\text{th}}$ UE

[3]With higher order constellations, the performance is similar, with an expected shift in the SNR required to achieve a given coded BER.

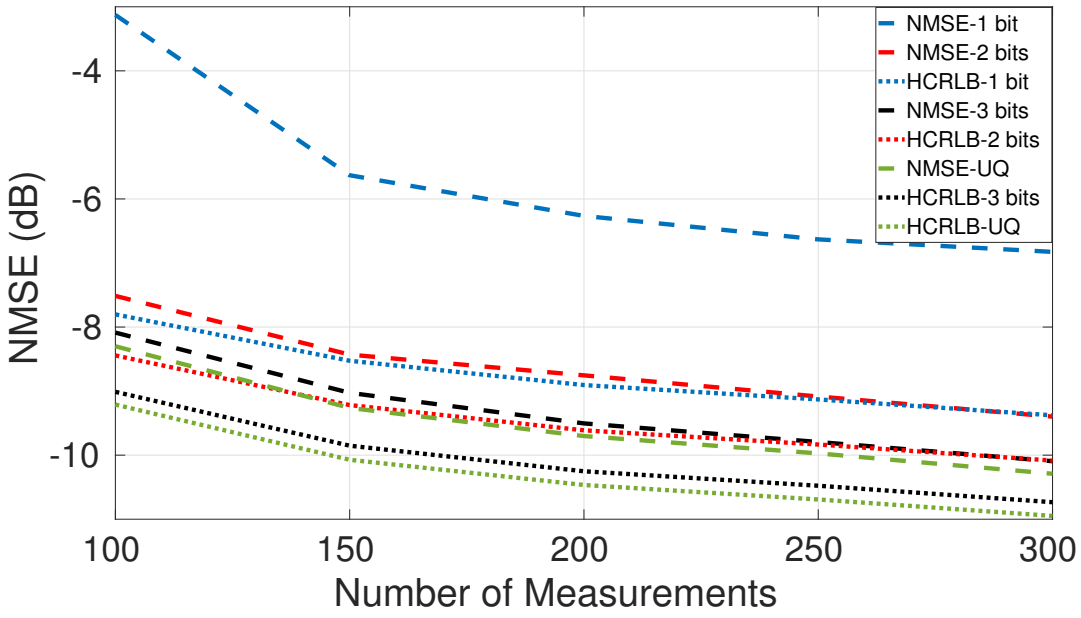


Figure 3.8: NMSE (dB) as a function of the number of quantized measurements. $N = 512$, $T = 20$, shape $= 0.55$, SNR $= 30$ dB.

and the BS antennas are the same due to the close spacing between the antennas compared to the BS-UE distance. The UEs adopt path loss inversion based transmit power control that compensates for the LSFCs, and therefore we set them to 1 in all our simulations.[4] Each tap of the frequency-selective channels is i.i.d. circularly symmetric complex normal distributed with mean 0 and variance 1. We include the details of the system parameters used for simulations in the captions of each simulation plot. We define the SNR as $1/\sigma_w^2$. We use the Frobenius norm of the difference between the channel estimates (and estimates of the unquantized received data symbols) in consecutive iterations as the stopping condition for the VB channel estimation (and soft symbol decoding) procedures. We set the maximum number of iterations for VB channel estimation and data detection algorithms to 25, and the total number of iterations in Algorithm 8 to 4.

Fig. 3.9 compares the CRLB derived for a quantized system with the analytical CRLB for an unquantized system (labeled UQ, from [119].) We see that, as the ADC resolution

[4]The LFSCs can be estimated at the UEs, for example, using the synchronization signals that are periodically transmitted by the BS.

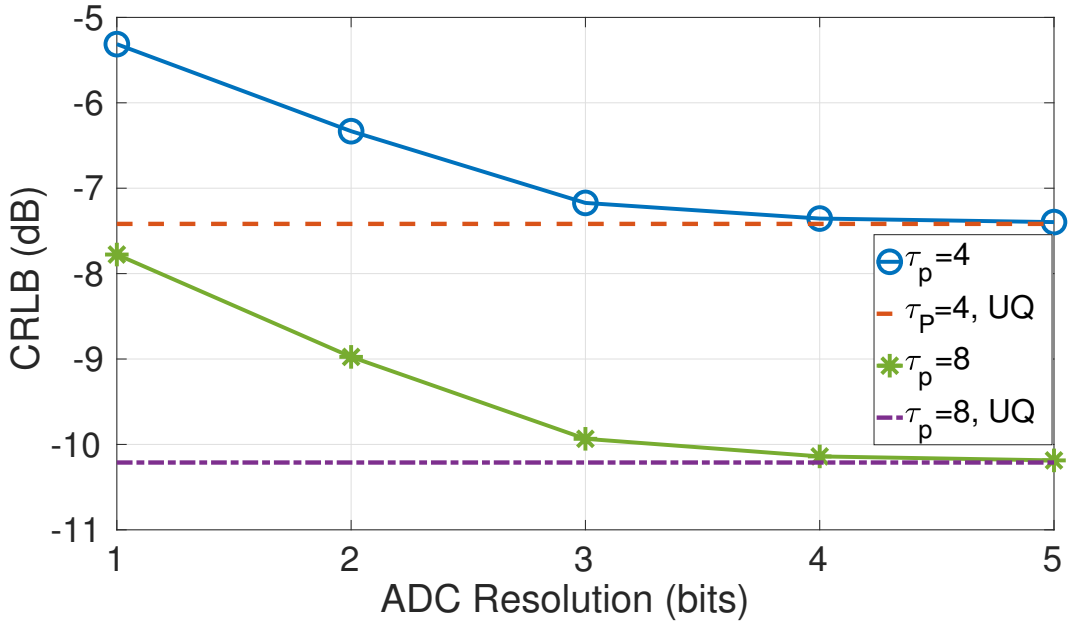


Figure 3.9: CRLB (dB) as a function of ADC resolution (bits) for $N_r = 64$, $K = 16$, $L = 64$, $L_{\text{sp}} = 8$, SNR = 1 dB.

increases, the gap between the quantized and unquantized Bayesian CRLBs decreases, and the bounds meet beyond an ADC resolution of 4 bits. In fact, if the ADC resolution is infinite, our CRLB matches with the analytical expression derived for an unquantized system. Therefore, our derived Bayesian CRLB captures the effect of low-resolution ADCs well, and can serve as a benchmark to evaluate the NMSE performance of any estimator in a quantized setup.

Fig. 3.10 shows the NMSE performance of the quantized VB algorithm and the Bayesian CRLB when $N_r = 40$, $K = 10$, $L_{sp} = 8$, and 3 bits quantization. We observe that when $\tau_p = 1$ and $\tau_p = 4$ OFDM symbols, the NMSE of VB overlaps with the Bayesian CRLB beyond 4 dB and $-4$ dB SNR, respectively. At low SNRs and $\tau_p = 1$, the gap between the Bayesian CRLB and the NMSE of VB is slightly more than that at high SNRs, which can be attributed to the fact that there can be support recovery errors in VB due to high noise. In our wireless communication application, we typically operate at medium to high SNRs, where the Bayesian CRLB and VB almost overlap. Therefore, the Bayesian CRLB serves as a good benchmark to characterize the NMSE performance

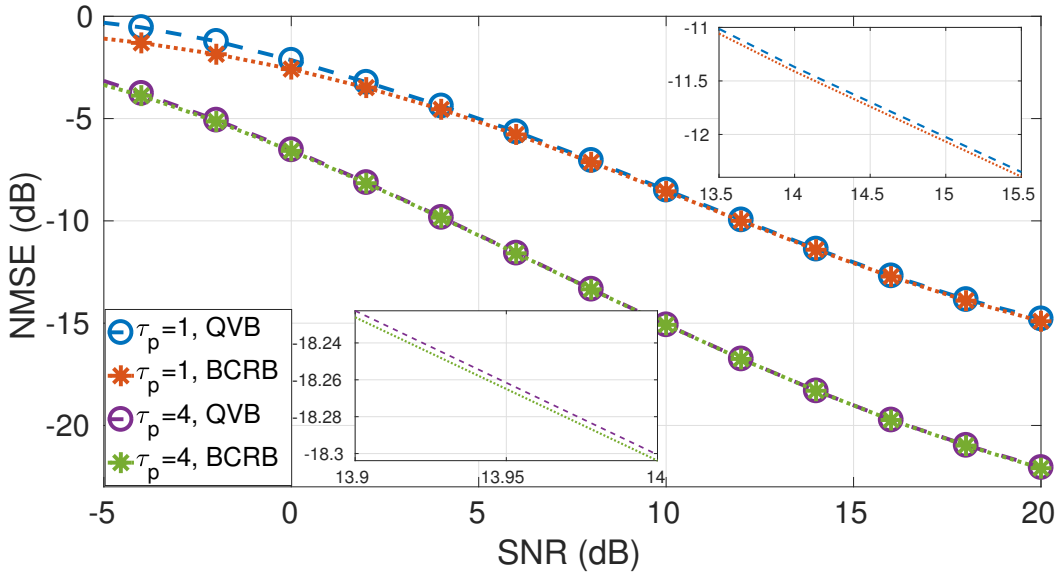


Figure 3.10: NMSE (dB) as a function of SNR (dB) for $N_r = 40$, $K = 10$, $N_c = 256$, $L = 64$, $L_{\text{sp}} = 8$, 3 bits quantization.

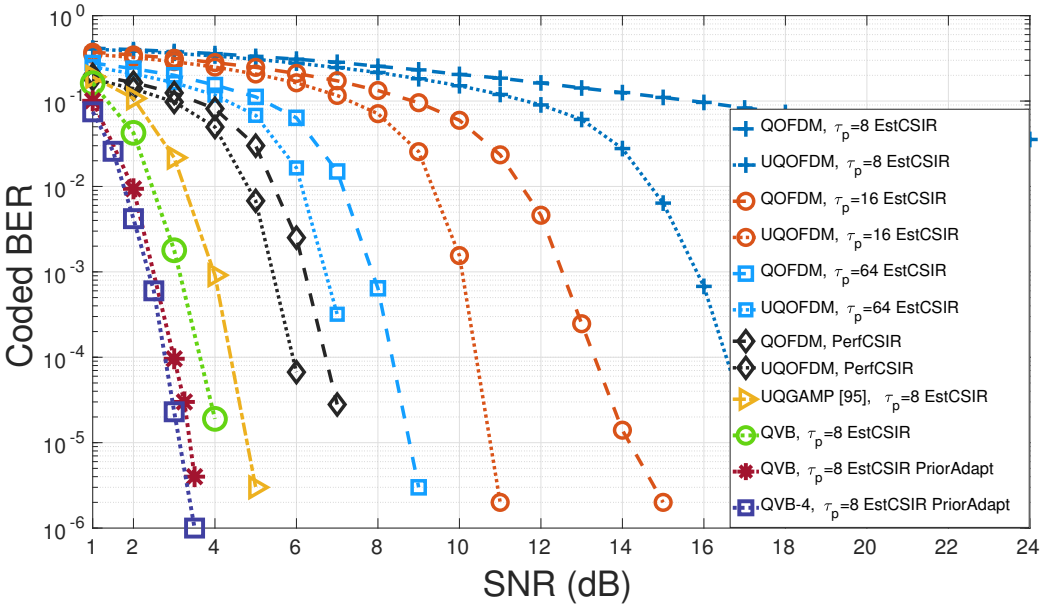


Figure 3.11: Coded BER as a function of SNR (dB) for $N_r = 32$, $K = 8$, $N_c = 256$, $L = 32$, $L_{\text{sp}} = 8$, $\tau_d = 10$, 3-bits quantization. The conventional OFDM receiver (curves labelled "UQOFDM" and "QOFDM") uses the soft-detection procedure from [1].

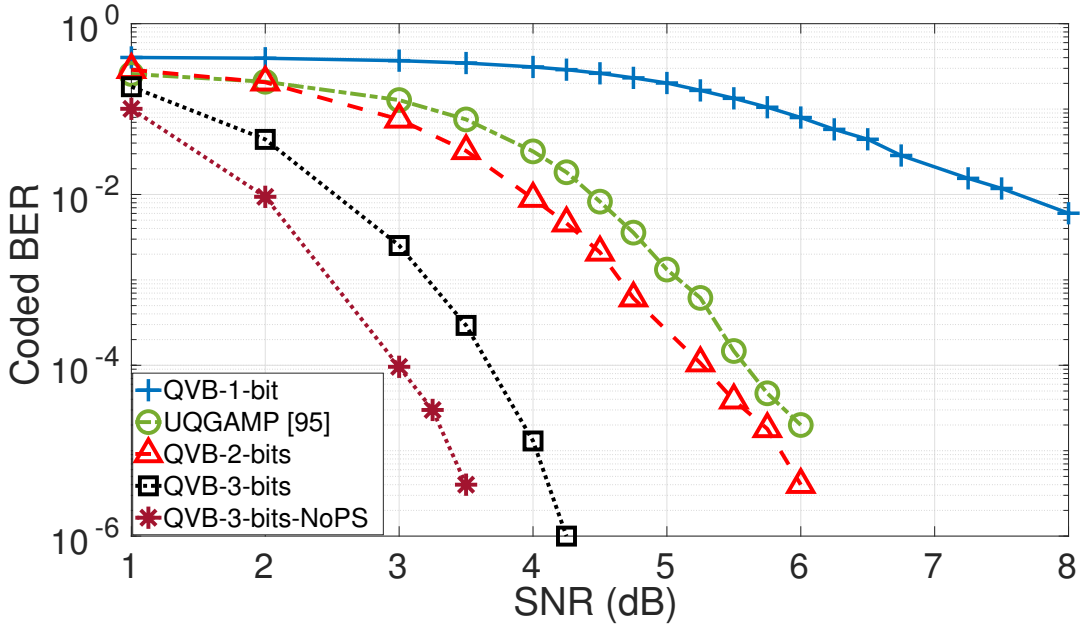


Figure 3.12: Coded BER as a function of SNR (dB) for $N_r = 32$, $K = 8$, $N_c = 256$, $L = 32$, $L_{\text{sp}} = 8$, $\tau_d = 10$ with a square root raised cosine pulse shaping transmit and receive filters.

of a Bayesian-inspired channel estimator.

Fig. 3.11 compares the coded BER performance of the VB algorithm with that of an unquantized joint channel estimation and data detection algorithm based on BiGAMP [95], and MMSE channel estimator and soft-detector [1]. For the quantized MMSE receiver, we compute the DFT after the quantization, and perform the equalization. We set the number of outer iterations of the iterative channel estimator and soft symbol decoder to 8. An advantage of the VB algorithm is that it can recover the channel with only one pilot OFDM symbol. However, for fair comparison, we set $\tau_p = 8$ because the conventional OFDM receiver cannot estimate the channel in an underdetermined setting. We see that, at a BER of $10^{-4}$, the quantized VB algorithm (labelled "QVB $\tau_p = 8$, EstCSIR") outperforms conventional OFDM receiver with unquantized observations and channels estimated using $\tau_p = 8$ pilot OFDM symbols by around 13 dB. In fact, it *even outperforms* the conventional OFDM receiver with *unquantized observations and perfect CSIR* by 2.5 dB. Moreover, the VB algorithm with only 3 bits quantization performs better

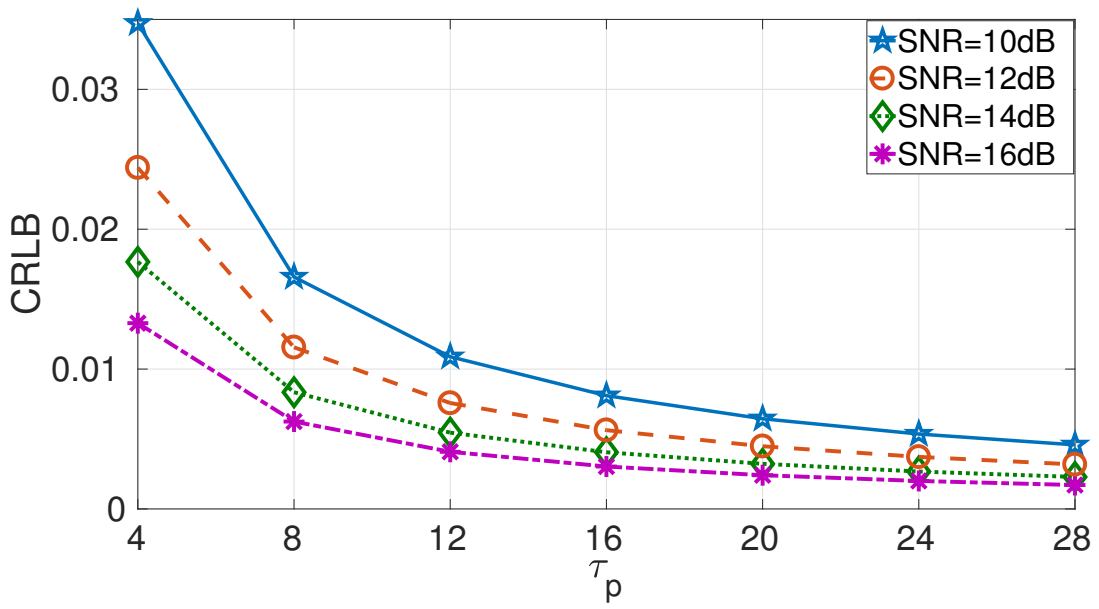


Figure 3.13: CRLB (dB) as a function of $\tau_p$ for $N_r = 32$, $K = 8$, $L = 64$, $L_{\text{sp}} = 16$, $N_c = 256$, 3-bits quantization.

than an unquantized BiGAMP by around 1 dB. This shows the importance of directly inferring the posterior distributions of the data symbols. We also see that the VB algorithm with data prior adaptation performs better than the VB algorithm without any adaptation in this scenario by more than 0.5 dB. Finally, the VB algorithm with 4-bits quantization (labelled "QVB-4 $\tau_p = 4$ EstCSIR PriorAdapt") is only marginally better than the VB with 3 bits quantization. Therefore, a 3 bits quantizer is sufficient to achieve good performance in a low-resolution ADC based MIMO-OFDM system.

Fig. 3.12 compares the coded BER of the VB algorithm for the ADC resolution set to $\{1, 2, 3\}$ bits with square root raised cosine transmit and receive pulse shaping filters. The roll-off factors for the transmit and receive filters are set to 0.3. The system bandwidth is set to 2 GHz, so the sampling period $T_s$ is 0.5 ns. We set the cyclic prefix length to the maximum delay spread of $L = 32$ symbols. The number of nonzero taps $L_{sp}$ is set to 8, with the corresponding delays generated uniformly at random between 0 and $(L-2)T_s$. The channel gains of the nonzero taps are i.i.d. complex normal with zero mean and unit variance. For comparison, we include the quantized VB algorithm with ideal pulse shaping filters (labeled "QVB-3-bits-NoPS") and the BiGAMP algorithm [95] with unquantized

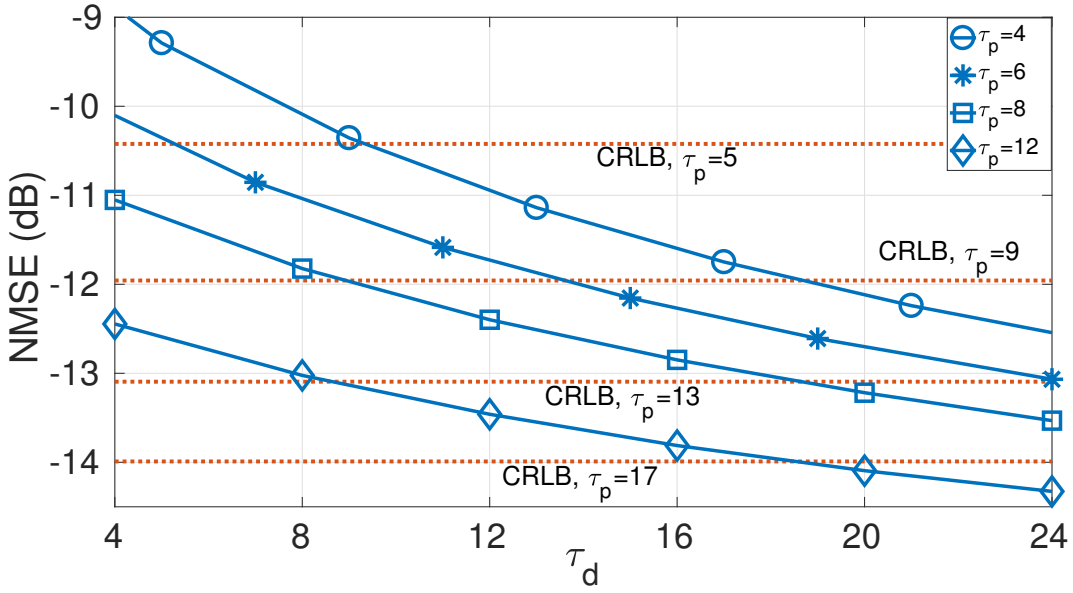


Figure 3.14: NMSE (dB) as a function of $\tau_d$ for $N_r = 64$, $K = 16$, $L = 64$, $L_{\text{sp}} = 8$, $N_c = 256$, SNR = 1 dB, 3-bits quantization.

pilot and data received signals (labeled "UQGAMP"). Although the pulse shaping filters introduce inter-symbol interference and noise correlation, it only results in a marginal performance loss of around $0.5 - 0.7$ dB in the VB algorithm with 3 bits quantization. This shows that the VB algorithm is robust to non-idealities. Moreover, VB algorithm with only 2 bits quantization outperforms UQGAMP.

Fig. 3.13 shows the CRLB as a function of the pilot length $\tau_p$ for various values of SNR. As $\tau_p$ increases, the CRLB decreases, which is due to the increase in the number of observations. The slope of the CRLB curves decreases as $\tau_p$ increases, and asymptotically becomes zero, which follows the law of diminishing returns. That is, if we vary only $\tau_p$ by fixing all the other parameters, we do not see any significant performance improvement beyond a point. This is because, irrespective of the number of measurements, the quantization noise floor limits the improvement obtainable by increasing $\tau_p$. Also, the value of the threshold decreases as SNR increases, which shows that when the noise power is low, we can potentially achieve better spectral efficiency with fewer number of pilots.

Figures 3.14 and 3.15 show the NMSE (dB) and coded BER, respectively, as a function of the data duration, $\tau_d$, for the iterative VB algorithm. We also plot the CRLB in

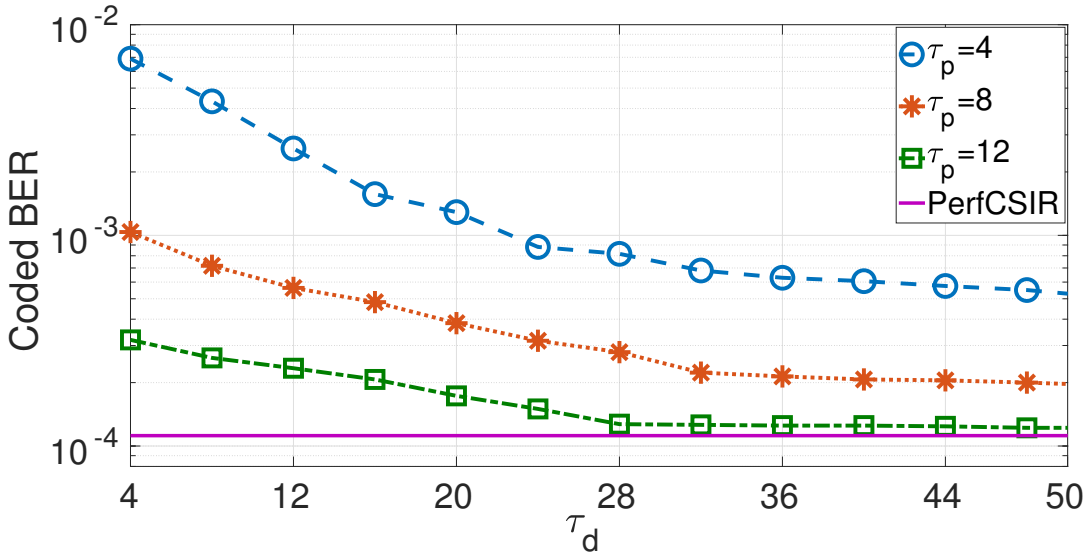


Figure 3.15: Coded BER as a function of $\tau_d$ for $N_r = 64$, $K = 16$, $L = 64$, $L_{\text{sp}} = 8$, SNR=0 dB, 3-bits quantization.

Fig. 3.14. The NMSE decreases with $\tau_d$, as expected, due to the increase in the number of virtual pilot symbols. In fact, the NMSE goes below the CRLB beyond a particular $\tau_d$ in all the $\tau_p$ configurations. For e.g., when $\tau_p$ is 12, the NMSE is around $-13$ dB for $\tau_d = 8$, whereas the CRLB of $-13$ dB is achieved at $\tau_p$ set to 13. That is, the VB channel estimator can attain an NMSE even lower than the CRLB computed using a larger training overhead, since it uses the data symbols as virtual pilots. On the other hand, when $\tau_p$ is 12 and $\tau_d$ is 5, the VB algorithm achieves an NMSE of around $-12.5$ dB, which is higher than the CRLB when $\tau_p$ is 17 by around 1.5 dB. Thus, the NMSE of the channel estimator is higher than the CRLB computed using $\tau_p + \tau_d$ as the pilot duration. This is because the data symbols are also estimated using the received symbols. Nonetheless, the iterative data-aided channel estimation and channel estimation assists in reducing the training overhead and increasing the spectral efficiency. Further, the slope of the NMSE curves decreases with $\tau_d$, which reiterates our observation in Fig. 3.13 about the error floor due to the quantization noise.

In Fig. 3.15, we see that, as $\tau_p$ increases, the coded BER decreases due to better channel estimation performance. We also include the coded BER performance for a quantized VB

soft symbol decoding algorithm with perfect CSIR, which serves as a lower bound for the iterative quantized VB channel estimator and soft symbol decoder. We see that, when $\tau_p$ is 12, and $\tau_d$ is greater than 28, our iterative VB algorithm almost meets the performance of the perfect CSIR case. Given the coherence interval, such studies can guide system designers to configure $\tau_p$ (and thus $\tau_d$) and obtain the same BER performance as with perfect CSIR, or to choose $\tau_p$ to attain the right trade off between training overhead and data duration and thereby achieve maximal spectral efficiency.

## 3.9 Conclusions

In this chapter, we derived the Bayesian, hybrid, and support-aware CRLBs for an estimator of a compressible signal using quantized lower dimensional measurements. Next, we developed a pilot-based channel estimator and a soft symbol decoder using a VB framework, which directly infers the posterior distributions of the channel and data given the quantized received signals. We utilized the posterior statistics of the decoded data symbols to develop an iterative VB data-aided channel estimator and soft symbol decoder. We marginally improved the performance by proposing a variant of the iterative algorithm that used the posterior bit LLRs from the channel decoder for data prior adaptation. We benchmarked the NMSE performance of the VB estimator with that of the derived Bayesian CRLB, and numerically showed that it is efficient. We also evaluated the NMSE and coded BER performances of the iterative VB channel estimator and soft symbol decoder, and compared with the state-of-the-art. Finally, we provided interesting insights into the impact of various parameters on the system performance. Future work could consider extending these ideas to millimeter-wave channels exploiting spatial sparsity, or account for carrier frequency and timing offsets across users by modeling them using latent variables that are estimated using the VB framework, and so on.

## 3.10 Proofs

### 3.10.1 Proof of Theorem 1

The conditional probability distribution of $\widetilde{\mathbf{Y}}$ given $\widetilde{\mathbf{X}}$, $\widetilde{\mathbf{\Phi}}$ and $\widetilde{\sigma}_w$ is

$$\begin{aligned} p(\widetilde{\mathbf{Y}}|\widetilde{\mathbf{X}};\widetilde{\mathbf{\Phi}},\widetilde{\sigma}_w^2) &= \prod_{\ell=1}^{T} p(\widetilde{\mathbf{y}}_\ell|\widetilde{\mathbf{x}}_\ell;\widetilde{\mathbf{\Phi}},\widetilde{\sigma}_w^2) \\ &= \prod_{\ell=1}^{T} \int_{\widetilde{\mathbf{z}}_\ell\in\mathbb{R}^{\widetilde{M}}} p(\widetilde{\mathbf{y}}_\ell|\widetilde{\mathbf{z}}_\ell) p(\widetilde{\mathbf{z}}_\ell|\widetilde{\mathbf{x}}_\ell;\widetilde{\mathbf{\Phi}},\widetilde{\sigma}_w^2)\,\mathrm{d}\widetilde{\mathbf{z}}_\ell \\ &= \prod_{\ell=1}^{T}\prod_{m=1}^{\widetilde{M}} B_{m\ell}, \end{aligned} \tag{3.60}$$

where

$$B_{m\ell} \triangleq \int_{\frac{\tilde{z}_{m\ell}^{(\mathrm{lo})}-\sum_{n=1}^{\widetilde{N}}\widetilde{\Phi}_{mn}\widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}}^{\frac{\tilde{z}_{m\ell}^{(\mathrm{hi})}-\sum_{n=1}^{\widetilde{N}}\widetilde{\Phi}_{mn}\widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}} \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{\widetilde{z}_{m\ell}^2}{2}\right]\,\mathrm{d}\widetilde{z}_{m\ell} \tag{3.61}$$

and $\widetilde{z}_{m\ell}^{(\mathrm{lo})}$ and $\widetilde{z}_{m\ell}^{(\mathrm{hi})}$ are the lower and upper quantization thresholds for the $(m,\ell)^{\mathrm{th}}$ entry of $\widetilde{\mathbf{Y}}$, respectively. Also, $\widetilde{\Phi}_{mn}$ and $\widetilde{x}_{n\ell}$ denote the $(m,n)^{\mathrm{th}}$ and $(n,\ell)^{\mathrm{th}}$ entries of $\widetilde{\mathbf{\Phi}}$ and $\widetilde{\mathbf{X}}$, respectively.

Note that, since we estimate $\widetilde{N}T$ parameters in total, the FIM is block diagonal matrix of size $\widetilde{N}T\times\widetilde{N}T$, with $T$ blocks each of size $\widetilde{N}\times\widetilde{N}$. Computing it requires the gradient and Hessian of the joint probability distribution w.r.t. $\mathbf{x}_\ell\ \forall\ell$. Since the columns of $\widetilde{\mathbf{X}}$ are independent of each other, we express the logarithm of the joint distribution using the chain rule as shown in (3.62).

$$\begin{aligned} \log p(\widetilde{\mathbf{Y}},\widetilde{\mathbf{X}},\widetilde{\mathbf{P}};\widetilde{\mathbf{\Phi}},\widetilde{\sigma}_w^2) &= \log p(\widetilde{\mathbf{Y}}|\widetilde{\mathbf{X}};\widetilde{\mathbf{\Phi}},\widetilde{\sigma}_w^2) + \log p(\widetilde{\mathbf{X}};\widetilde{\mathbf{P}}) + \log p(\boldsymbol{\alpha};a,b), \\ &\propto \sum_{\ell=1}^{T}\sum_{m=1}^{\widetilde{M}} \log \int_{\frac{\tilde{z}_{m\ell}^{(\mathrm{lo})}-\sum_{n=1}^{\widetilde{N}}\widetilde{\Phi}_{mn}\widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}}^{\frac{\tilde{z}_{m\ell}^{(\mathrm{hi})}-\sum_{n=1}^{\widetilde{N}}\widetilde{\Phi}_{mn}\widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}} \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{\widetilde{z}_{m\ell}^2}{2}\right]\,\mathrm{d}\widetilde{z}_{m\ell} - \sum_{\ell=1}^{T}\frac{\widetilde{\mathbf{x}}_\ell^T\widetilde{\mathbf{P}}\widetilde{\mathbf{x}}_\ell}{2}. \end{aligned} \tag{3.62}$$

In (3.62), we omit the terms that do not depend on $\widetilde{\mathbf{X}}$ for brevity. We can verify that the joint probability distribution in (3.62) fall into the exponential family of distributions which satisfies the regularity conditions in Sec. 5.2.3 of [125]. We apply Leibniz integral rule to compute the first and second derivatives of $\log B_{ml}$ with respect to (w.r.t.) $\widetilde{x}_{k\ell}$ and $\widetilde{x}_{j\ell}$ shown in (3.63) and (3.64), respectively,

$$\begin{aligned}&\frac{\partial}{\partial \widetilde{x}_{k\ell}} \log B_{m\ell} \\ &= -\frac{\widetilde{\Phi}_{mk}}{\sqrt{2\pi} B_{m\ell} \widetilde{\sigma}_w} \left[ \exp\left( -\frac{(\widetilde{z}_{m\ell}^{(\mathrm{hi})} - \sum_{n=1}^{\widetilde{N}} \widetilde{\Phi}_{mn} \widetilde{x}_{n\ell})^2}{2\widetilde{\sigma}_w^2} \right) - \exp\left( -\frac{(\widetilde{z}_{m\ell}^{(\mathrm{lo})} - \sum_{n=1}^{\widetilde{N}} \widetilde{\Phi}_{mn} \widetilde{x}_{n\ell})^2}{2\widetilde{\sigma}_w^2} \right) \right].\end{aligned} \tag{3.63}$$

$$-\frac{\partial^2}{\partial \widetilde{x}_{j\ell} \widetilde{x}_{k\ell}} \log B_{m\ell} = \frac{\widetilde{\Phi}_{mk} \widetilde{\Phi}_{mj}}{\widetilde{\sigma}_w^2} \left[ \frac{\widetilde{\eta}_{m\ell}^{(\mathrm{hi})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - \widetilde{\eta}_{m\ell}^{(\mathrm{lo})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})}{F(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - F(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})} + \left( \frac{f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})}{F(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - F(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})} \right)^2 \right], \tag{3.64}$$

where

$$\widetilde{\eta}_{m\ell}^{(\mathrm{hi})} \triangleq \frac{\widetilde{z}_{m\ell}^{(\mathrm{hi})} - \sum_{n=1}^{\widetilde{N}} \widetilde{\Phi}_{mn} \widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}, \tag{3.65}$$

$$\widetilde{\eta}_{m\ell}^{(\mathrm{lo})} \triangleq \frac{\widetilde{z}_{m\ell}^{(\mathrm{lo})} - \sum_{n=1}^{\widetilde{N}} \widetilde{\Phi}_{mn} \widetilde{x}_{n\ell}}{\widetilde{\sigma}_w}. \tag{3.66}$$

In (3.64), $f(\cdot)$ and $F(\cdot)$ denote the PDF and CDF of a standard normal random variable, respectively. Writing in matrix form, the $\ell^{\mathrm{th}}$ diagonal block of the FIM, denoted by $\widetilde{\mathbf{M}}_\ell(a, b, \widetilde{\sigma}_w^2)$, is shown in (3.67).

$$\begin{aligned}&\widetilde{\mathbf{M}}_\ell(a, b, \widetilde{\sigma}_w^2) = \mathbb{E}\left[ -\frac{\partial^2}{\partial \widetilde{\mathbf{x}}_\ell \widetilde{\mathbf{x}}_\ell^T} \log p(\widetilde{\mathbf{Y}}, \widetilde{\mathbf{X}}; \widetilde{\mathbf{P}}, \widetilde{\mathbf{\Phi}}, \widetilde{\sigma}_w^2) \right] \\ &= \widetilde{\mathbf{\Phi}}^T \mathtt{diag}\left( \frac{1}{\widetilde{\sigma}_w^2} \mathbb{E}\left[ \frac{\widetilde{\eta}_{m\ell}^{(\mathrm{hi})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - \widetilde{\eta}_{m\ell}^{(\mathrm{lo})} f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})}{F(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - F(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})} + \left( \frac{f(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - f(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})}{F(\widetilde{\eta}_{m\ell}^{(\mathrm{hi})}) - F(\widetilde{\eta}_{m\ell}^{(\mathrm{lo})})} \right)^2 \right] \right)_{m=1}^{\widetilde{M}} \widetilde{\mathbf{\Phi}} + \mathbb{E}[\widetilde{\mathbf{P}}].\end{aligned} \tag{3.67}$$

### 3.10.2 Proof of Lemma 1

As $\delta \to 0$, both (3.13) and (3.14) become indeterminate forms. Further, both the numerators and denominators in the left hand sides of (3.13) and (3.14) are differentiable at 0. Applying L'Hôpital's and Leibniz integral rules (for differentiating the denominators), we get the right hand sides in (3.13) and (3.14). □

### 3.10.3 Proof of Lemma 2

The proof follows by using the lower and upper thresholds of the 1-bit quantizer as follows:

$$\widetilde{z}_{m\ell}^{(\mathrm{lo})} = \begin{cases} 0 & \text{if } \widetilde{y}_{m\ell} = +1 \\ -\infty & \text{if } \widetilde{y}_{m\ell} = -1 \end{cases} \tag{3.68}$$

and

$$\widetilde{z}_{m\ell}^{(\mathrm{hi})} = \begin{cases} \infty & \text{if } \widetilde{y}_{m\ell} = +1 \\ 0 & \text{if } \widetilde{y}_{m\ell} = -1. \end{cases} \tag{3.69}$$

Substituting (3.68) and (3.69) in (3.7), after straightforward algebraic manipulation and using the facts that $F(\infty) = 1$, $F(-\infty) = 0$ and $F(\widetilde{\eta}_{m\ell}) = 1 - F(-\widetilde{\eta}_{m\ell})\ \forall\ m, \ell$, we get (3.16). □

### 3.10.4 Proof of Lemma 3

To obtain the approximate posterior distribution $q_{\mathbf{H}}(\mathbf{H})$, we first compute the approximate posterior distribution of each of the factors $q_{\mathbf{h}_n}(\mathbf{h}_n)$, $n = \{1, \ldots, N_r\}$. We calculate the expectation of the joint distribution in (3.38) with respect to the approximate posterior distributions

$\{q_{\mathbf{Z}}(\mathbf{Z}^{(p)}), q_{\mathbf{h}_1}(\mathbf{h}_1), \ldots, q_{\mathbf{h}_{n-1}}(\mathbf{h}_{n-1}), q_{\mathbf{h}_{n+1}}(\mathbf{h}_{n+1}), \ldots, q_{\mathbf{h}_{N_r}}(\mathbf{h}_{N_r}), q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})\}$ as follows:

$$\ln q_{\mathbf{h}_n}(\mathbf{h}_n) = \left\langle \ln\, p(\mathbf{Y}^{(p)} \,|\, \mathbf{Z}^{(p)}) + \ln\, p(\mathbf{Z}^{(p)} \,|\, \mathbf{H}; \boldsymbol{\Phi}^{(p)}, \sigma_w^2) + \ln\, p(\mathbf{H}|\mathbf{P}) + \ln\, p(\boldsymbol{\alpha}; a, b) \right\rangle \tag{3.70}$$

$$\propto \langle \ln\, p(\mathbf{Z}^{(p)} \,|\, \mathbf{H}; \boldsymbol{\Phi}^{(p)}, \sigma_w^2) + \ln\, p(\mathbf{H}|\mathbf{P}) \rangle, \tag{3.71}$$

where $\langle \cdot \rangle$ denotes the expectation operation w.r.t. the approximate posterior distributions of all the latent variables except $q_{\mathbf{h}_n}(\mathbf{h}_n)$. We obtain (3.71) from (3.70) by including only the terms that do not depend on $\mathbf{h}_n$ as proportionality constants such that $q_{\mathbf{h}_n}(\mathbf{h}_n)$ becomes a probability distribution. Simplifying (3.71) by separating the terms that depend only on $\mathbf{h}_n$, we get

$$\begin{aligned} &\ln q_{\mathbf{h}_n}(\mathbf{h}_n) \\ &\quad \propto \left\langle -\frac{1}{\sigma_w^2} \left( \mathbf{h}_n^H {\boldsymbol{\Phi}^{(p)}}^H \boldsymbol{\Phi}^{(p)} \mathbf{h}_n - {\mathbf{z}_n^{(p)}}^H \boldsymbol{\Phi}^{(p)} \mathbf{h}_n - \mathbf{h}_n^H {\boldsymbol{\Phi}^{(p)}}^H \mathbf{z}_n^{(p)} \right) - \mathbf{h}_n^H \mathbf{P} \mathbf{h}_n \right\rangle \qquad (3.72) \\ &\quad \propto -\left( \mathbf{h}_n^H \left( \frac{1}{\sigma_w^2} {\boldsymbol{\Phi}^{(p)}}^H \boldsymbol{\Phi}^{(p)} + \langle \mathbf{P} \rangle \right) \mathbf{h}_n - \frac{1}{\sigma_w^2} \left\langle \mathbf{z}_n^{(p)} \right\rangle^H \boldsymbol{\Phi}^{(p)} \mathbf{h}_n - \frac{1}{\sigma_w^2} \mathbf{h}_n^H {\boldsymbol{\Phi}^{(p)}}^H \left\langle \mathbf{z}_n^{(p)} \right\rangle \right), \qquad (3.73) \end{aligned}$$

where $\left\langle \mathbf{z}_n^{(p)} \right\rangle$ is the mean of $q_{\mathbf{z}_n}\left(\mathbf{z}_n^{(p)}\right)$, $\langle \mathbf{P} \rangle = \mathtt{diag}(\langle \boldsymbol{\alpha} \rangle)$, and $\langle \boldsymbol{\alpha} \rangle$ is the mean of $q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$. Taking exponentials on both sides of (3.73), and by completing the squares, we can deduce from the structure of the resulting expression that $q_{\mathbf{h}_n}(\mathbf{h}_n)$ is complex normal distributed with covariance matrix and mean given by

$$\boldsymbol{\Sigma}_{\mathbf{H}} = \left( \frac{1}{\sigma_w^2} {\boldsymbol{\Phi}^{(p)}}^H \boldsymbol{\Phi}^{(p)} + \langle \mathbf{P} \rangle \right)^{-1}, \tag{3.74}$$

$$\langle \mathbf{h}_n \rangle = \frac{1}{\sigma_w^2} \boldsymbol{\Sigma}_{\mathbf{H}} {\boldsymbol{\Phi}^{(p)}}^H \left\langle \mathbf{z}_n^{(p)} \right\rangle, \tag{3.75}$$

respectively. Note that, the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{H}}$ is independent of $n$. So, we can write the posterior mean of $q_{\mathbf{h}_n}(\mathbf{h}_n)$, $n = \{1, \ldots, N_r\}$ in a matrix form to get (3.44) and (3.45).

$$\boldsymbol{\Sigma}_{\mathbf{H}} = \left( \frac{1}{\sigma_w^2} {\boldsymbol{\Phi}^{(p)}}^H \boldsymbol{\Phi}^{(p)} + \langle \mathbf{P} \rangle \right)^{-1} \qquad \text{and} \qquad \langle \mathbf{h}_n \rangle = \frac{1}{\sigma_w^2} \boldsymbol{\Sigma}_{\mathbf{H}} {\boldsymbol{\Phi}^{(p)}}^H \left\langle \mathbf{z}_n^{(p)} \right\rangle, \tag{3.76}$$

□

### 3.10.5 Proof of Lemma 4

To obtain the approximate posterior distribution $q_{\mathbf{Z}}(\mathbf{Z}^{(p)})$, we first compute the approximate posteriors of each of its factors $q_{\mathbf{z}_n}(\mathbf{z}_n^{(p)})$, $n = \{1, \ldots, N_r\}$. We calculate the expectation of the log of the joint probability distribution in (3.38) with respect to the approximate posterior distributions of all the latent variables except $q_{\mathbf{z}_n}(\mathbf{z}_n^{(p)})$ as follows:

$$\ln q_{\mathbf{z}_n}(\mathbf{z}_n^{(p)}) = \left\langle \ln\, p(\mathbf{Y}^{(p)} \,|\, \mathbf{Z}^{(p)}) + \ln\, p(\mathbf{Z}^{(p)} \,|\, \mathbf{H}; \mathbf{\Phi}^{(p)}, \sigma_w^2) + \ln\, p(\mathbf{H}|\mathbf{P}) + \ln\, p(\boldsymbol{\alpha}; a, b) \right\rangle \tag{3.77}$$

$$\propto \left\langle \ln\, p(\mathbf{y}_n^{(p)} \,|\, \mathbf{z}_n^{(p)}) + \ln\, p(\mathbf{z}_n^{(p)} \,|\, \mathbf{h}_n; \mathbf{\Phi}^{(p)}, \sigma_w^2) \right\rangle \tag{3.78}$$

$$\propto \ln \mathbb{1}(\mathbf{z}_n^{(p)} \in (\mathbf{z}_n^{(\mathrm{lo})}, \mathbf{z}_n^{(\mathrm{hi})})) - \frac{1}{\sigma_w^2} \left\langle \left\| \mathbf{z}_n^{(p)} - \mathbf{\Phi}^{(p)} \mathbf{h}_n \right\|^2 \right\rangle. \tag{3.79}$$

By expanding the second term in (3.79), completing the squares, and taking exponential on both sides, $q_{\mathbf{z}_n}(\mathbf{z}_n^{(p)})$ is truncated complex normal distributed with mean given by

$$\left\langle \mathbf{z}_n^{(p)} \right\rangle = \mathbf{\Phi}^{(p)} \left\langle \mathbf{h}_n \right\rangle + \frac{\sigma_w}{\sqrt{2}} \frac{f\left(\frac{\mathbf{z}_n^{(\mathrm{lo})} - \mathbf{\Phi}^{(p)} \langle \mathbf{h}_n \rangle}{\sigma_w/\sqrt{2}}\right) - f\left(\frac{\mathbf{z}_n^{(\mathrm{hi})} - \mathbf{\Phi}^{(p)} \langle \mathbf{h}_n \rangle}{\sigma_w/\sqrt{2}}\right)}{F\left(\frac{\mathbf{z}_n^{(\mathrm{hi})} - \mathbf{\Phi}^{(p)} \langle \mathbf{h}_n \rangle}{\sigma_w/\sqrt{2}}\right) - F\left(\frac{\mathbf{z}_n^{(\mathrm{lo})} - \mathbf{\Phi}^{(p)} \langle \mathbf{h}_n \rangle}{\sigma_w/\sqrt{2}}\right)}. \tag{3.80}$$

We have included the second order terms of $q_{\mathbf{h}_n}(\mathbf{h}_n)$ as part of the proportionality constant to arrive at (3.80). By writing the posterior means of $q_{\mathbf{z}_n}(\mathbf{z}_n^{(p)})$ in matrix form, we get (3.46). □

### 3.10.6 Proof of Lemma 5

We follow similar steps that are used to compute $q_{\mathbf{Z}}(\mathbf{Z}^{(p)})$ and $q_{\mathbf{H}}(\mathbf{H})$ to obtain $q_{\alpha_k}(\alpha_k)$, $1 \leq k \leq KL$ as follows:

$$\ln q_{\alpha_k}(\alpha_k) = \left\langle \ln\, p(\mathbf{Y}^{(p)} \,|\, \mathbf{Z}^{(p)}) + \ln\, p(\mathbf{Z}^{(p)} \,|\, \mathbf{H}; \mathbf{\Phi}^{(p)}, \sigma_w^2) + \ln\, p(\mathbf{H}|\mathbf{P}) + \ln\, p(\boldsymbol{\alpha}; a, b) \right\rangle \tag{3.81}$$

$$\propto \langle \ln\, p(\mathbf{H}|\mathbf{P}) + \ln\, p(\boldsymbol{\alpha}; a, b)\rangle \tag{3.82}$$

$$\propto (a + N_r - 1)\ln \alpha_k - \Big(b + \sum_{n=1}^{N_r} \langle |h_{kn}|^2 \rangle\Big)\alpha_k. \tag{3.83}$$

From the structure of (3.83), we see that $q_{\boldsymbol{\alpha_k}}(\alpha_k)\ \forall k$ is Gamma distributed with shape and rate parameters given by (3.47). The mean $\langle \alpha_k \rangle$ is computed as $\widetilde{a}_k/\widetilde{b}_k$, where $\widetilde{a}_k$ and $\widetilde{b}_k$ are as defined in (3.47). □

### 3.10.7 Proof of Lemma 6

We obtain the approximate posterior distribution $q_x\left(x_{kt}^{(d)}\right)$ by computing the expectation of the log of the joint distribution with respect to the approximate posterior distributions of all the latent variables except $q_{x_{kt}}$.

$$\begin{aligned}\ln\, q_{x_{kt}}\left(x_{kt}^{(d)}\right) &= \left\langle \ln p\left(\mathbf{Y}^{(d)}|\mathbf{Z}^{(d)}\right) + \ln p\left(\mathbf{Z}^{(d)}|\mathbf{X}^{(d)}, \mathbf{D}, \sigma_w^2\right) + \ln p\left(\mathbf{X}^{(d)}\right)\right\rangle \\ &\propto -\frac{1}{\sigma_w^2}\Big( \|\mathbf{D}_{:,k}\|^2 \ \Big|\ x_{kt}^{(d)}\ \Big|^2 - 2\Re\Big[\mathbf{D}_{:,k}^H\Big( \Big\langle \mathbf{z}_t^{(d)} \Big\rangle - \sum_{\substack{k'=1 \\ k'\neq k}}^{KN_c} \mathbf{D}_{:,k'} \Big\langle x_{k't}^{(d)} \Big\rangle \Big) x_{kt}^{(d)*}\Big]\Big) \\ &\quad + \ln p(x_{kt}^{(d)}). \end{aligned} \tag{3.84}$$

We include all the terms that do not depend on $x_{kt}^{(d)}$ as part of the proportionality constant. Now, we substitute the values of $x_{kt}^{(d)}$ from a discrete constellation set in (3.84) and take exponential on both the sides to get an expression for the probability mass at a constellation point, and normalize it to obtain the approximate posterior probability mass function of $x_{kt}^{(d)}$ given in (3.54). □

# Part 2

# Chapter 4

# Site-Specific Millimeter-Wave Compressive Channel Estimation with Hybrid MIMO Architectures

## 4.1 Introduction

Millimeter-wave (mmWave) wireless communication is one of the potential technologies proposed for the next generation communication systems (5G and beyond) to meet the ever-increasing demand for high data rates. The mmWave frequency spectrum, ranging from 30 GHz to 300 GHz, is attractive because it offers large bandwidths ($\sim$ 2GHz), resulting in very high data rates and low latency. These advantages come at a cost of higher path loss due to several factors, such as blockages and oxygen absorption at mmWave frequencies, which in turn bring several engineering challenges in adopting this technology in commercial wireless communication systems.

A potential solution to overcome this problem is beamforming, which leverages the availability of multiple antennas at the transmitter and receiver. In particular, millimeter wavelengths enable one to accommodate a larger number of antennas into the same physical space, and thereby attain high beamforming gains. However, a fully digital architecture in a MIMO system, i.e., one Radio Frequency (RF) chain per antenna, and one

complex-valued Analog-to-Digital Converter (ADC) per RF chain is less appealing both from commercial and engineering perspectives due to its high cost and energy requirements. Therefore, a hybrid MIMO architecture is proposed in the literature as a potential solution to solve this problem [45].

In a hybrid MIMO system, multiple antennas are connected to an RF chain using a phase shifter network (RF precoder/combiner), and a digital precoder/combiner is employed in the complex baseband side of the transceiver. The RF and digital precoders/combiners are configured by optimizing a system performance metric such as the sum rate or signal to interference noise ratio. Unlike a fully analog architecture, a hybrid architecture allows one to reduce the number of RF chains, while supporting multi-stream and multi-user transmissions. The major challenges then are in estimating the mmWave wireless channel and configuring the RF and digital precoders/combiners based on the channel estimate. The problem is exacerbated by the fact that only the low dimensional RF combined signals at the baseband are available for estimating the channel. Since the system does not have any knowledge of the channel state during the channel estimation phase, the baseband precoders/combiners are set to the identity matrix and random phase shifts are chosen for the RF precoders/combiners.

Mmwave channel estimation in a hybrid MIMO architecture is a well studied problem, and we provide a brief overview of some of the key existing literature here. The simplest channel estimation method in hybrid MIMO systems is the Least Squares (LS)-based approach [126], which is inherited from conventional MIMO [127]. A more refined solution to channel estimation is to exploit both the delay and angular domain sparsity that mmWave channels exhibit. In this approach, the channel estimation problem is formulated as a sparse recovery problem [57]. Such compressive sensing based estimation techniques were first developed for frequency-flat hybrid mmWave MIMO systems [50,128]. Recently, frequency-selective channels with OFDM-based communications leading to a more complex estimation problem have also been considered, with different approaches to exploit the sparse channel characteristics [57, 129, 130]. Several model-based signal processing techniques for mmWave channel estimation under various system settings can be found

in [131–145].

Machine Learning and Artificial Intelligence (ML/AI) have been shown to be powerful tools in diverse areas such as natural language processing, speech processing, and image recognition, where it is challenging to design specific model-based algorithms. However, the impact of ML/AI on the design and optimization of communication systems is yet to be extensively studied, especially under realistic and practically meaningful settings. We aim to address some of the aspects of ML/AI in wireless communications here.

In this chapter, we study the potential advantage of using data-driven approaches for channel estimation in hybrid MIMO systems. The model-cum-data driven algorithm we develop in this chapter was selected as one of the **top three** solutions in the "**ML5G-PHY Channel Estimation Global Challenge 2020**" organized by the International Telecommunication Union (ITU)[1]. It also won the "**Popular Vote**" award in the "ITU Challenge Grand Finale" conference organized in December 2020. Our main goal in this chapter is to present our algorithm, and contrast it with the other two winning solutions for estimating an mmWave channel in a hybrid MIMO system. We utilize the channel training datasets generated using the *Raymobtime* tool to customize our algorithm so that it performs well for a test dataset generated in a similar environment [146].

We provide a brief overview of our solution here. We integrate a fast greedy search with a high-performing Bayesian inference method in the first approach. We develop a multi-level greedy search (MLGS) to learn the sparsifying virtual beamspace dictionary which reduces the dimensionality of the problem. We then use the learned dictionary to estimate the channel using a reduced dimensional sparse Bayesian learning (SBL) method. We finally exploit the delay-domain sparsity learnt from the training dataset to de-noise the estimated channels. We name the algorithm as MLGS-SBL. We compare the NMSE of our algorithm with the other two competing solutions named Pattern-Coupled Sparse Bayesian Learning for Channel Estimation with Dominating Delay Taps (PCSBL-DDT) and Projection Cuts Orthogonal Matching Pursuit (PC-OMP) [63].

[1]https://www.itu.int/en/ITU-T/AI/challenge/2020/Pages/default.aspx

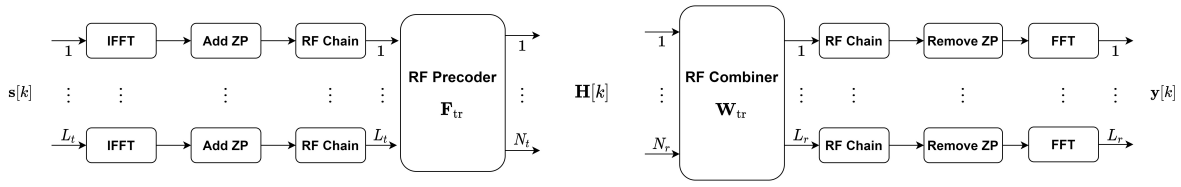


Figure 4.1: mmWave MIMO system based on a hybrid analog-digital architecture.

## Notation

The operator $(\cdot)^*$ represents the conjugate transpose or conjugate for a matrix or a scalar, respectively. $\bar{\mathbf{A}}$, $\mathbf{A}^T$, and $\mathbf{A}^\dagger$ denote the conjugate, transpose, and Moore-Penrose pseudoinverse of a matrix $\mathbf{A}$, respectively. The multivariate complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ is denoted by $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$ and its probability density function (pdf) of a random vector $\mathbf{x}$ is denoted by $\mathcal{CN}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{C})$. $\texttt{blkdiag}(\cdot)$ represents the block-diagonal part of a matrix. $\texttt{diag}(\mathbf{X})$ or $\texttt{diag}(\mathbf{x})$ represents a vector obtained by the diagonal elements of the matrix $\mathbf{X}$ or the diagonal matrix obtained with the elements of $\mathbf{x}$ in the diagonal, respectively. $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of the matrices $\mathbf{A}$ and $\mathbf{B}$. $||\mathbf{A}||_F$ denotes the Frobenius norm of a matrix $\mathbf{A}$. $\langle \mathbf{a}, \mathbf{b} \rangle$ is the inner product of the two vectors $\mathbf{a}$ and $\mathbf{b}$. The trace of a matrix $\mathbf{A}$ is denoted by $\text{tr}(\mathbf{A})$. Tx and Rx denote the transmitter and receiver, respectively. We use $\texttt{vec}(\mathbf{A})$ to vectorize the matrix $\mathbf{A}$ column-wise. $\mathbb{E}[\cdot]$ denotes the expectation.

# 4.2 System Model

We consider a single cell mmWave hybrid MIMO-OFDM system with $N_t$ antennas at the transmitter (Tx) and $N_r$ antennas at the receiver (Rx), as shown in Fig. 4.1. The Tx and Rx are equipped with $L_t$ and $L_r$ RF chains, respectively. The training input signal $\mathbf{s}[k] \in \mathbb{C}^{L_t \times 1}$ on the $k^{\text{th}}$ subcarrier is OFDM modulated, up-converted to RF, and analog precoded using $\mathbf{F}_{\text{tr}} \in \mathbb{C}^{N_t \times L_t}$, and transmitted over the air to the Rx via an mmWave channel denoted by $\mathbf{H}[k]$ on the $k^{\text{th}}$ subcarrier. The received signal is filtered using an RF combining matrix $\mathbf{W}_{\text{tr}} \in \mathbb{C}^{N_r \times L_r}$, down-converted to baseband, OFDM demodulated

to obtain the $k^{\text{th}}$ subcarrier's complex baseband signal $\mathbf{y}[k] \in \mathbb{C}^{L_r \times 1}$. We denote the total number of subcarriers by $K$.

In the initial access phase, the system has no prior knowledge of the channel, and therefore the precoder and combiner matrices cannot be designed to optimize any chosen performance metric. Hence, we choose random analog precoding and combining matrices (with unit modulus entries). In our system model, we adopt a fully connected phase shifter network for analog precoding/combining. The analog precoders and combiners are frequency-flat, and thus are the same for each subcarrier $k = 1, \ldots, K$. The system operates with Uniform Linear Arrays (ULAs) at both the Tx and Rx with half wavelength spacing between consecutive antennas. The total number of training frames is denoted by $M$.

After RF combining, down-conversion, zero/cyclic prefix removal and DFT, the complex baseband signal received during the $m^{\text{th}}$ training frame for the $k^{\text{th}}$ subcarrier, denoted by $\mathbf{y}^{(m)}[k] \in \mathbb{C}^{L_r \times 1}$ is given by

$$\mathbf{y}^{(m)}[k] = \mathbf{W}_{\text{tr}}^{(m)^*}\big(\mathbf{H}[k]\mathbf{F}_{\text{tr}}^{(m)}\mathbf{q}^{(m)}t^{(m)}[k] + \mathbf{n}^{(m)}[k]\big), \tag{4.1}$$

for $m = 1, \ldots, M$ where $\mathbf{H}[k] \in \mathbb{C}^{N_r \times N_t}$ represents the frequency domain MIMO channel matrix for the $k^{\text{th}}$ subcarrier. We choose the $m^{\text{th}}$ training signal as $\mathbf{s}^{(m)}[k] = \mathbf{q}^{(m)}t^{(m)}[k]$, where $t^{(m)}[k] \in \mathbb{C}$ is a subcarrier-dependent pilot symbol, and $\mathbf{q}^{(m)} \in \mathbb{C}^{L_t \times 1}$ is a frequency-flat vector whose entries are chosen as $\frac{1}{\sqrt{2L_t}}(a+jb)$, where $a, b \in \{-1, 1\}$ and are uniformly distributed. The noise vector $\mathbf{n}^{(m)}[k]$ is independently and identically distributed across $K$ subcarriers as $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_r})$. We define the transmit Signal-to-Noise Ratio (SNR) as $\rho = \frac{1}{\sigma_n^2}$. After compensating for $t^{(m)}[k]$, and vectorizing (4.1), we use the result $\texttt{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})\texttt{vec}(\mathbf{X})$ to obtain

$$\texttt{vec}(\mathbf{y}^{(m)}[k]) = \underbrace{\Big(\mathbf{q}^{(m)^T}\mathbf{F}_{\text{tr}}^{(m)^T} \otimes \mathbf{W}_{\text{tr}}^{(m)^*}\Big)}_{\mathbf{\Phi}^{(m)}}\texttt{vec}(\mathbf{H}[k]) + \mathbf{W}_{\text{tr}}^{(m)^*}\mathbf{n}^{(m)}[k]. \tag{4.2}$$

Next, we describe the mmWave channel model.

### 4.2.1 Channel model

We consider a frequency-selective geometric channel model that is constant across $M$ training frames, and has $N_c$ delay taps [57,74]. The $d^{\text{th}}$ delay tap is modeled as a clustered channel with $L$ paths as

$$\mathbf{H}_d = \sqrt{\frac{N_t N_r}{L\rho_L}} \sum_{\ell=1}^{L} \alpha_\ell p(dT_s - \tau_\ell)\mathbf{a}_{\text{R}}(\phi_\ell)\mathbf{a}_{\text{T}}^*(\theta_\ell), \tag{4.3}$$

where $\rho_L$ is the path loss between Tx and Rx, $\alpha_\ell$ represents the complex path gain, $\phi_\ell$ is the AoA, $\theta_\ell$ is the AoD, $\tau_\ell$ denotes the delay of the $\ell^{\text{th}}$ path. The corresponding Rx and Tx array steering vectors are denoted by $\mathbf{a}_{\text{R}}(\phi_\ell) \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{a}_{\text{T}}(\theta_\ell) \in \mathbb{C}^{N_t \times 1}$, respectively. The pulse shaping and other low pass filtering evaluated at $\tau$ is represented by $p(\tau)$, and $T_s$ is the sampling interval. We represent the MIMO channel $\mathbf{H}_d$ in a matrix form as

$$\mathbf{H}_d = \mathbf{A}_{\text{R}}\mathbf{\Delta}_d\mathbf{A}_{\text{T}}^*, \tag{4.4}$$

where $\mathbf{A}_{\text{R}} \in \mathbb{C}^{N_r \times L}$ and $\mathbf{A}_{\text{T}} \in \mathbb{C}^{N_t \times L}$ contain the Rx and Tx array steering vectors $\mathbf{a}_{\text{R}}(\phi_\ell)$ and $\mathbf{a}_{\text{T}}(\theta_\ell)$ as their columns for $\ell = 1, \ldots, L$, respectively. $\mathbf{\Delta}_d \in \mathbb{C}^{L \times L}$ is a diagonal matrix containing the complex channel gains. We take a $K$-point DFT of the delay-domain channel to get the frequency domain representation as

$$\mathbf{H}[k] = \sum_{d=0}^{N_c-1} \mathbf{H}_d \exp\left(-\frac{j2\pi kd}{K}\right) = \mathbf{A}_{\text{R}}\mathbf{\Delta}[k]\mathbf{A}_{\text{T}}^*, \tag{4.5}$$

for $k = 0, \ldots, K-1$, and

$$\mathbf{\Delta}[k] = \sum_{d=0}^{N_c-1} \mathbf{\Delta}_d \exp\left(-\frac{j2\pi kd}{K}\right). \tag{4.6}$$

We adopt the extended virtual channel model in [74] to represent $\mathbf{H}_d$ as

$$\mathbf{H}_d \approx \tilde{\mathbf{A}}_{\text{R}}\mathbf{\Delta}_d^v\tilde{\mathbf{A}}_{\text{T}}^*, \tag{4.7}$$

where the dictionary matrices $\tilde{\mathbf{A}}_{\mathrm{R}} \in \mathbb{C}^{N_r \times G_r}$ and $\tilde{\mathbf{A}}_{\mathrm{T}} \in \mathbb{C}^{N_t \times G_t}$ contain the Rx and Tx array steering vectors evaluated on a grid of size $G_r$ for the AoA and a grid of size $G_t$ for the AoD, respectively. When $G_r$ and $G_t$ are chosen properly, i.e., much greater than $L$, $\boldsymbol{\Delta}_d^v \in \mathbb{C}^{G_r \times G_t}$ becomes a sparse matrix containing the channel path gains on the locations that match with the actual AoDs and AoAs. We represent (4.7) in the frequency domain as

$$\mathbf{H}[k] \approx \tilde{\mathbf{A}}_{\mathrm{R}} \boldsymbol{\Delta}^v[k] \tilde{\mathbf{A}}_{\mathrm{T}}^*, \tag{4.8}$$

for $k = 0, \ldots, K-1$, and

$$\boldsymbol{\Delta}^v[k] = \sum_{d=0}^{N_c-1} \boldsymbol{\Delta}_d^v \exp\left(-\frac{j2\pi kd}{K}\right). \tag{4.9}$$

Note that the dictionary matrices $\tilde{\mathtt{vec}}(A)_{\mathrm{R}}$ and $\tilde{\mathtt{vec}}(A)_{\mathrm{T}}$ are common to all the subcarriers due to the frequency-flat array response vectors. Hence, the sparse matrices $\boldsymbol{\Delta}^v[k]$ for $k = 1, \ldots, K$ have the non-zero elements at the same indices. This means that they share a common sparsity pattern [57].

Now, we vectorize (4.8) to get

$$\mathtt{vec}(\mathbf{H}[k]) = \left(\bar{\tilde{\mathbf{A}}}_{\mathrm{T}} \otimes \tilde{\mathbf{A}}_{\mathrm{R}}\right) \mathtt{vec}(\boldsymbol{\Delta}^{\mathrm{v}}[k]). \tag{4.10}$$

We define $\boldsymbol{\Psi} = \bar{\tilde{\mathbf{A}}}_{\mathrm{T}} \otimes \tilde{\mathbf{A}}_{\mathrm{R}} \in \mathbb{C}^{N_t N_r \times G_t G_r}$ and $\mathbf{H}^{\mathrm{v}}[k] = \mathtt{vec}(\boldsymbol{\Delta}^{\mathrm{v}}[k]) \in \mathbb{C}^{G_t G_r}$, and substitute $\mathtt{vec}(\mathbf{H}[k])$ in (4.2) to get

$$\mathtt{vec}(\mathbf{y}^{(m)}[k]) = \boldsymbol{\Phi}^{(m)} \boldsymbol{\Psi} \mathbf{H}[k] + \mathbf{n}_c^{(m)}[k], \tag{4.11}$$

where $\mathbf{n}_c^{(m)}[k] = \mathbf{W}_{\mathrm{tr}}^{(m)*} \mathbf{n}^{(m)}[k]$. By concatenating the RF combined signals of $M$ training

frames, we get

$$\underbrace{\begin{bmatrix} \mathbf{y}^{(1)}[k] \\ \vdots \\ \mathbf{y}^{(M)}[k] \end{bmatrix}}_{\mathbf{y}[k]} = \underbrace{\begin{bmatrix} \mathbf{\Phi}^{(1)} \\ \vdots \\ \mathbf{\Phi}^{(M)} \end{bmatrix}}_{\mathbf{\Phi}} \mathbf{\Psi}\mathbf{h}^{\mathrm{v}}[k] + \underbrace{\begin{bmatrix} \mathbf{n}_c^{(1)}[k] \\ \vdots \\ \mathbf{n}_c^{(M)}[k] \end{bmatrix}}_{\mathbf{n}_c[k]}. \tag{4.12}$$

Now, by stacking the received signals of $K$ subcarriers, we get the final system equation

$$\begin{aligned} \mathbf{Y} &= \begin{bmatrix} \mathbf{y}[1] & \dots & \mathbf{y}[K] \end{bmatrix} \\ &= \mathbf{\Phi}\mathbf{\Psi} \begin{bmatrix} \mathbf{h}^{\mathrm{v}}[1] & \dots & \mathbf{h}^{\mathrm{v}}[K] \end{bmatrix} + \begin{bmatrix} \mathbf{n}_c[1] & \dots & \mathbf{n}_c[K] \end{bmatrix} \\ &= \mathbf{\Phi}\mathbf{\Psi}\mathbf{H}^{\mathrm{v}} + \mathbf{N}_c. \end{aligned} \tag{4.13}$$

Our goal is to estimate $\mathbf{H}[k]$, for $k = 0, \dots, K-1$, given $\mathbf{Y}$ and $\mathbf{\Phi}$. As the AoDs and AoAs are the same for all the subcarriers, $\mathbf{H}^{\mathrm{v}} \in \mathbb{C}^{G_t G_r \times K}$ has a joint row sparse structure, i.e., the support set of each column of $\mathbf{H}^{v}$ are the same. Also, we do not have the knowledge of the sparsifying dictionary $\mathbf{\Psi}$ and the noise variance, which makes the channel estimation problem more challenging. In the following section, we present our solution to this channel estimation problem.

## 4.3 MLGS-SBL

In this section, we present a model-based approach using the framework of Compressed Sensing (CS), to estimate the mmWave channel given the received pilot measurements and the frequency-flat transmit vector. We integrate a fast greedy search procedure and a high performing statistical inference method to estimate the channel. The algorithm consists of the following steps:

1. Preconditioning
2. Multi-level greedy search for dictionary learning
3. Noise variance estimation

4. Sparse Bayesian learning for channel estimation

5. Channel de-noising

We provide a detailed description of each step below.

### 4.3.1 Preconditioning

Sparse signal recovery using greedy algorithms, such as OMP, are likely to choose the correct support set when the noise covariance matrix is diagonal. In our mmWave channel estimation problem, RF combining by $\mathbf{W}_{\text{tr}}$ at the front end of the receiver results in correlated noise, which needs to be whitened using a preconditioning filter [57].

The scaled noise covariance matrix before whitening is

$$\mathbf{C}_{\text{w}} = \frac{\mathbb{E}\left[\mathbf{n}_c[k]\mathbf{n}_c^*[k]\right]}{\sigma^2} = \texttt{blkdiag}\{\mathbf{W}_{\text{tr}}^{(1)^*}\mathbf{W}_{\text{tr}}^{(1)}, \ldots, \mathbf{W}_{\text{tr}}^{(M)^*}\mathbf{W}_{\text{tr}}^{(M)}\}. \tag{4.14}$$

We get the above by noting that

$$\mathbb{E}\left[\mathbf{n}_c^{(i)}[k]\mathbf{n}_c^{(j)^*}[k]\right] = \sigma^2\mathbf{W}_{\text{tr}}^{(i)^*}\mathbf{W}_{\text{tr}}^{(j)}\delta[i-j]. \tag{4.15}$$

We perform a Cholesky decomposition of $\mathbf{C}_{\text{w}}$ to obtain $\mathbf{C}_{\text{w}} = \mathbf{D}_{\text{w}}^*\mathbf{D}_{\text{w}}$, where $\mathbf{D}_{\text{w}} \in \mathbb{C}^{ML_r \times ML_r}$ is upper triangular. Let us define $\mathbf{D}_{\text{w}}^{-*}$ to denote the inverse of $\mathbf{D}_{\text{w}}^*$. Now, we multiply the RF combined received signal (4.12) by $\mathbf{D}_{\text{w}}^{-*}$ to obtain the noise-whitened received signal:

$$\begin{aligned}\mathbf{y}_{\text{w}}[k] = \mathbf{D}_{\text{w}}^{-*}\mathbf{y}[k] &= \mathbf{D}_{\text{w}}^{-*}\mathbf{\Phi}\mathbf{\Psi}\mathbf{h}^{\text{v}}[k] + \mathbf{D}_{\text{w}}^{-*}\mathbf{n}_c[k] \\ &= \mathbf{\Upsilon}_{\text{w}}\mathbf{h}^{\text{v}}[k] + \mathbf{D}_{\text{w}}^{-*}\mathbf{n}_c[k],\end{aligned} \tag{4.16}$$

where $\mathbf{\Upsilon}_{\text{w}} = \mathbf{D}_{\text{w}}^{-*}\mathbf{\Phi}\mathbf{\Psi} \in \mathbb{C}^{ML_r \times G_tG_r}$. Concatenating the noise-whitened received signals of all the $K$ subcarriers, we get

$$\mathbf{Y}_{\text{w}} = \begin{bmatrix}\mathbf{y}_{\text{w}}[1] & \ldots & \mathbf{y}_{\text{w}}[K]\end{bmatrix} = \mathbf{\Phi}_{\text{w}}\mathbf{\Psi}\mathbf{H}^{\text{v}} + \mathbf{N}_{\text{w}}\,, \tag{4.17}$$

where $\mathbf{Y}_{\mathrm{w}} \in \mathbb{C}^{ML_r \times K}$, $\mathbf{\Phi}_{\mathrm{w}} = \mathbf{D}_{\mathrm{w}}^{-*}\mathbf{\Phi} \in \mathbb{C}^{ML_r \times N_t N_r}$, and $\mathbf{N}_{\mathrm{w}} = \mathbf{D}_{\mathrm{w}}^{-*}\begin{bmatrix}\mathbf{n}[1] & \dots & \mathbf{n}[K]\end{bmatrix} \in \mathbb{C}^{ML_r \times K}$. Thus, we need to estimate the row sparse matrix $\mathbf{H}^{\mathrm{v}}$ given $\mathbf{Y}_{\mathrm{w}}$ and $\mathbf{\Phi}_{\mathrm{w}}$.

### 4.3.2 Multi-level Greedy Search

We obtain an initial channel estimate using the MLGS procedure with a coarsely quantized beamspace dictionary. We adopt the Simultaneously Weighted Orthogonal Matching Pursuit (SW-OMP) algorithm as our base algorithm to form an initial estimate of the channel [57]. As the sparsifying dictionary $\mathbf{\Psi}$ is unknown a priori, we use row-truncated DFT matrices of size $N_t \times G_t$ and $N_r \times G_r$ as the Tx and Rx array steering matrices, respectively. Let $\widetilde{\mathbf{\Psi}}$ be the initial sparsifying dictionary.

In the first step of MLGS, we select a column from $\widetilde{\mathbf{\Psi}}$ that is maximally correlated with the received signal. Mathematically,

$$\widehat{i} = \arg\max_{i} \sum_{k=1}^{K} \left| \left( \mathbf{\Phi}_{\mathrm{w}} \widetilde{\mathbf{\Psi}}[:, i] \right)^{*} \mathbf{y}_{\mathrm{w}}[k] \right|^{2}, \tag{4.18}$$

where $|\cdot|$ denotes an element-wise modulus operation, and $\widetilde{\mathbf{\Psi}}[:, i]$ is the $i^{\text{th}}$ column of $\widetilde{\mathbf{\Psi}}$. Once we select $\widehat{i}$, we extract AoD $\theta_{\widehat{i}}$ and AoA $\phi_{\widehat{i}}$ using the structure of $\widetilde{\mathbf{\Psi}}$, and form a finely spaced dictionary of range $(\theta_{\widehat{i}} - \Delta\theta, \theta_{\widehat{i}} + \Delta\theta)$ and $(\phi_{\widehat{i}} - \Delta\phi, \phi_{\widehat{i}} + \Delta\phi)$, where $\Delta\theta$ and $\Delta\phi$ are appropriately chosen based on the spatial quantization of the previously chosen dictionary. We repeat (4.18) with $\widetilde{\mathbf{\Psi}}$ replaced by the newly formed dictionary, and choose a new {AoD, AoA} pair. We repeat this process $N$ times and select one set of AoD and AoA. Then, we compute

$$\widehat{\mathbf{H}}^{\mathrm{v}} = \left( \mathbf{\Phi}_{\mathrm{w}} \widehat{\mathbf{\Psi}} \right)^{\dagger} \mathbf{Y}_{\mathrm{w}}, \tag{4.19}$$

where $\widehat{\mathbf{\Psi}}$ is formed using the currently chosen AoD and AoA. This whole procedure constitutes the first out of $S$ iterations of the MLGS algorithm in which we recover a single tap.

In the $s^{\text{th}}$ iteration of MLGS, we recover $s$ channel taps by following the same steps as

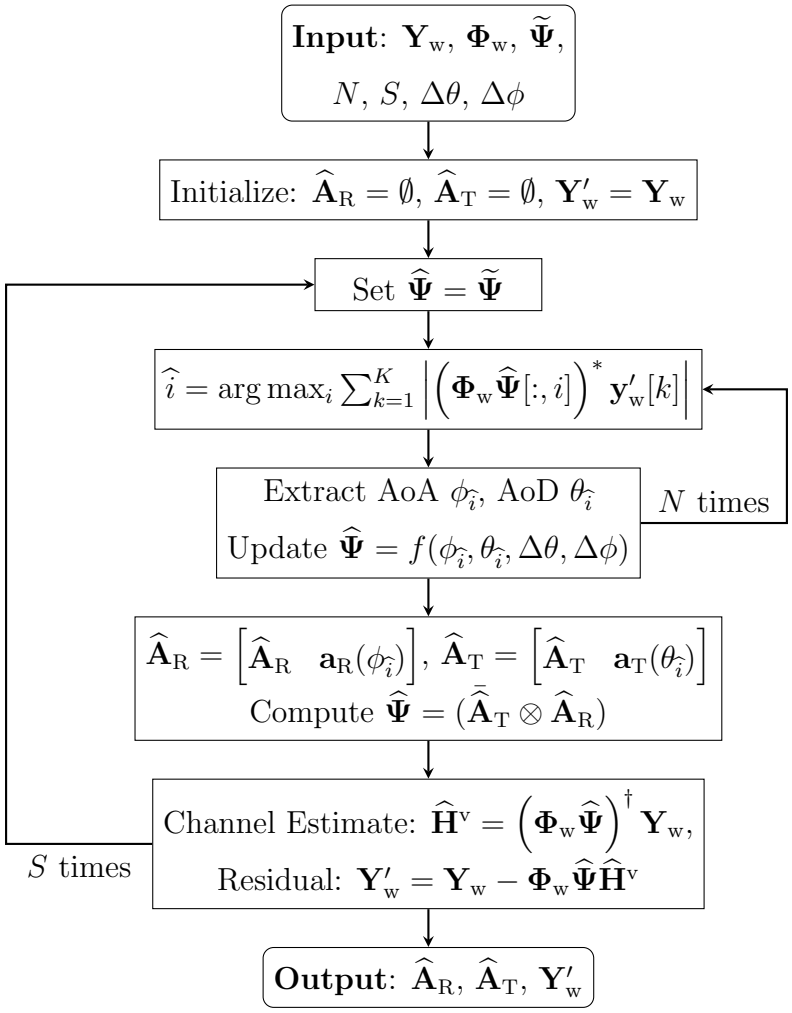


Figure 4.2: Flow diagram of MLGS.

above, but with the residual $\mathbf{Y}'_{\mathrm{w}} = \mathbf{Y}_{\mathrm{w}} - \mathbf{\Phi}_{\mathrm{w}} \widehat{\mathbf{\Psi}} \widehat{\mathbf{H}}^{\mathrm{v}}$ as observations, where $\widehat{\mathbf{\Psi}}$ comprises the set of {AoD, AoA} pairs chosen in the first $s-1$ iterations. Therefore, after $S$ iterations, we recover $S$ virtual beamspace channel taps. We summarize MLGS as a flow diagram in Fig. 4.2.

### 4.3.3 Noise Variance Estimation

We estimate the noise variance $\widehat{\sigma}_n^2$ using the residual output from MLGS. The noise variance is computed as

$$\widehat{\sigma}_n^2 = \frac{1}{MKL_r}||\mathbf{Y}'_{\mathrm{w}}||_F^2. \tag{4.20}$$

### 4.3.4 Sparse Bayesian Learning

In this step, our goal is to refine the channel estimates output by the MLGS procedure. For convenience, we recall the measurement equation:

$$\mathbf{Y}_{\mathrm{w}} = \mathbf{\Phi}_{\mathrm{w}}\widehat{\mathbf{\Psi}}\mathbf{H}^{\mathrm{v}} + \mathbf{N}_{\mathrm{w}}\,, \tag{4.21}$$

where $\widehat{\mathbf{\Psi}} = (\widehat{\bar{\mathbf{A}}}_{\mathrm{T}} \otimes \widehat{\mathbf{A}}_{\mathrm{R}})$ is the dictionary output by MLGS. We adopt a statistical inference approach to infer the posterior distribution of $\mathbf{H}^{\mathrm{v}}$ given the measurements $\mathbf{Y}_{\mathrm{w}}$, measurement matrix $\mathbf{\Phi}_{\mathrm{w}}\widehat{\mathbf{\Psi}}$, and noise variance $\widehat{\sigma}_n^2$. Let us denote $\mathbf{\Phi}_{\mathrm{w}}\widehat{\mathbf{\Psi}}$ by $\widehat{\mathbf{\Upsilon}}_{\mathrm{w}}$.

We use sparse Bayesian learning, a type-II maximum likelihood estimation procedure to obtain the channel estimate [32, 147]. In this method, we consider $\mathbf{H}^{\mathrm{v}}$ as a hidden variable, and obtain its posterior statistics given the observations. We impose a parameterized complex Gaussian prior on each column of the channel as $\mathcal{CN}(\mathbf{0}, \mathbf{\Gamma})$, where $\mathbf{\Gamma} = \mathtt{diag}(\boldsymbol{\gamma})$. Using a common hyperparameter $\boldsymbol{\gamma}$ across all the columns of $\mathbf{H}^{\mathrm{v}}$ aids in promoting common row sparsity in the solution. Now, we need to obtain the posterior distribution of $\mathbf{H}^{\mathrm{v}}$, and the hyperparameter $\boldsymbol{\gamma}$. Since the prior and the noise are both Gaussian, obtaining the posterior statistics of $\mathbf{H}^{\mathrm{v}}$ is straightforward. But, computing $\boldsymbol{\gamma}$ requires computing the marginal probability distribution $p(\mathbf{Y}_{\mathrm{w}}; \boldsymbol{\gamma})$ and maximizing it w.r.t. $\boldsymbol{\gamma}$, which is called evidence maximization or type-II maximum likelihood estimation.

We adopt the Expectation Maximization (EM) algorithm to learn the beamspace sparse channel $\mathbf{H}^{\mathbf{v}}$ and the hyperparameters $\boldsymbol{\gamma}$. In the EM formulation, we treat $\mathbf{H}^{\mathbf{v}}$ as a hidden variable, and iteratively maximize a lower bound on the posterior distribution $p(\boldsymbol{\gamma}|\mathbf{Y}_{\mathbf{w}})$. We alternate between E and M steps till a suitable convergence criterion is

satisfied. We explain the steps below:

#### 4.3.4.1 Expectation Step

In the E step, we compute the expectation of the log likelihood $p(\mathbf{Y}_{\mathrm{w}}, \mathbf{H}^{\mathrm{v}}|\boldsymbol{\gamma})$ w.r.t. the posterior probability distribution $p(\mathbf{H}^{\mathrm{v}}|\mathbf{Y}_{\mathrm{w}}, \boldsymbol{\gamma}^{(i)})$, where $i$ denotes the iteration index. Given the hyperparameters $\boldsymbol{\gamma}^{(i)}$, $p(\mathbf{H}^{\mathrm{v}}|\mathbf{Y}_{\mathrm{w}}, \boldsymbol{\gamma}^{(i)})$ is complex normal distributed with covariance and mean given by

$$\boldsymbol{\Sigma}_{\mathbf{H}} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\widehat{\boldsymbol{\Upsilon}}_{\mathrm{w}}^{*}{\boldsymbol{\Sigma}_{\mathbf{Y}}}^{-1}\widehat{\boldsymbol{\Upsilon}}_{\mathrm{w}}\boldsymbol{\Gamma}, \tag{4.22}$$

$$\widehat{\mathbf{H}}^{\mathrm{v}} = \frac{1}{\widehat{\sigma}_n^2}\boldsymbol{\Sigma}_{\mathbf{H}}\widehat{\boldsymbol{\Upsilon}}_{\mathrm{w}}^{*}\mathbf{Y}_{\mathrm{w}}, \tag{4.23}$$

where $\boldsymbol{\Gamma} = \mathtt{diag}(\boldsymbol{\gamma}^{(i)})$, and

$$\boldsymbol{\Sigma}_{\mathbf{Y}} = \widehat{\sigma}_n^2\mathbf{I}_{ML_r} + \widehat{\boldsymbol{\Upsilon}}_{\mathrm{w}}\boldsymbol{\Gamma}\widehat{\boldsymbol{\Upsilon}}_{\mathrm{w}}^{*} \tag{4.24}$$

is the covariance matrix of the circularly symmetric complex normal marginal probability distribution $p(\mathbf{Y}_{\mathbf{w}})$.

#### 4.3.4.2 Maximization Step

In the M step, we compute the hyperparameter $\boldsymbol{\gamma}$ by maximizing the function obtained in the E step:

$$\boldsymbol{\gamma}^{(i+1)} = \underset{\boldsymbol{\gamma}\geq\mathbf{0}}{\arg\max}\ \mathbb{E}_{p(\mathbf{H}^{\mathrm{v}}|\mathbf{Y}_{\mathrm{w}},\boldsymbol{\gamma}^{(i)})}\left[p(\mathbf{Y}_{\mathrm{w}}, \mathbf{H}^{\mathrm{v}}|\boldsymbol{\gamma})\right]. \tag{4.25}$$

Solving (4.25), we get the hyperparameter update as

$$\gamma_n^{(i+1)} = \frac{1}{K}\sum_{k=1}^{K}|\widehat{\mathbf{H}}^{\mathrm{v}}[n,k]|^2 + \boldsymbol{\Sigma}_{\mathbf{H}}[n,n], \qquad n = \{1,\ldots,S^2\}. \tag{4.26}$$

We alternate between E and M steps till a suitable convergence criterion is satisfied. We use the Frobenius norm of the difference between the channel estimates of consecutive

iterations as the convergence criterion in our simulations.

More theoretical details of SBL and type-II ML estimation can be found in [32, 123]. We provide a flow diagram of Multiple Measurement Vector SBL (MSBL) to compute the posterior mean and covariance of the channel, and the hyperparameters, in Fig. 4.3. Specifically, in Fig. 4.3, the E-step of the EM algorithm corresponds to the computation of $\mathbf{\Sigma_Y}$, $\mathbf{\Sigma_H}$ and $\widehat{\mathbf{H}}^{\text{v}}$, and the M-step corresponds to the computation of $\mathbf{\Gamma}$.

Once we obtain the frequency domain channel estimate $\widehat{\mathbf{H}}^{\text{v}}$, we estimate the support of the row sparse matrix and the channel coefficients using the hyperparameters obtained using SBL. We re-estimate the noise variance using the Frobenius norm of the residual $\widetilde{\mathbf{Y}}_{\text{w}} = \mathbf{Y}_{\text{w}} - \mathbf{\Phi}_{\text{w}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{H}}^{\text{v}}$.

### 4.3.5 Denoising

By analyzing the training dataset, we observed that the channel is sparse in both the virtual beamspace and delay domains. We exploited the beamspace sparsity and obtained the frequency domain channel estimates using MLGS and SBL. In this final step, we exploit the delay domain sparsity to denoise the channel to further reduce the MSE between the original and estimated channels.

For each subcarrier $k$, we compute $(\bar{\widehat{\mathbf{A}}}_{\text{T}} \otimes \widehat{\mathbf{A}}_{\text{R}})\mathbf{H}^{\text{v}}[:, k]$, and reshape it to form $k^{\text{th}}$ subcarrier's channel matrix of size $N_r \times N_t$. Then, for each transmit and receive antenna pair, we compute a $K$-point inverse DFT to obtain a delay-domain channel estimate. We retain the $P$ dominant taps in the delay-domain channel estimate, and set the other $K-P$ taps to 0. We fix $P$ based on the estimated noise variance, and the number of training frames $M$. The value of $P$ is inversely proportional to $\widehat{\sigma}_n^2$, and the training dataset is used to choose an appropriate $P$. From our experiments on the training dataset, we found that this denoising step leads to an approximately 2 dB reduction in NMSE. This concludes the description of the MLGS-SBL approach.

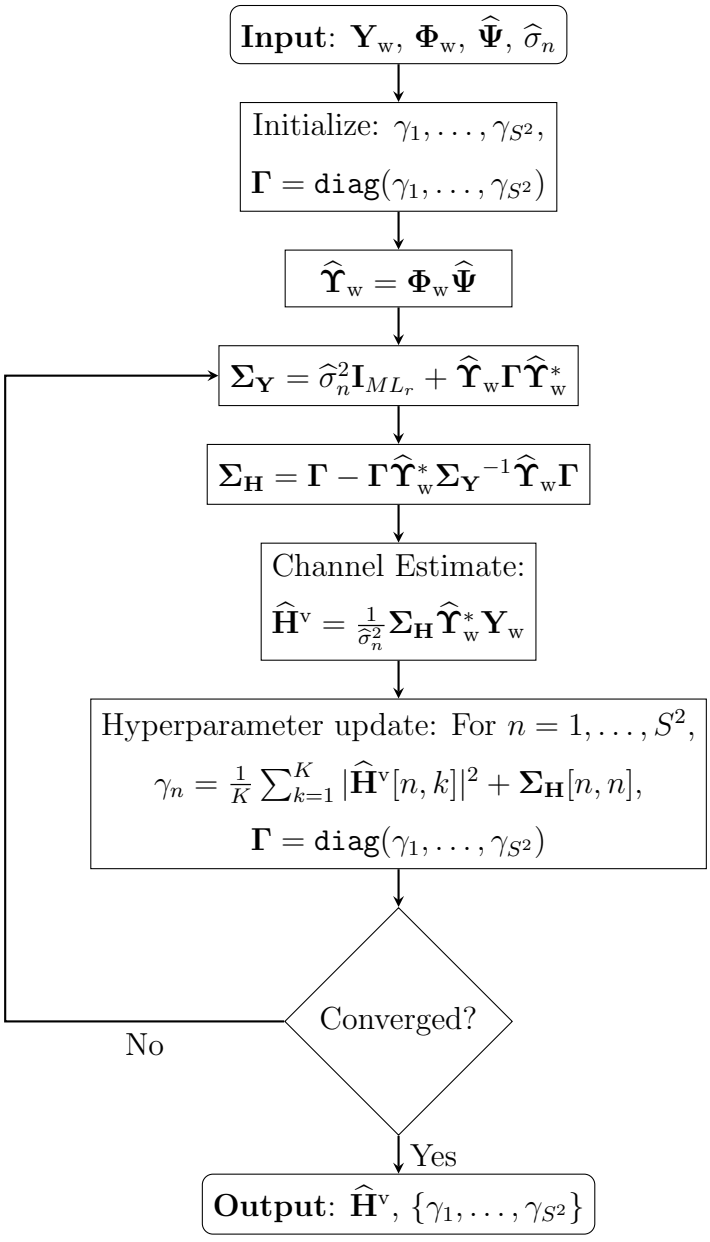


Figure 4.3: Flow diagram of MSBL.

Table 4.1: NMSE Table for Training Data

| **SNR (dB)** | Algorithm | **−15** | **−10** | **−5** |
|---|---|---|---|---|
| **Pilot Frames:** 20 | SW-OMP | −1.45 dB | −5.70 dB | −9.68 dB |
| | MLGS-SBL | −4.29 dB | −9.13 dB | −12.34 dB |
| | PCSBL-DDT | −8.16 dB | −10.62 dB | −11.07 dB |
| | PC-OMP | −8.34 dB | −12.36 dB | −16.15 dB |
| **Pilot Frames:** 40 | SW-OMP | −3.95 dB | −7.95 dB | −11.87 dB |
| | MLGS-SBL | −7.55 dB | −11.19 dB | −14.15 dB |
| | PCSBL-DDT | −10.56 dB | −12.14 dB | −12.62 dB |
| | PC-OMP | −12.66 dB | −16.33 dB | −19.78 dB |
| **Pilot Frames:** 80 | SW-OMP | −7.33 dB | −11.60 dB | −15.63 dB |
| | MLGS-SBL | −13.02 dB | −16.37 dB | −18.94 dB |
| | PCSBL-DDT | −11.90 dB | −13.10 dB | −13.63 dB |
| | PC-OMP | −18.70 dB | −21.49 dB | −24.48 dB |

Table 4.2: NMSE Table for Test Data

| **SNR (dB)** | Algorithm | [**−20**, **−11**) | [**−11**, **−6**) | [**−6**, **0**] |
|---|---|---|---|---|
| **Pilot Frames:** 20 | MLGS-SBL | −7.66 dB | −10.97 dB | −12.34 dB |
| | PCSBL-DDT | −8.94 dB | −9.99 dB | −10.31 dB |
| | PC-OMP | −9.09 dB | −12.45 dB | −14.22 dB |
| **Pilot Frames:** 40 | MLGS-SBL | −11.87 dB | −12.79 dB | −14.20 dB |
| | PCSBL-DDT | −10.82 dB | −11.33 dB | −11.89 dB |
| | PC-OMP | −13.79 dB | −15.24 dB | −16.79 dB |
| **Pilot Frames:** 80 | MLGS-SBL | −13.62 dB | −16.23 dB | −20.08 dB |
| | PCSBL-DDT | −11.74 dB | −12.47 dB | −12.98 dB |
| | PC-OMP | −16.32 dB | −19.07 dB | −23.91 dB |

## 4.4 Simulation Results

In this section, we discuss the NMSE performance of our algorithm with the training and testing data generated using *Raymobtime,* a ray tracing based mmWave channel generation tool. We train the mmWave channel estimation algorithm using $10,000$ independent channel realizations, each consisting of 100 paths between the Tx and Rx. More details about the channel generation methodology can be found in [146]. We used 20, 40, and 80 pilot frames during both the training and testing phases of the developed algorithm. For the training phase, we used SNR values of $\{-15, -10, -5\}$ dB. We benchmark the NMSE performance of our MLGS-SBL algorithm with that of PCSBL-DDT, PC-OMP and a reference state-of-the-art model-based greedy search algorithm called SW-OMP [57].

We note that while MLGS-SBL, PCSBL-DDT, and PC-OMP have been fine-tuned based on the training dataset, the baseline algorithm, SW-OMP, has been implemented as-is from the literature. On the other hand, in our implementation of SW-OMP, we consider the case where the true AoDs and AoAs are contained in the sparsifying dictionary. While the derived algorithm does suffer from the off-grid effects, the SW-OMP algorithm is insulated from the performance degradation caused by them.

Upon analyzing the training channels, we set the maximum number of paths obtained from MLGS to $S = 10$, and the number of levels in MLGS to $N = 5$. We set the number of columns in the initial AoD/AoA steering matrices to 256. We use the estimated noise variance after SBL to threshold the number of dominant delay taps of the channel denoiser. As this approach is primarily a model-based method, and uses few statistics from the training data, it is suitable for general mmWave channel estimation problems also. Further, the thresholds are set keeping in mind the computational complexity of the MLGS-SBL algorithm. By increasing the number of paths output by MLGS, we can potentially improve the performance of the algorithm, but at the cost of higher computational complexity. We include the NMSE values obtained for the training and testing datasets in Table 4.1 and Table 4.2, respectively. The final performance score achieved, which is a weighted combination of the NMSE performance in Table 4.2 when the number of pilot frames is 20, using our algorithm in the mmWave channel estimation challenge is

$-9.16$ dB.

The final performance score on the test dataset in the channel estimation challenge of the PC-OMP algorithm is $-10.64$ dB, outperforming the MLGS-SBL and the PCSBL-DDT methods by 1.48 dB and 1.15 dB, respectively. Also, from Table 4.1, we can see that the PCSBL-DDT and the PC-OMP algorithms are tuned better than the MLGS-SBL method for the training dataset that result in their better NMSE performances at SNR $-15$ dB and pilot frames $\{20, 40\}$. But the performance gap between MLGS-SBL and PCSBL-DDT reduces for the testing data for SNR $[-20, -11)$ dB and 20 pilot frames. Moreover, MLGS-SBL performs better than PCSBL-DDT at SNR $[-20, -11)$ dB and 40 pilot frames. This can be attributed to the fact that extracting more features from a training dataset may result in an excellent performance during training but slightly inferior performance while testing. This shows that a good model-based signal processing solution has to be combined with appropriate training, while taking into account the training and testing performance trade-off.

## 4.5 Conclusions

In this chapter, we have presented a novel signal processing approach to estimate an mmWave channel in a hybrid analog-digital MIMO setup. We have adapted model-driven procedures to utilize the AoD, AoA, and channel gain information from a training dataset, and fine-tuned the algorithms to reduce the NMSE in the testing dataset. We empirically showed that our algorithm unanimously performed better than a purely model-based approach by a large margin on a given training data set. Hence, machine learning approaches can be potentially used in conjunction with model-driven based approaches to fine-tune them and thereby obtain better performance in physical layer wireless communication problems in realistic channel environments.

# Part 3

# Chapter 5

# Codebook Based Precoding and Power Allocation for Sum Rate Maximization in Downlink MU-MIMO Systems

## 5.1 Introduction

The design of downlink precoding and beamforming schemes for multiuser muliple input multiple output (MU-MIMO) systems with a large number of antennas at the base station (BS) or access point (AP) has attracted significant research interest in recent years [148–154]. In typical frequency division duplex systems, the channel state information (CSI) is first obtained at the UEs using downlink (DL) training, i.e., from pilot symbols transmitted by the AP. Then, the UEs send their channel estimates back to the AP over an uplink (UL) control channel. The AP, upon receiving the channel estimates, computes a precoding matrix for data transmission to each of the users. In this chapter, we consider an approach where the AP selects the columns of the precoding matrix from a predetermined codebook of beamforming vectors. This allows the AP the flexibility of either conveying the selected codebook indices to the UEs over a DL control channel, or

using dedicated pilots to enable the UEs to estimate their respective effective channels. Codebook based precoding is also relevant because it is employed (for example) in the IEEE 802.15.3 and IEEE 802.15ac standards [76, 155]. Our goal in this context is to determine the optimal selection of beamforming vectors and power allocation across users, with possibly *multi-stream* data transmission to each of the UEs. This is a non-convex and combinatorial problem, and therefore hard to solve. We present two novel algorithms based on the minorization-maximization (MM) framework for maximizing the sum rate under the codebook constraint. In the process, we also develop new matrix inequalities that facilitate the use of the MM approach for optimization. These latter results could be of independent interest in many other non-convex optimization problems.

### 5.1.1 Prior Art

Most of the existing studies on sum rate maximization in MU-MIMO systems do not consider the problem when the transmitter is constrained to select its precoding vectors from a codebook of candidate vectors [156–163]. In codebook based transmission, the columns of the precoding matrices need to be selected from the codebook. This makes our problem fundamentally different from, and intrinsically harder, than the (unconstrained) design of precoding matrices, as the underlying problem becomes one of allocating beamforming vectors to users, i.e., an integer optimization problem.

A codebook based approach for beamforming and power allocation in multiuser multiple input *single* output (MU-MISO) systems is considered in [164], where the authors transform the underlying mixed integer optimization problem into a structured mixed integer second-order cone program. They also customize a convex continuous relaxation based branch-and-cut algorithm to compute an optimal solution to the beamforming problem. Considering the MU-MIMO system (where multiple data streams are transmitted to each user) significantly changes the problem, because interference between streams assigned to the same UE can be handled via joint processing of the signals received at the UE antennas. This is unlike the single antenna UE case, where all inter-stream interference negatively impacts the data rate.

In [165], the authors consider the beamforming assignment and power allocation (BAPA) problem for MU-MISO systems. They introduce a virtual uplink (VUL) to decouple the power allocations across different UEs, thereby admitting an iterative solution. In the case of imperfect CSI at the transmitter, rate-splitting is shown to be sum-rate optimal [166–169], but it requires successive interference cancellation to remove the interference caused by the common messages and decode the private messages of all the users. In contrast, we consider linear receivers and linear precoding of data using a beamforming vector codebook, which leads to a solution that is easy to implement in practical systems.

In this work, we approach the problem of codebook based precoding for sum rate maximization using the iterative technique of minorization-maximization (MM). We bound the original objective function in multiple stages, which simplifies the optimization problem and helps in finding a closed-form analytical solution. Note that, when using an MM approach for optimizing a non-convex objective function, the key novelty is to bound the cost function by a surrogate function that is tight at the current iterate and is easy to optimize. Different bounds can lead to different convergence and complexity tradeoffs. We present and compare two alternatives for bounding the cost function. The resulting algorithms are computationally simple (e.g., they do not involve any matrix inversion operations), making them attractive for implementation. Further, as they are based on the MM principle, they are guaranteed to converge to a local optimum from any initialization.

### 5.1.2 Main Contributions

Our main contributions in this chapter are as follows:

1. We present a closed-form expression for the achievable sum rate of a codebook based precoding MU-MIMO broadcast system with minimum mean squared error (MMSE) channel estimation at the receiver and the feedback of imperfect CSI to the transmitter via an error-free control channel. The achievable sum rate expression provides us with the objective function for beamforming vector selection and power allocation.

2. We develop two algorithms for solving the sum rate maximization problem in MU-MIMO systems. The two algorithms differ in the way they bound the non-convex sum rate cost function to arrive at the surrogate cost that needs to be optimized.

   (a) Square root MM (SMM) algorithm: Here, we consider the square-root of the power allocation as the optimization variable, and apply minorization three times to lower-bound the objective function with a quadratic-form cost function. The surrogate cost function so obtained admits a closed-form optimal solution.

   (b) Inverse MM (IMM) algorithm: We use a matrix inequality to deal with the matrix inverse term in the objective function. After two rounds of minorization, this again leads to a surrogate quadratic lower bound, and admits a closed-form optimal solution.

   We analytically show that the closed-form solutions of the SMM and IMM algorithms are optimal with respect to their corresponding surrogate optimization problems.

3. We empirically study the performance of the SMM and IMM algorithms with respect to the number of users, codebook size, data SNR, pilot SNR etc. Further, we illustrate the performance advantage offered by the SMM and IMM algorithms compared to the WMMSE [157] and WSRMax [161] approaches as well as a single-user-optimal codebook based precoding approach, where the IMM algorithm is used to select the beamforming vector and power allocation on a per-user basis. The results demonstrate that jointly choosing beamforming vectors is necessary to realize the full potential of MU-MIMO transmission. We compare the sum rate performance and run times of the MM algorithms with that of CVX [170, 171], a convex optimization package available online. The IMM algorithm has significantly lower run time compared to CVX in the interference-limited regime, which is the primary domain of interest of our work.

We note that, even in the single-user context, our solution to the problem of selecting

*multiple* beamforming vectors in a codebook along with their corresponding power allocation is novel, and a similar solution does not exist in the literature, to the best of our knowledge. Moreover, our approach can easily accommodate additional constraints such as a minimum rate per user for a selected subset of users, etc. Also, codebook based precoding will necessarily play a role in the next generation wireless systems like mmWave massive MIMO, where it is customary to adopt a hybrid precoding architecture [45] due to the high cost and power consumption in the power amplifiers and ADCs. For instance, in the analog precoding stage, due to the finite angular resolution of the analog phase shifters, codebook based precoding arises naturally. The solution presented in this work can be easily adapted to mmWave hybrid analog-digital beamforming based systems, by imposing a constraint on the total number of spatial streams to which nonzero power is allocated. Finally, the guaranteed convergence and simple implementation makes the novel bounding technique developed in this work a potentially attractive approach for a variety of optimization problems which arise in MU-MIMO systems.

## 5.2 System Model & Problem Statement

We consider a MU-MIMO system comprised of a AP equipped with $N_t$ antennas and $K$ users each equipped with $N_r$ antennas. The UEs and the AP share a codebook $\mathbf{C} \in \mathbb{C}^{N_t \times N}$, whose columns consist of $N$ unit-norm beamforming vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N$, with $\mathbf{c}_j \in \mathbb{C}^{N_t}$. The complex baseband channel between the AP and the $k^{\text{th}}$ UE is denoted by $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$. The AP sends the data symbol $s_k(l)$ to the $k^{\text{th}}$ UE by precoding it using the $l^{\text{th}}$ beamforming vector $\mathbf{c}_l$, and the composite signal $\mathbf{x} \in \mathbb{C}^{N_t}$ transmitted by the AP is

$$\mathbf{x} = \sqrt{\rho_{dl}} \sum_{k=1}^{K} \sum_{l=1}^{N} \sqrt{P_k(l)} \mathbf{c}_l s_k(l), \tag{5.1}$$

where $\rho_{dl}$ is the data signal to noise ratio (SNR). In the sequel, all powers are normalized with respect to the noise variance, and we use the SNR and transmit power interchangeably.

Note that this general model allows multiple users to receive data on the same beamforming vector or multiple beamforming vectors to be assigned to a given user. Ultimately, the beamforming vector selection and power allocation solution will ensure that the objective function, namely, the sum rate, is maximized. Therefore, there is no need to explicitly impose constraints such as each beamforming vector should be allotted to at most one user, or that a user should not be allocated more than a given number of beamforming vectors. In (5.1), the data symbols $\{s_k(l)\}$ for $k = 1, \ldots, K$, $l = 1, \ldots, L$ are assumed to be independent and identically distributed (i.i.d.) Gaussian, with zero mean and unit variance. Let $\mathbf{\Phi}_k \triangleq \texttt{diag}\,(()\, P_k(1), P_k(2), \ldots, P_k(N))$ denote a diagonal matrix whose entries contain the fraction of the available power at the AP that is allocated to $k^{\text{th}}$ user on the $N$ beamforming vectors (hence, $\text{tr}\left(\sum_{k=1}^{K} \mathbf{\Phi}_k\right) = 1$). Then, the goal at the AP is to determine $\mathbf{\Phi}_k$, based on $\mathbf{H}_k$, $k = 1, 2, \ldots, K$, to maximize the achievable sum rate in the system. Note that, $P_k(l) = 0$ is equivalent to *not allotting* the $l^{\text{th}}$ beamforming vector in the codebook to the $k^{\text{th}}$ user.

Past works in the area, e.g., [164, 165], assume that $\mathbf{H}_k$ is perfectly known at the AP. However, in practice, channel is estimated using training symbols, which results in imperfect CSI. Therefore, we first describe the MMSE channel estimation at the UEs using common pilots transmitted by the AP.

### 5.2.1 Downlink Training and Channel Estimation

In the downlink training phase, the AP transmits $\tau_p$ orthogonal pilot symbols ($\tau_p \geq N_t$) over its $N_t$ antennas. The pilot signal $\mathbf{X}_p \in \mathbb{C}^{N_t \times \tau_p}$ satisfies $\mathbf{X}_p \mathbf{X}_p^H = \mathbf{I}_{N_t}$. The received pilot sequence at the $k^{\text{th}}$ user, $\mathbf{Y}_k^{(p)} \in \mathbb{C}^{N_r \times \tau_p}$, is given by

$$\mathbf{Y}_k^{(p)} = \sqrt{\rho_{dl}^{(p)} \tau_p} \mathbf{H}_k \mathbf{X}_p + \mathbf{W}_k, \tag{5.2}$$

where $\rho_{dl}^{(p)}$ is the pilot signal to noise ratio (SNR), and $\mathbf{W}_k \in \mathbb{C}^{N_r \times \tau_p}$ is the complex additive white Gaussian noise (AWGN) whose columns are i.i.d. with mean $\mathbf{0}$ and covariance matrix $\mathbf{I}_{N_r}$, denoted $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$. The multiplication of the transmit symbols by

$\sqrt{\tau_p}$ above is to ensure that the total energy expended over the entire pilot duration is $\tau_p$. Also, $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ denotes the channel matrix of the $k^{\text{th}}$ user, which contains i.i.d. entries drawn from $\mathcal{CN}(0, \beta_k)$, where $\beta_k$ denotes the combined effect of long term pathloss and large scale shadowing between the AP and user $k$.

By multiplying (5.2) by $\mathbf{X}_p^H$ on the right, we get

$$\mathbf{Y}_k^{(p)'} \triangleq \mathbf{Y}_k^{(p)} \mathbf{X}_p^H = \sqrt{\rho_{dl}^{(p)} \tau_p} \mathbf{H}_k + \mathbf{W}_k', \tag{5.3}$$

where $\mathbf{W}_k^{'} = \mathbf{W}_k \mathbf{X}_p^H \in \mathbb{C}^{N_r \times N_t}$ is the effective noise whose columns are also distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$. The MMSE estimate of the channel is given by [4]

$$\widehat{\mathbf{H}}_k \triangleq \sqrt{\rho_{dl}^{(p)} \tau_p} \beta_k \mathbf{Y}_k^{(p)'} / \left(1 + \rho_{dl}^{(p)} \tau_p \beta_k\right), \tag{5.4}$$

and the mean square value of each entry of $\widehat{\mathbf{H}}_k$ is given by

$$\gamma_k \triangleq \rho_{dl}^{(p)} \tau_p \beta_k^2 / \left(1 + \rho_{dl}^{(p)} \tau_p \beta_k\right). \tag{5.5}$$

Note that, $\widehat{\mathbf{H}}_k$ is Gaussian distributed and is uncorrelated with the channel estimation error $\widetilde{\mathbf{H}}_k \triangleq \mathbf{H}_k - \widehat{\mathbf{H}}_k$. This is useful in computing the noise plus interference covariance matrix, in the next subsection.

### 5.2.2 Derivation of the Achievable Rate

Consider a power allocation matrix $\mathbf{\Phi}_k$, $k = 1, 2, \ldots, K$. From (5.1), the composite signal transmitted by the AP to all the users can be written compactly as $\sqrt{\rho_{dl}} \sum_{j=1}^{K} \mathbf{C} \mathbf{\Phi}_j^{\frac{1}{2}} \mathbf{s}_j \in \mathbb{C}^{N_t}$ where $\mathbf{s}_j = [s_j(1), s_j(2), \ldots, s_j(N)]^T$ is the data transmitted data to the $j^{\text{th}}$ user using the beamforming codebook $\mathbf{C}$, and $\rho_{dl}$ is the downlink SNR. The received signal $\mathbf{y}_k \in \mathbb{C}^{N_r}$ at the $k^{\text{th}}$ user is given by

$$\mathbf{y}_k = \mathbf{H}_k \left( \sqrt{\rho_{dl}} \sum_{j=1}^{K} \mathbf{C} \mathbf{\Phi}_j^{\frac{1}{2}} \mathbf{s}_j \right) + \mathbf{w}_k,$$

where $\mathbf{w}_k \in \mathbb{C}^{N_r}$ is the complex AWGN at the $k^{\text{th}}$ user with distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$.

Given the channel estimate $\widehat{\mathbf{H}}_k$ at the receiver, the received signal can be rewritten as

$$\mathbf{y}_k = \underbrace{\sqrt{\rho_{dl}}\widehat{\mathbf{H}}_k\mathbf{C}\mathbf{\Phi}_k^{\frac{1}{2}}\mathbf{s}_k}_{\text{Desired signal}} + \sqrt{\rho_{dl}}\widehat{\mathbf{H}}_k \sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{C}\mathbf{\Phi}_j^{\frac{1}{2}}\mathbf{s}_j + \sqrt{\rho_{dl}}\widetilde{\mathbf{H}}_k \sum_{j=1}^{K} \mathbf{C}\mathbf{\Phi}_j^{\frac{1}{2}}\mathbf{s}_j + \mathbf{w}_k. \tag{5.6}$$

In order to compute the achievable rate from the above equation, we need to compute the signal and noise plus interference covariance matrices, find the signal to interference plus noise ratio (SINR) and then use the worst case noise theorem [172]. The covariance of the desired signal is $\rho_{dl}\widehat{\mathbf{H}}_k\mathbf{C}\mathbf{\Phi}_k\mathbf{C}^H\widehat{\mathbf{H}}_k^H$. We denote the covariance matrix of the noise and interference of the $k^{\text{th}}$ user by $\mathbf{V}_k$. Using the fact that the terms involved are uncorrelated, it is easy to show that

$$\mathbf{V}_k = \mathbf{I}_{N_r} + \rho_{dl}\widehat{\mathbf{H}}_k\mathbf{C}\sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{\Phi}_j\mathbf{C}^H\widehat{\mathbf{H}}_k^H + \underbrace{\rho_{dl}\mathbb{E}\left[\widetilde{\mathbf{H}}_k\mathbf{x}\mathbf{x}^H\widetilde{\mathbf{H}}_k^H\right]}_{\text{Due to channel est. errors}}. \tag{5.7}$$

It is shown in Sec. 5.6.1 that

$$\mathbb{E}\left[\widetilde{\mathbf{H}}_k\mathbf{x}\mathbf{x}^H\widetilde{\mathbf{H}}_k^H\right] = (\beta_k - \gamma_k)\,\mathbf{I}_{N_r}. \tag{5.8}$$

Substituting (5.8) in (5.7) and simplifying, we get

$$\mathbf{V}_k = \sigma_k^2\mathbf{I}_{N_r} + \widehat{\widehat{\mathbf{H}}}_k \sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{\Phi}_j\widehat{\widehat{\mathbf{H}}}_k^H, \tag{5.9}$$

where $\widehat{\widehat{\mathbf{H}}}_k \triangleq \sqrt{\rho_{dl}}\widehat{\mathbf{H}}_k\mathbf{C}$ and $\sigma_k^2 \triangleq (1 + \rho_{dl}\,(\beta_k - \gamma_k))$. Now, since the interference terms are uncorrelated with the desired signal by virtue of MMSE estimation, using the worst case noise theorem [172], the achievable rate of the $k^{\text{th}}$ user and the downlink sum rate are given by

$$R_k = \log\left|\mathbf{I}_{N_r} + \mathbf{V}_k^{-1}\widehat{\widehat{\mathbf{H}}}_k\mathbf{\Phi}_k\widehat{\widehat{\mathbf{H}}}_k^H\right|. \tag{5.10}$$

Our goal is to maximize the sum rate $R_{\text{tot}} = \sum_{k=1}^{K} R_k$ under a total power constraint:

$$\underset{\substack{\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_2,\ldots,\boldsymbol{\Phi}_K \\ \boldsymbol{\Phi}_k \text{ diagonal, p.s.d.}}}{\text{maximize}} \sum_{k=1}^{K} \log \left| \mathbf{I}_{N_r} + \mathbf{V}_k^{-1} \widehat{\widehat{\mathbf{H}}}_k \boldsymbol{\Phi}_k \widehat{\widehat{\mathbf{H}}}_k^H \right|, \tag{5.11}$$

$$\text{subject to } \text{tr}\Big( \sum_{k=1}^{K} \boldsymbol{\Phi}_k \Big) = 1.$$

The optimization problem in (5.11) is nonconvex in $\boldsymbol{\Phi}_1, \ldots, \boldsymbol{\Phi}_K$ due to the $\mathbf{V}_k^{-1}$ term, and cannot be solved in closed-form. Note that, we restrict the power allocation matrices to be diagonal in order to be implementable under codebook based precoding. This constrains the precoding matrices to belong to the finite set of matrices that can be expressed as the sum of outer products of codebook vectors weighted by the corresponding power allocation, and makes the problem significantly harder than unconstrained designs of precoding matrices [156, 157].

In this work, we present two algorithms based on the MM principle, which proceeds by finding a surrogate function that is a lower bound on the objective function, followed by maximizing the surrogate cost function, iteratively, until convergence to a local optimum. An excellent tutorial on the MM principle can be found in [64].

## 5.3 Minorization-Maximization Algorithms for Sum Rate Maximization

In this section, we present our solutions to the beamforming vector selection and power allocation problem stated in (5.11). We develop two algorithms, namely, the square root MM (SMM) and inverse MM (IMM) algorithms. These algorithms start with a common minorization step, and then solve the resulting optimization problem by two different approaches.

The first step in finding a computationally efficient solution to (5.11) is to find a surrogate function which is a lower bound on the sum rate, and is tight at the current iterate. To this end, consider the function $f(\mathbf{Z}, \mathbf{Y}) = \log |\mathbf{Z}^{-1}\mathbf{Y}|$, for $\mathbf{Z}$, $\mathbf{Y} \succeq 0$. This

function is convex in $\mathbf{Z}$, $\mathbf{Y}^{-1}$. Hence, we can bound it from below using the first order Taylor series expansion, as given by the following Lemma:

**Lemma 8.** *For matrices $\mathbf{Z}$, $\mathbf{Y} \succeq 0$, the function*

$$f(\mathbf{Z}, \mathbf{Y}) = \log \left| \mathbf{Z}^{-1}\mathbf{Y} \right|$$

*can be lower bounded by*

$$\begin{aligned} f(\mathbf{Z}, \mathbf{Y}) \geq -\Big( & \log \left|\mathbf{Z}^{(m)}\right| + \mathrm{tr}\left(\mathbf{Z}^{(m)^{-1}}\left(\mathbf{Z}-\mathbf{Z}^{(m)}\right)\right) \\ & + \log \left|\mathbf{Y}^{(m)^{-1}}\right| + \mathrm{tr}\left(\mathbf{Y}^{(m)}\left(\mathbf{Y}^{-1}-\mathbf{Y}^{(m)^{-1}}\right)\right)\Big) \end{aligned}$$

*with equality at $\mathbf{Z} = \mathbf{Z}^{(m)}$ and $\mathbf{Y} = \mathbf{Y}^{(m)}$. (Later, $m$ will be used to denote the iteration index).*

Returning to our problem, we define an intermediate matrix

$$\mathbf{B}_k \triangleq \sigma_k^2 \mathbf{I}_{N_r} + \sum_{j=1}^{K} \hat{\hat{\mathbf{H}}}_k \mathbf{\Phi}_j \hat{\hat{\mathbf{H}}}_k^H . \tag{5.12}$$

The rate of the $k^{\text{th}}$ user in (5.10) can then be written as $R_k = \log \left| \mathbf{V}_k^{-1}\mathbf{B}_k \right|$. Using Lemma 8, we get the following surrogate optimization problem for (5.11):

$$\begin{aligned} & \{\mathbf{\Phi}_1^{(m+1)}, \ldots, \mathbf{\Phi}_K^{(m+1)}\} \\ & = \underset{\mathbf{\Phi}_1,\ldots,\mathbf{\Phi}_K}{\mathrm{argmax}} \sum_{k=1}^{K} \Bigg\{ - \mathrm{tr}\left( \mathbf{V}_k^{(m)^{-1}} \Big( \sigma_k^2 \mathbf{I}_{N_r} + \sum_{\substack{j=1 \\ j\neq k}}^{K} \hat{\hat{\mathbf{H}}}_k \mathbf{\Phi}_j \hat{\hat{\mathbf{H}}}_k^H \Big) \right) \\ & \qquad\qquad - \mathrm{tr}\left( \mathbf{B}_k^{(m)} \Big[ \sigma_k^2 \mathbf{I}_{N_r} + \sum_{j=1}^{K} \hat{\hat{\mathbf{H}}}_k \mathbf{\Phi}_j \hat{\hat{\mathbf{H}}}_k^H \Big]^{-1} \right) \Bigg\}, \\ & \qquad \text{subject to} \quad \mathrm{tr}\Big( \sum_{k=1}^{K} \mathbf{\Phi}_k \Big) \leq 1, \end{aligned} \tag{5.13}$$

where $m$ is the iteration index. Here, we omit the $\log_e 2$ term in the denominator, as it does not affect the solution. In (5.13), the quantities $\mathbf{V}_k^{(m)}$ and $\mathbf{B}_k^{(m)}$ are computed by

substituting $\mathbf{\Phi}_k^{(m)}$ for $\mathbf{\Phi}_k$ in (5.9) and (5.12), respectively. Now, if we are able to solve the surrogate problem in (5.13), then, starting from an arbitrary initialization for $\boldsymbol{\Phi_k}$, the MM procedure iterates between solving (5.13) and updating $\mathbf{V}_k$ and $\mathbf{B}_k$. By virtue of the fact that the cost function increases in each iteration and is bounded above (for example, by the sum of the best rates achievable by each individual user), such a procedure is guaranteed to converge to a local optimum from any initialization.

Now, the optimization problem in (5.13) is a semidefinite program (SDP). However, the matrices $\{\mathbf{\Phi}_k\}_{k=1}^{k=K}$ are coupled in the objective function and constraints, making it a large dimensional problem. Due to this, SDP based methods such as `sdpsol` to solve (5.13) can quickly become computationally prohibitive as the number of users, the size of the codebook, and/or number of antennas gets large. Hence, there is a need to find alternative, computationally inexpensive approaches to solving (5.13). The SMM and IMM algorithms employ two different surrogate functions to further lower bound the objective function, in turn, leading to a surrogate cost function that is more amenable to optimization. In fact, we are able to solve the final surrogate problem in closed-form.

Before discussing the SMM and IMM algorithms further, we define some notation and simplify the first term in the objective function in (5.13). Let

$$\mathbf{\Phi} \triangleq \operatorname{diag}\left(\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_K\right), \tag{5.14}$$

$$\mathbf{\Psi}_k \triangleq \left[\widehat{\widehat{\mathbf{H}}}_k, \ldots, \widehat{\widehat{\mathbf{H}}}_k\right], k = 1, \ldots, K \tag{5.15}$$

denote the augmented power allocation and the $k^{\text{th}}$ user's channel matrices, respectively. In (5.15), $\widehat{\widehat{\mathbf{H}}}_k$ is repeated $K$ times. Also, let

$$\mathbf{Q} \triangleq \sum_{k=1}^{K} \operatorname{diag}\left(\widehat{\widehat{\mathbf{H}}}_k^H \mathbf{V}_k^{-1} \widehat{\widehat{\mathbf{H}}}_k, \ldots, \mathbf{0}_N, \ldots, \widehat{\widehat{\mathbf{H}}}_k^H \mathbf{V}_k^{-1} \widehat{\widehat{\mathbf{H}}}_k\right). \tag{5.16}$$

In the above, the $N \times N$ all zero matrix $\mathbf{0}_N$ is in the $k^{\text{th}}$ block diagonal position of $\mathbf{Q}$.

Excluding the constant noise variance part, we can rewrite the first term of (5.13) as

$$\sum_{k=1}^{K} \mathrm{tr}\left( \mathbf{V}_k^{(m)^{-1}} \Big( \sum_{\substack{j=1 \\ j\neq k}}^{K} \hat{\hat{\mathbf{H}}}_k \mathbf{\Phi}_j \hat{\hat{\mathbf{H}}}_k^H \Big) \right) = \mathrm{tr}\left( \mathbf{Q}^{(m)} \mathbf{\Phi} \right), \tag{5.17}$$

where the superscript $m$ denotes the iteration index, and $\mathbf{Q}^{(m)}$ is obtained by substituting $\mathbf{V}_k^{(m)}$ for $\mathbf{V}_k$ in (5.16). We are now ready to describe the SMM and IMM algorithms in detail.

### 5.3.1 Square-Root Minorization Maximization Procedure

The square root MM procedure involves working with the square root of the power allocation matrix $\mathbf{\Phi}$. It also involves two stages of minorization. The result is a surrogate objective function that is a lower bound on the cost function in (5.13), is tight at the current iterate, and is easy to optimize. First, note that, with the notation defined in (5.14) and (5.15), the second term in (5.13) can be written as

$$-\sum_{k=1}^{K} \mathrm{tr}\left( \mathbf{F}_k^{(m)} \left( \sigma_k^2 \mathbf{I}_{N_r} + \mathbf{\Psi}_k \mathbf{\Phi} \mathbf{\Psi}_k^H \right)^{-1} \mathbf{F}_k^{(m)^H} \right), \tag{5.18}$$

where $\mathbf{F}_k$ is such that $\mathbf{B}_k = \mathbf{F}_k^H \mathbf{F}_k$, and can be computed via the Cholesky decomposition of $\mathbf{B}_k$. The above cost function cannot be directly optimized due to the matrix inversion involved. Hence, we minorize it using the following lemma.

**Lemma 9.** *Let* $\mathbf{R}$ *denote a diagonal p.s.d. square matrix, and consider the function*

$$f(\mathbf{R}) \triangleq -\mathrm{tr}\left( \mathbf{A} \left( \mathbf{B} + \mathbf{C}\mathbf{R}\mathbf{C}^H \right)^{-1} \mathbf{A}^H \right), \tag{5.19}$$

*where* $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{C}$ *are matrices of compatible dimensions, and* $\mathbf{B} \succ 0$*, so that* $\mathbf{B} + \mathbf{C}\mathbf{R}\mathbf{C}^H$ *is invertible. Then, for a given diagonal p.s.d. matrix* $\mathbf{R}^{(m)}$*,* $f(\mathbf{R})$ *can be lower bounded by*

$$f(\mathbf{R}) \geq g(\mathbf{R}|\mathbf{R}^{(m)})$$

$$\triangleq -\mathrm{tr}(\hat{\mathbf{K}}) + \mathrm{tr}\left(\left(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\mathbf{A}\mathbf{B}^{-1}\mathbf{C} + \mathbf{C}^H\mathbf{B}^{-H}\mathbf{A}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\right)\mathbf{R}^{\frac{1}{2}} - \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\right), \tag{5.20}$$

*where*

$$\hat{\mathbf{X}} \triangleq \mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{(m)\frac{1}{2}}, \quad \hat{\mathbf{Y}} \triangleq \mathbf{I} + \mathbf{R}^{(m)\frac{1}{2}}\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{(m)\frac{1}{2}},$$
$$\hat{\mathbf{K}} \triangleq \mathbf{A}\mathbf{B}^{-1}\mathbf{A}^H + \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}} - \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\hat{\mathbf{Y}} + \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1} + \hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H.$$

*Also,* $g(\mathbf{R}^{(m)}|\mathbf{R}^{(m)}) = f(\mathbf{R}^{(m)})$.

*Proof.* See Sec. 5.6.2. □

The objective function in (5.18) is in the same form as the function in Lemma 9. Applying Lemma 9 to (5.18), we get

$$-\mathrm{tr}\left(\mathbf{W}_{1,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_{2,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right), \tag{5.21}$$

where

$$\mathbf{W}_{1,k} \triangleq -\left\{\frac{\mathbf{Y}_k^{-1}\mathbf{X}_k^H\mathbf{F}_k\mathbf{\Psi}_k + \mathbf{\Psi}_k^H\mathbf{F}_k^H\mathbf{X}_k\mathbf{Y}_k^{-1}}{\sigma_k^2}\right\}, \tag{5.22}$$

$$\mathbf{W}_{2,k} \triangleq \mathbf{Y}_k^{-1}\mathbf{X}_k^H\mathbf{X}_k\mathbf{Y}_k^{-1}, \tag{5.23}$$

and $\mathbf{S}_k \in \mathbb{C}^{KN\times KN}, \mathbf{X}_k \in \mathbb{C}^{N_r\times KN}$ and $\mathbf{Y}_k \in \mathbb{C}^{KN\times KN}$ are defined as

$$\mathbf{S}_k \triangleq \frac{\mathbf{\Psi}_k^H\mathbf{\Psi}_k}{\sigma_k^2},\ \mathbf{X}_k \triangleq \frac{\mathbf{F}_K\mathbf{\Psi}_k\mathbf{\Phi}^{\frac{1}{2}}}{\sigma_k^2},\ \mathbf{Y}_k \triangleq \mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}.$$

Note that, $\mathbf{W}_{1,k}$ and $\mathbf{W}_{2,k}$ in (5.22) and (5.23), are negative and p.s.d. matrices, and hence, their diagonal entries are non-positive and non-negative, respectively. Also, $\mathbf{Y}_k^{-1}$ can be computed with low complexity using Woodbury matrix identity (requiring only a $N_r \times N_r$ matrix inverse instead of a $KN \times KN$ matrix inverse). Substituting (5.21) into (5.13), we get the surrogate cost function that needs to be maximized. Notice that

the matrix inversion in (5.18) has been circumvented by the use of the lower bound. However, the surrogate cost function is not yet amenable to a closed-form solution due to the $\mathbf{W}_{2,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}$ term in (5.21). Hence, we minorize the second term in (5.21) again to get a cost function that is easy to optimize. To this end, we need the following Lemma.

**Lemma 10.** *Suppose* $\mathbf{R}$ *is a p.s.d. diagonal matrix, and* $\mathbf{A}$ *and* $\mathbf{B}$ *are symmetric p.s.d. square matrices. Then, the function* $f(\mathbf{R}) \triangleq -\mathrm{tr}\left(\mathbf{ARBR}\right)$ *can be lower bounded by*

$$\begin{aligned} f(\mathbf{R}) \geq & -\mathrm{tr}\left(\mathbf{A}\mathbf{R}^{(m)}\mathbf{B}\mathbf{R}^{(m)} - \left(\left(\mathbf{B}-\lambda\mathbf{I}\right)\mathbf{R}^{(m)}\mathbf{A} + \mathbf{A}\mathbf{R}^{(m)}\left(\mathbf{B}-\lambda\mathbf{I}\right)\right)\mathbf{R}^{(m)}\right) \\ & - \mathrm{tr}\left(\left(\left(\mathbf{B}-\lambda\mathbf{I}\right)\mathbf{R}^{(m)}\mathbf{A} + \mathbf{A}\mathbf{R}^{(m)}\left(\mathbf{B}-\lambda\mathbf{I}\right)\right)\mathbf{R}\right) - \lambda\mathrm{tr}\left(\mathbf{A}\mathbf{R}^2\right), \end{aligned} \tag{5.24}$$

*where* $\lambda$ *is the largest eigenvalue of* $\mathbf{B}$*. Further, we have equality in* (5.24) *at* $\mathbf{R} = \mathbf{R}^{(m)}$*.*

*Proof.* See Sec. 5.6.3. □

Applying Lemma 10 to (5.21), we get the final lower bound for (5.18) as follows:

$$-\sum_{k=1}^{K}\mathrm{tr}\left(\mathbf{W}_{1,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_{2,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right) \geq -\mathrm{tr}\left(\mathbf{W}_A^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_B^{(m)}\mathbf{\Phi}\right), \tag{5.25}$$

where

$$\mathbf{W}_A \triangleq \sum_{k=1}^{K}\left(\mathbf{W}_{1,k} + \left(\mathbf{S}_k - \lambda_{\max}\left(\mathbf{S}_k\right)\mathbf{I}_{KN}\right)\mathbf{\Phi}^{\frac{1}{2}}\mathbf{W}_{2,k}\right), \tag{5.26}$$

$$\mathbf{W}_B \triangleq \sum_{k=1}^{K}\lambda_{\max}\left(\mathbf{S}_k\right)\mathbf{W}_{2,k}, \tag{5.27}$$

and $\lambda_{\max}(\mathbf{S}_k)$ is the largest eigenvalue of $\mathbf{S}_k$. Note that the superscript $m$ in (5.21) and (5.25) is the iteration index. Also, we can compute the eigenvalues of $\mathbf{S}_k$ by multiplying the eigenvalues of the smaller dimensional matrix $\widehat{\widehat{\mathbf{H}}}_k^H\widehat{\widehat{\mathbf{H}}}_k/\sigma_k^2$ by the number of users in the system. Thus, $\lambda_{\max}(\mathbf{S}_k)$ needs to be computed only once and stored in the memory. Combining (5.25) with $\mathrm{tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi}\right)$ in (5.17), the optimization problem we wish to solve

becomes

$$\{\mathbf{\Phi}^{(m+1)}\} = \underset{\mathbf{\Phi}}{\operatorname{argmax}} \left\{ -\operatorname{tr}\left( \mathbf{Q}^{(m)}\mathbf{\Phi} + \mathbf{W}_A^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_B^{(m)}\mathbf{\Phi} \right) \right\} \quad \text{subject to} \quad \operatorname{tr}(\mathbf{\Phi}) \leq 1. \tag{5.28}$$

**Lemma 11.** *The optimization problem in* (5.28) *has a locally optimal solution given by*

$$P(i) = \left( \frac{\left[\mathbf{W}_A^{(m)}\right]_{(i,i)}}{2\left( \left[\mathbf{W}_B^{(m)}\right]_{(i,i)} + \left[\mathbf{Q}^{(m)}\right]_{(i,i)} + \eta \right)} \right)^2, \quad \forall i, \tag{5.29}$$

*where $\eta$ is chosen to satisfy $\sum_{i=1}^{KN} P(i) = 1$.*

*Proof.* See Sec. 5.6.4. □

Using the solution for $P(i)$, one can construct the new surrogate function that needs to be optimized in the next iteration. Iterating the process of computing $P(i)$, we arrive at a locally optimal joint power and beamforming vector allocation solution for maximizing the sum rate.

We next present an alternative bounding approach which leads to a different MM procedure for sum rate maximization.

### 5.3.2 Inverse Minorization Maximization Procedure

We now return to the original optimization problem in (5.13). Recall that $\mathbf{\Phi} \in \mathbb{R}^{KN \times KN}$ is the augmented transmit power allocation matrix defined in (5.14). For convenience, let us define an augmented covariance matrix $\widetilde{\mathbf{\Phi}} \in \mathbb{R}^{(KN+N_r)\times(KN+N_r)}$, an augmented channel matrix $\widetilde{\mathbf{\Psi}}_k \in \mathbb{C}^{N_r \times (KN+N_r)}$ and the matrix $\mathbf{\Xi}_k \in \mathbb{C}^{N_r \times N_r}$ as follows:

$$\widetilde{\mathbf{\Phi}} \triangleq \operatorname{diag}\left(\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_K, \sigma^2 \mathbf{I}_{N_r}\right), \tag{5.30}$$

$$\widetilde{\mathbf{\Psi}}_k \triangleq \left[\widehat{\widehat{\mathbf{H}}}_k, \ldots, \widehat{\widehat{\mathbf{H}}}_k, \mathbf{I}_{N_r}\right], \tag{5.31}$$

$$\mathbf{\Xi}_k \triangleq \widetilde{\mathbf{\Psi}}_k \widetilde{\mathbf{\Phi}} \widetilde{\mathbf{\Psi}}_k^H, \tag{5.32}$$

$k = 1, \ldots, K$. In the definition of $\widetilde{\mathbf{\Psi}}$ above, the matrix $\widehat{\widehat{\mathbf{H}}}_k$ is repeated $K$ times. Then, we can rewrite the term inside the square brackets in (5.13) as $\mathbf{B}_k^{(m)}\mathbf{\Xi}_k^{-1}$. Note that the matrix $\mathbf{\Xi}_k \,\forall\, k$ is p.s.d., which will be useful in showing that the optimization problem has a feasible solution. For simplicity, we assumed that $\sigma_k$ is same for all the users to define $\widetilde{\mathbf{\Phi}}$. This is a valid assumption in an interference dominated regime, with path loss inversion based power control.

In order to develop the IMM procedure, we start with the following proposition from [173].

**Proposition 1.** *Let* $\mathbf{R}$ *be an* $m \times n$ *matrix and* $\mathbf{A}$ *be an* $m \times m$ *p.s.d. matrix. We can upper bound the function* $f(\mathbf{U}) \triangleq \mathrm{tr}\left(\mathbf{A}\left(\mathbf{R}\mathbf{U}\mathbf{R}^H\right)^{-1}\right)$ *as*

$$f(\mathbf{U}) \leq \mathrm{tr}\Big(\mathbf{A}\big(\mathbf{R}\mathbf{U}^{(m)}\mathbf{R}^H\big)^{-1}\mathbf{R}\mathbf{U}^{(m)}\mathbf{U}^{-1}\mathbf{U}^{(m)}\mathbf{R}^H\big(\mathbf{R}\mathbf{U}^{(m)}\mathbf{R}^H\big)^{-1}\Big), \tag{5.33}$$

*with equality at* $\mathbf{U} = \mathbf{U}^{(m)}$.

Since $\mathbf{B}_k^{(m)} \succeq 0 \,\forall\, k$, we can apply proposition 1 to $\mathrm{tr}\left(\mathbf{B}_k^{(m)}\mathbf{\Xi}_k^{-1}\right)$, which leads to

$$\sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{B}_k^{(m)}\mathbf{\Xi}_k^{-1}\right) \leq \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{B}_k^{(m)}{\mathbf{\Xi}_k^{(m)}}^{-1}\widetilde{\mathbf{\Psi}}_k\widetilde{\mathbf{\Phi}}^{(m)}\widetilde{\mathbf{\Phi}}^{-1}\widetilde{\mathbf{\Phi}}^{(m)}\widetilde{\mathbf{\Psi}}_k^H{\mathbf{\Xi}_k^{(m)}}^{-1}\right) \tag{5.34}$$

$$= \mathrm{tr}\left(\sum_{k=1}^{K}\widetilde{\mathbf{\Phi}}^{(m)}\widetilde{\mathbf{\Psi}}_k^H{\mathbf{\Xi}_k^{(m)}}^{-1}\widetilde{\mathbf{\Psi}}_k\widetilde{\mathbf{\Phi}}^{(m)}\widetilde{\mathbf{\Phi}}^{-1}\right), \tag{5.35}$$

where (5.35) is obtained by recognizing that $\mathbf{B}_k^{(m)}{\mathbf{\Xi}_k^{(m)}}^{-1}$ is the identity matrix, cyclically permuting the terms, and pulling the summation over $k$ into the trace function. In (5.35), the matrix $\tilde{\mathbf{\Phi}}$ is diagonal and positive semi-definite, which may become singular. On the other hand, Proposition 1 assumes it to be an invertible matrix for deriving the upper bound to the objective function of the optimization problem and obtain a closed form solution. However, this does not pose a problem in practice, because if some of the diagonal entries of $\tilde{\mathbf{\Phi}}$ become 0, we can remove the corresponding columns and rows of the matrices in the left hand side and right hand side of $\tilde{\mathbf{\Phi}}$ in (5.34), and form a new nonsingular $\tilde{\mathbf{\Phi}}$. During initialization, we allocate equal or random powers to all the users

across all the beamforming vectors, which makes $\tilde{\mathbf{\Phi}}$ invertible. Therefore, without loss of generality, we can assume that $\tilde{\mathbf{\Phi}}$ is invertible. Further, in the final iterative algorithm, we compute the power allocations using closed form expressions, which converge to a stationary point of the optimization problem. Now, let

$$\mathbf{Z} \triangleq \sum_{k=1}^{K} \widetilde{\mathbf{\Phi}} \widetilde{\mathbf{\Psi}}_k^H \mathbf{\Xi}_k^{-1} \widetilde{\mathbf{\Psi}}_k \widetilde{\mathbf{\Phi}}. \tag{5.36}$$

Substituting $\mathrm{tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi}\right)$ (from (5.17)) and (5.35) into (5.11), we get the following surrogate optimization problem:

$$\begin{aligned} \mathbf{\Phi}^{(m+1)} = & \arg\max_{\mathbf{\Phi}\succeq 0} \left\{ -\mathrm{tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi} + \mathbf{Z}^{(m)}\widetilde{\mathbf{\Phi}}^{-1}\right)\right\} \\ & \text{subject to} \quad \mathrm{tr}\left(\mathbf{\Phi}\right) \leq 1, \end{aligned} \tag{5.37}$$

where $m$ is the iteration index.

**Lemma 12.** *The optimization problem in* (5.37) *has a locally optimal solution given by*

$$P(i) = \left( \frac{\left[\mathbf{Z}^{(m)}\right]_{(i,i)}}{\left[\mathbf{Q}^{(m)}\right]_{(i,i)} + \eta} \right)^{\frac{1}{2}}, \quad i = 1, \ldots, KN, \tag{5.38}$$

*where $\eta$ is chosen to satisfy $\sum_{i=1}^{KN} P(i) = 1$.*

*Proof.* See Sec. 5.6.5. □

In each iteration, we bound the cost function using proposition 1, and maximize the corresponding surrogate cost function using the solution in (5.38). Then, we recompute the bounding function, and the process repeats till convergence. Since (5.38) is strictly decreasing in $\eta$, we can determine the value of $\eta$ for which the solution satisfies the power constraint using a line search or bisection method [174].

The outcome of both the SMM and IMM procedures is the matrix $\mathbf{\Phi}$, which gives the individual users' powers across all the beamforming vectors. The power allocated to the $k^{\text{th}}$ user on the $j^{\text{th}}$ beamforming vector can be written using the solutions from the SMM

Table 5.1:

Flop count order of SMM per iteration

| **Matrix** | **Size** | **Flop Count** |
|---|---|---|
| $\mathbf{S}_k$ | $KN \times KN$ | $K^2N^2N_r$ |
| $\mathbf{X}_k$ | $N_r \times KN$ | $KNN_r^2$ |
| $\mathbf{Y}_k^{-1}$ | $KN \times KN$ | $KNN_r^3$ |
| $\mathbf{W}_{1,k}$ | $KN \times KN$ | $KN$ |
| $\mathbf{W}_{2,k}$ | $KN \times KN$ | $K^2N^2N_r$ |
| $\mathbf{W}_{A,k}$ | $KN \times KN$ | $K^3N^3$ |
| $\mathbf{W}_{B,k}$ | $KN \times KN$ | $KN$ |
| $\mathbf{W}_A$ | $KN \times KN$ | $K^2N$ |
| $\mathbf{W}_B$ | $KN \times KN$ | $K^2N$ |

or IMM procedure as

$$P_k(j) = P\left((k-1)\,N + j\right). \tag{5.39}$$

Pseudo-codes for SMM and IMM are shown in Sec. 5.7.

### 5.3.3 Computational Complexity

We use floating point operations (flops) to quantify the computational complexity of the algorithms. We assume that the multiplication of a $p \times q$ matrix with a $q \times r$ matrix requires $\mathcal{O}(pqr)$ flops. The per-iteration computational complexity of the SMM and IMM algorithms are provided in Tables 5.1 and 5.2, respectively. The flop counts account for the structural properties of the various matrices. For example, while computing the flop counts for the matrix $\mathbf{Z}$ or $\mathbf{Q}$, we only consider the computations involved in finding

Table 5.2:

Flop count order of IMM per iteration

| **Matrix** | **Size** | **Flop Count** |
|---|---|---|
| $\mathbf{\Xi}_k$ | $N_r \times N_r$ | $KNN_r^2$ |
| $\mathbf{Z}$ | $(KN + N_r)$ $\times(KN + N_r)$ | $KNN_r^2$ |
| $\mathbf{Q}$ | $KN \times KN$ | $(N_r + K)N^2$ |

the diagonal entries of those matrices. Also, although the flop count for computing the matrix $\mathbf{Q}$ is mentioned only in Table 5.1, it is common for both the SMM and IMM algorithms. The overall computational complexities of the SMM and IMM algorithms are of the order $\mathcal{O}(K^4N^3)$ and $\mathcal{O}(K^2N^2)$, respectively. Thus, the per-iteration complexity of IMM is lower than that of the SMM algorithm. Finally, we note that the complexity of SMM and IMM are independent of the number of transmit antennas. In practice, one would typically scale the size of the codebook with the number of transmit antennas, for example, as $N = 2^{N_t B}$ when a codebook with resolution $B$-bits per antenna is used. This can be substituted in the tables to infer the dependence of the complexity on the number of antennas.

## 5.4 Simulation Results

In this section, we evaluate the performance of the SMM and IMM algorithms using Monte Carlo simulations. We consider $N_t = \{8, 16, 32, 64, 128\}$ transmit antennas at the AP and $N_r = \{1, 2\}$ antennas at each UE. The number of UEs is varied from $K = 1$ to 10. The channel coefficients are drawn i.i.d. from $\mathcal{CN}(0, 1)$. The AWGN at the receivers is also distributed as $\mathcal{CN}(0, 1)$ and is independent across receive antennas. For channel state feedback, we quantize the CSI using 4 to 8 bits per channel coefficient. The dynamic

range of the channel coefficients typically vary around three times the standard deviation from the mean, and we set the maximum and minimum values of the quantization levels accordingly, and use uniform quantization. We consider the size of the beamforming codebook varying from $N = 64$ (6 bits) to 1024 (10 bits), uniformly distributed on the $N$ dimensional complex unit sphere. The algorithms are initialized randomly for all the users across all the beams, such that the total power constraint is satisfied with equality. The algorithms are run till the normalized increase in the sum rate between two consecutive iterations is less than $10^{-4}$.

We compare the sum rate of the IMM algorithm against the WMMSE [157] and WSRMax [161] algorithms. These algorithms use perfect CSI to design the precoder matrix, and do not consider the codebook constraint in the optimization. Hence, for comparison with our work, we quantize the precoding vectors output by the above algorithms to the nearest vector in the codebook, and compare the sum rates achieved. We also compare against the eigen-mode beamforming (EBF), a heuristic precoding approach (User-BFVec Selection) in which each UE maximizes its achievable sum rate using the IMM algorithm and feeds back the selected beamforming vectors to the AP, and the use of CVX [170,171] to solve (5.13).

Figure 5.1 shows the complementary cumulative distribution function (CCDF) of the achieved sum rates for the SMM and IMM algorithms at an SNR of 10 dB. We see that the two approaches offer similar sum rates. Figures 5.2 and 5.3 illustrate that although both SMM and IMM algorithms exhibit monotonic convergence, IMM converges much faster than SMM. This highlights the impact of the choice of the surrogate function on the rate of convergence [173]. In SMM, we apply the minorization three times to lower bound the objective function, whereas, in IMM, the minorization is applied twice. Also, SMM makes use of the first order Taylor series expansion, whereas, IMM uses a matrix inequality to find the surrogate function. These differences result in the different rates of convergence of the two procedures. We note that the associated matrix inequalities (Lemmas 9, 10, and Proposition 1) are potentially be useful in other problem scenarios. Given their similar performance, in the sequel, we do not include the SMM algorithm in

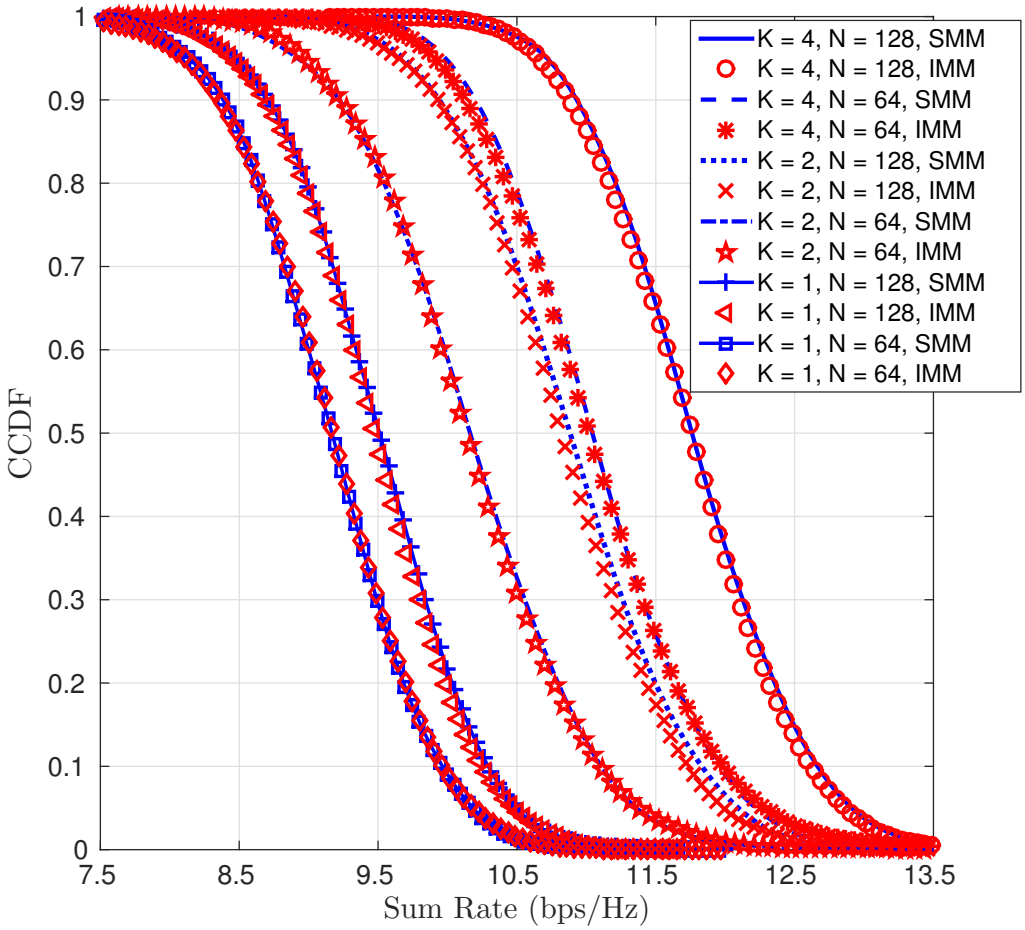


Figure 5.1: CCDF comparison between SMM and IMM for data SNR = 10 dB, $N_r = 2, N_t = 16$. The distribution of the sum rates achieved by SMM and IMM are similar.

the performance plots, to avoid clutter.

Figure 5.4 shows the achievable sum rate vs. the number of users, with $N = 512, N_r = 2, N_t = 16$ and data SNR 20 dB. The sum rate improves with finer quantization of channel estimates, as expected. However, increasing the number bits of quantization beyond 6 has a negligible impact on the performance. So, we have merged the curves for 6 bits and 8 bits to avoid clutter.

Figure 5.5 shows the achievable sum rate vs. the number of users, with $N = 512, N_r = 2, N_t = 16$ and CSI quantized to 6 bits per channel coefficient. We have merged the 10 dB and 20 dB pilot SNR plots, as their performances were very similar. While the sum rate

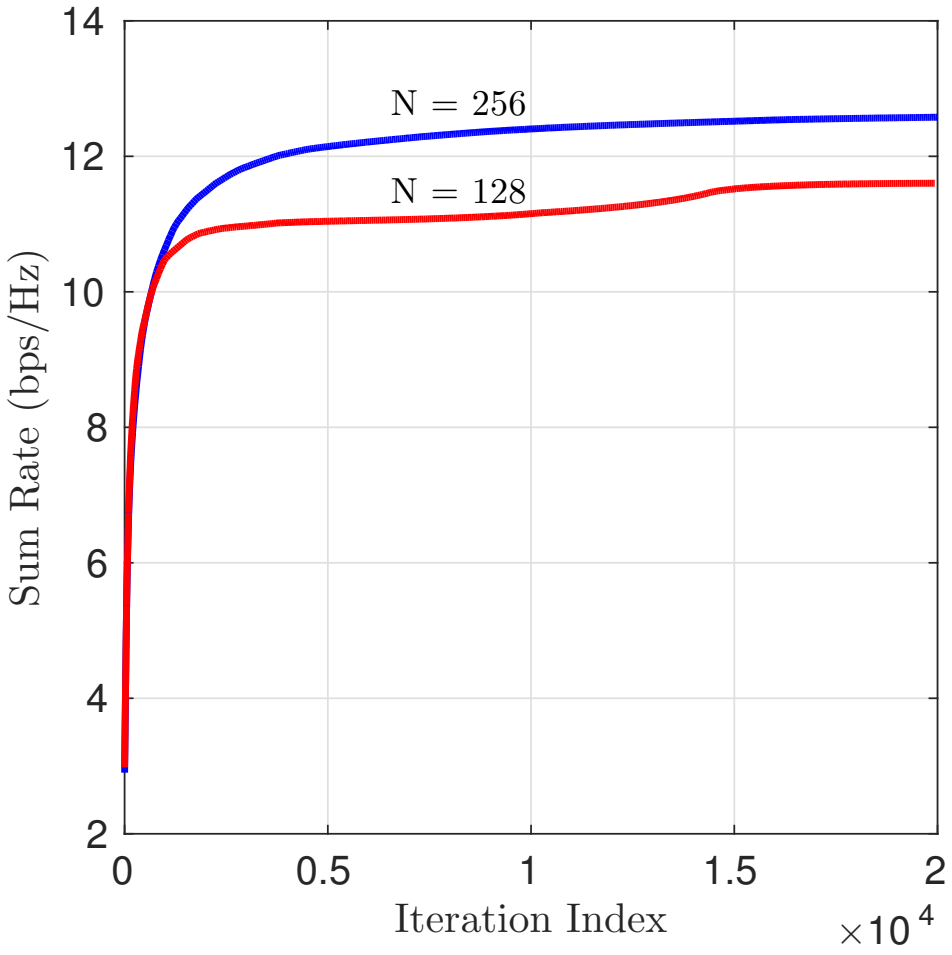


Figure 5.2: Convergence behavior of the SMM procedure, $K = 4, N_r = 2, N_t = 16$, Data SNR = 10 dB.

increases, the rate of increase decreases with the number of users. As $K$ is increased from $K = 1$, the algorithm has a larger number of choices to assign resources, which improves the sum rate. However, this only offers a marginal benefit at a given pilot/data SNR for large $K$. Also, the sum rate improvement with pilot SNR becomes marginal once it exceeds the data SNR. Thus, the AP can use these results to determine the pilot and data transmission powers in order to, for example, maximize energy efficiency while achieving a desired rate.

Figure 5.6 shows the sum rate vs. the pilot SNR, with $N = 512, K = 8, N_r = 2, N_t = 16$, and CSI quantized to 6 bits. The sum rate monotonically increases with the pilot

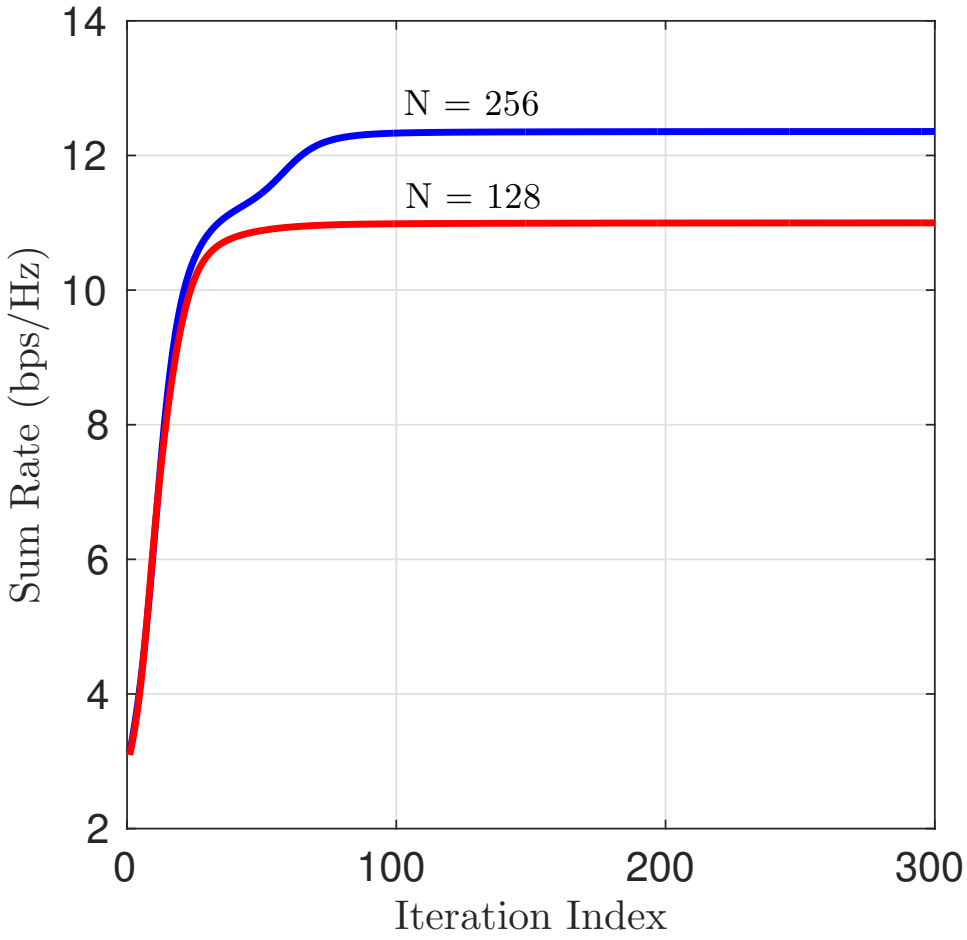


Figure 5.3: Convergence behavior of the IMM procedure, $K = 4, N_r = 2, N_t = 16$, Data SNR = 10 dB.

SNR, but saturates as the pilot SNR exceeds the data SNR, as observed earlier. From the achievable sum rate expression, the tipping point occurs when the channel estimation error variance is of the same order as that of the AWGN. The channel estimation error decreases linearly with the pilot SNR, while the residual interference caused due to channel estimation error increases linearly with the data SNR. Hence, if the pilot power scales linearly with the data power, it results in a roughly constant interference in the denominator of the SINR. Interestingly, this intuition continues to hold for the sum rate even after the beams and corresponding data powers are optimally chosen by the IMM algorithm. Hence, increasing the pilot SNR beyond the data SNR only marginally improves the sum

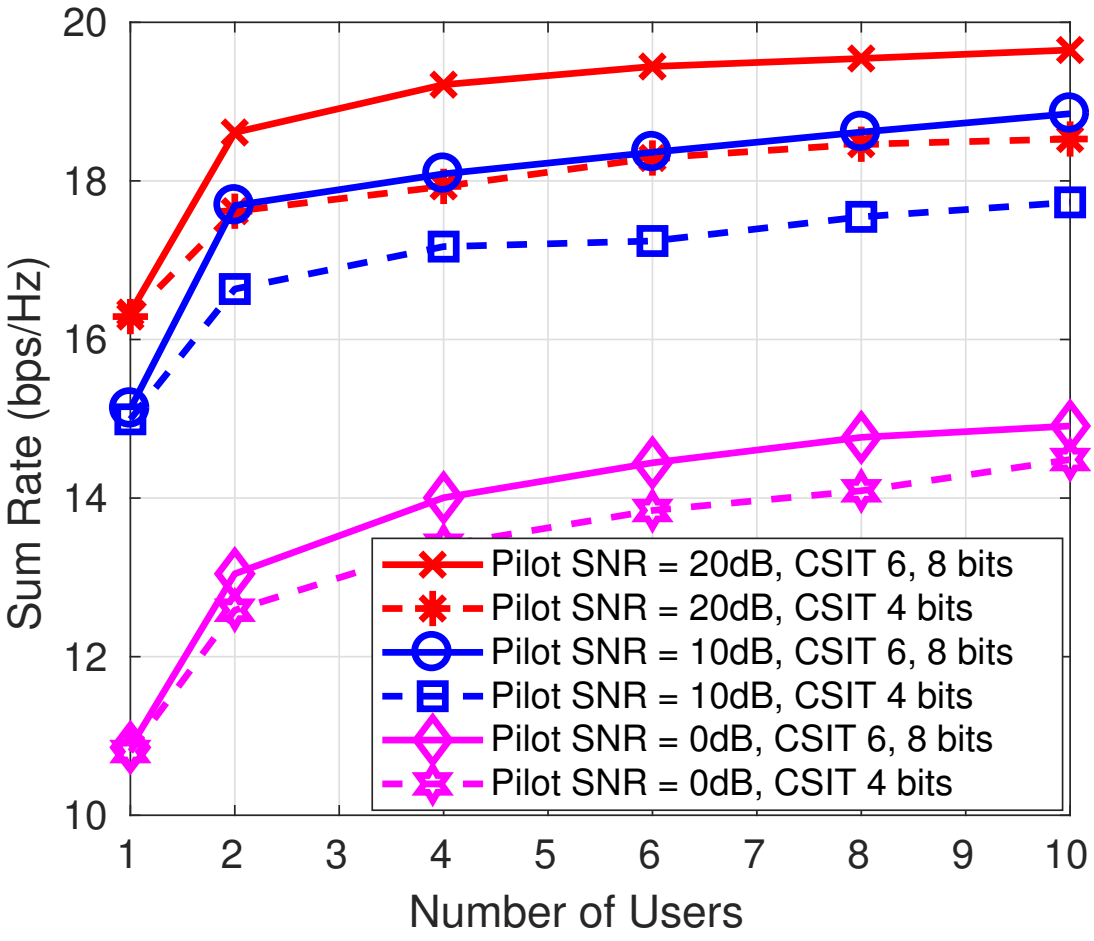


Figure 5.4: Sum rate vs. $K$, $N = 512, N_r = 2, N_t = 16$, data SNR = 20 dB. The sum rate improves with the number of CSI quantization bits, but beyond 6 bits, the performance improvement is negligible.

rate, and the system becomes noise and multiuser interference limited.

Figure 5.7 shows the sum rate vs. the data SNR, with $N = 512$, $K = 8, N_r = 2, N_t = 16$ and CSI quantized to 6 bits. The sum rate initially increases linearly with the data SNR, but begins to saturate once the data SNR exceeds the pilot SNR. For example, the sum rates achieved with pilot SNR = 10 dB and 20 dB match till a data SNR of 10 dB, with the sum rate increasing linearly with the data SNR. Beyond a data SNR of 10 dB, the performance with pilot SNR = 10 dB becomes limited by channel estimation errors, and the sum rate starts to saturate. This behavior is consistent with the observations made for Fig. 5.6. We also illustrate the sum rate performance of the WMMSE algorithm

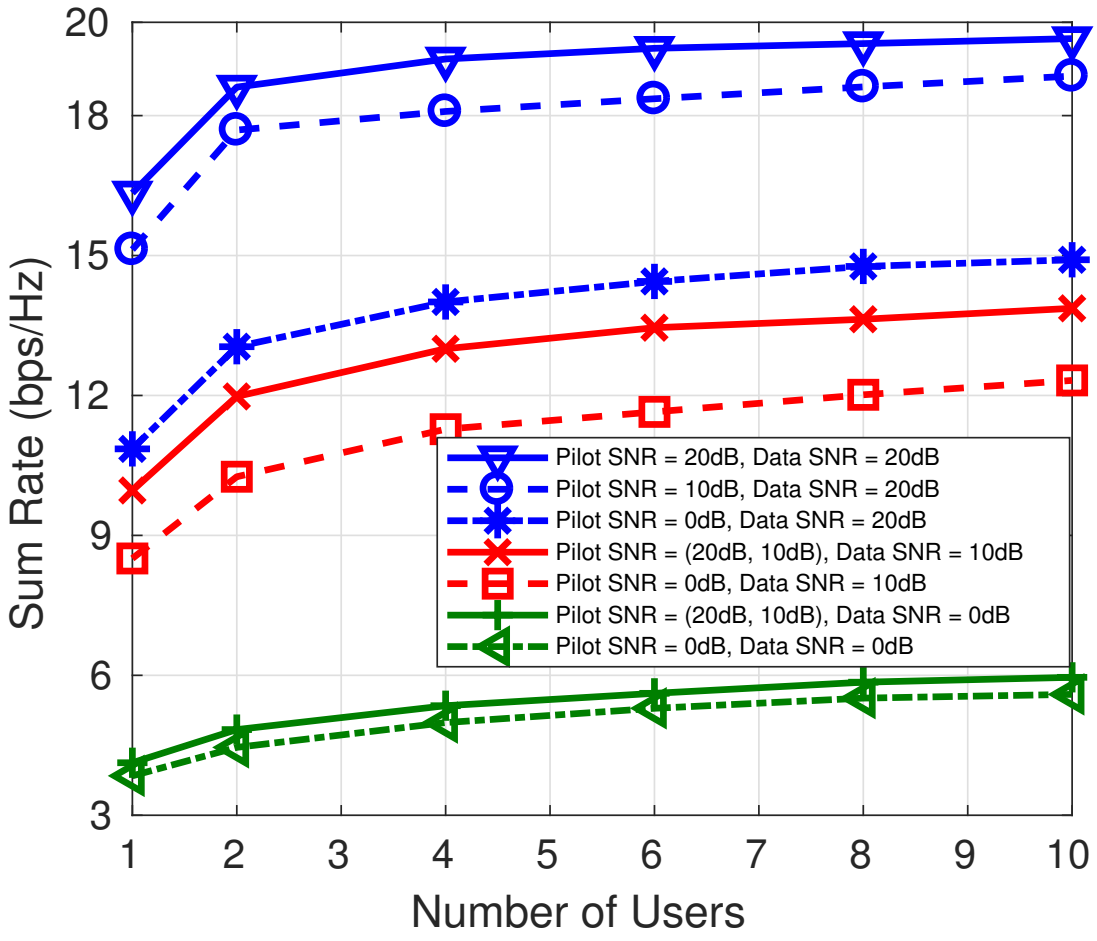


Figure 5.5: Sum rate vs. $K$, $N = 512, N_r = 2, N_t = 16$, CSI quantized to 6 bits. The sum rate improves with pilot SNR, but the improvement is marginal once the pilot SNR exceeds the data SNR.

in Fig. 5.7. We see that the sum rate performance of WMMSE is far inferior than that achieved by the MM based approaches when its precoder outputs are quantized to the nearest vectors in the codebook. The sum rate obtained using the non-codebook based WMMSE is higher than that of the MM algorithms, but it comes with the high overhead of conveying the precoding matrices to the users. For example, at a data SNR of 10 dB, the sum rate achieved by the unconstrained WMMSE is 36 bps/Hz (not shown in the figure), but when the precoders are quantized to the nearest beamforming vectors in the codebook, it drastically reduces to 6.46 bps/Hz. This illustrates the importance of considering the codebook constraint while solving the beamforming vector assignment

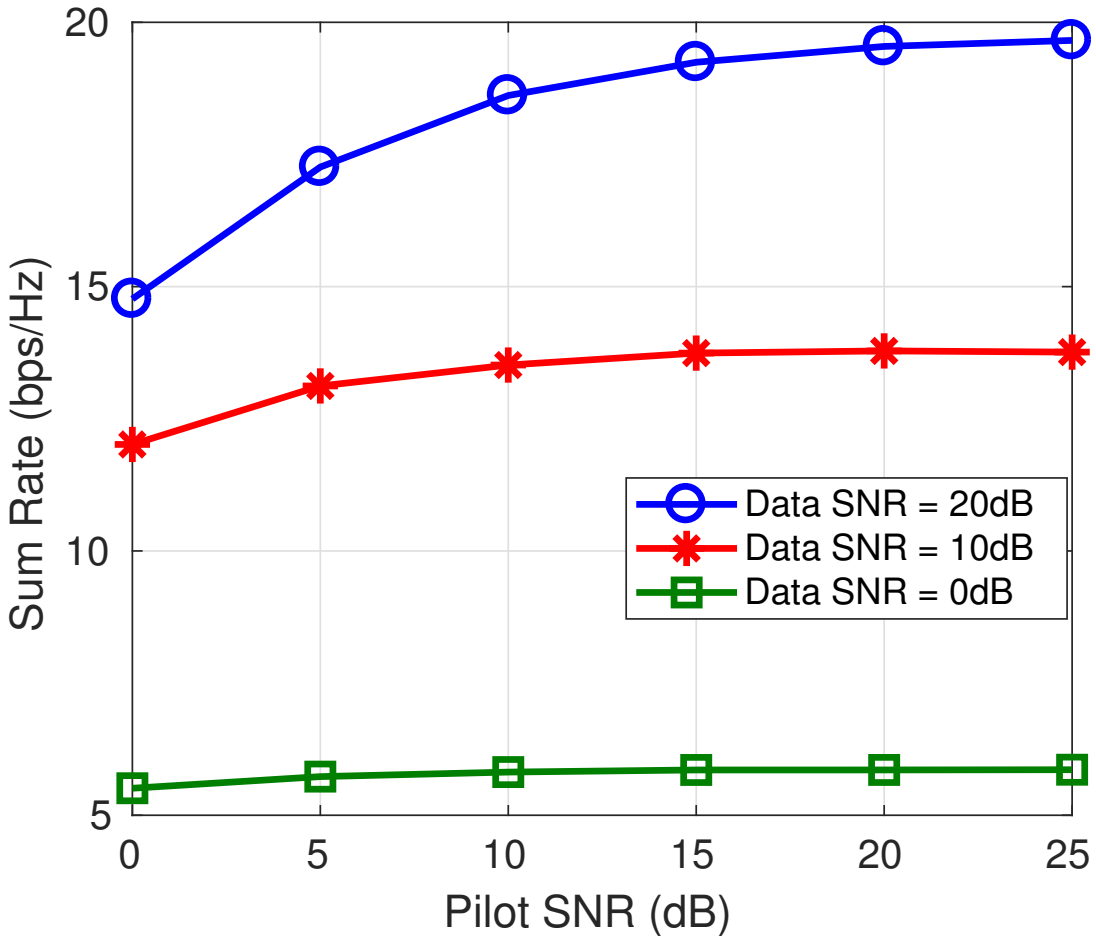


Figure 5.6: Sum rate vs. pilot SNR for $K = 8, N = 512, N_r = 2, N_t = 16$, CSI quantized to 6 bits. We observe a monotonic increase in the sum rate with pilot and data SNR. Eventually, the sum rate saturates beyond a certain level of the pilot SNR but the saturation point moves to the right as the data SNR increases.

and power allocation problem.

In Fig. 5.7, we also show the performance of the EBF and User-BFVec Selection. In the EBF method, each UE chooses the beamforming vectors as the $N_r$ dominant right singular vectors, and feeds back their quantized versions. In our simulations, we use 32 bits to quantize each selected beamforming vector. Note that, the feedback overhead associated with this is higher than the feedback overhead associated with the IMM approaches. The EBF approach outperforms the IMM algorithm at low SNRs. This is because the noise dominates the interference, and the multiuser interference term does not significantly

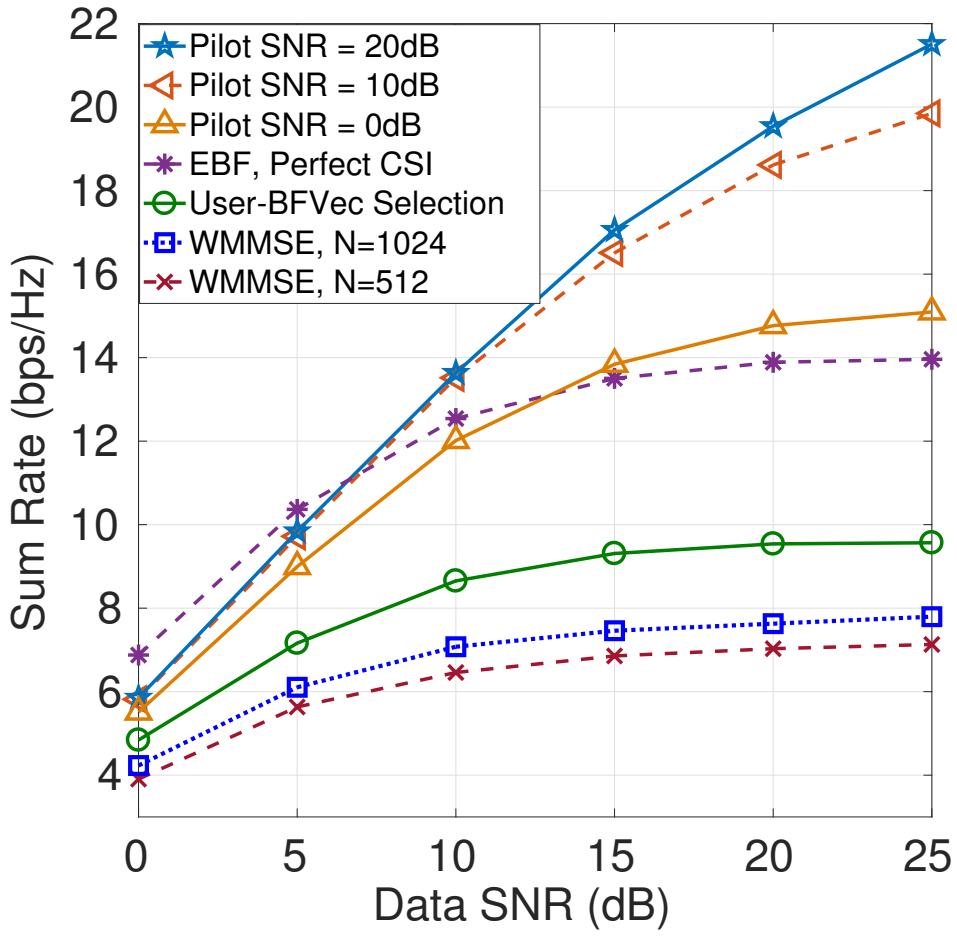


Figure 5.7: Sum rate vs. data SNR for $K = 8, N = 512, N_r = 2, N_t = 16$, CSI quantized to 6 bits. We observe a monotonic increase in the sum rate with pilot and data SNR. The sum rate performance for the values of pilot SNRs 10 dB and 20 dB are almost the same till the data SNR reaches 10 dB beyond which they diverge.

affect the achievable rate. At higher SNRs, the multiuser interference terms dominate the noise terms, and the IMM algorithm performs better than the other approaches as it is able to mitigate the multiuser interference through the joint selection of BF vectors. Also, User-BFVec selection approach performs worse than the IMM algorithm because of its inability to suppress the multiuser interference. Note that the feedback associated with this approach is the same as that of the IMM algorithm.

Figure 5.8 shows the sum rate performance vs. the codebook size (in bits), with $K = 10, N_r = 2$ and $N_t = 16$. The sum rate increases linearly with the codebook size,

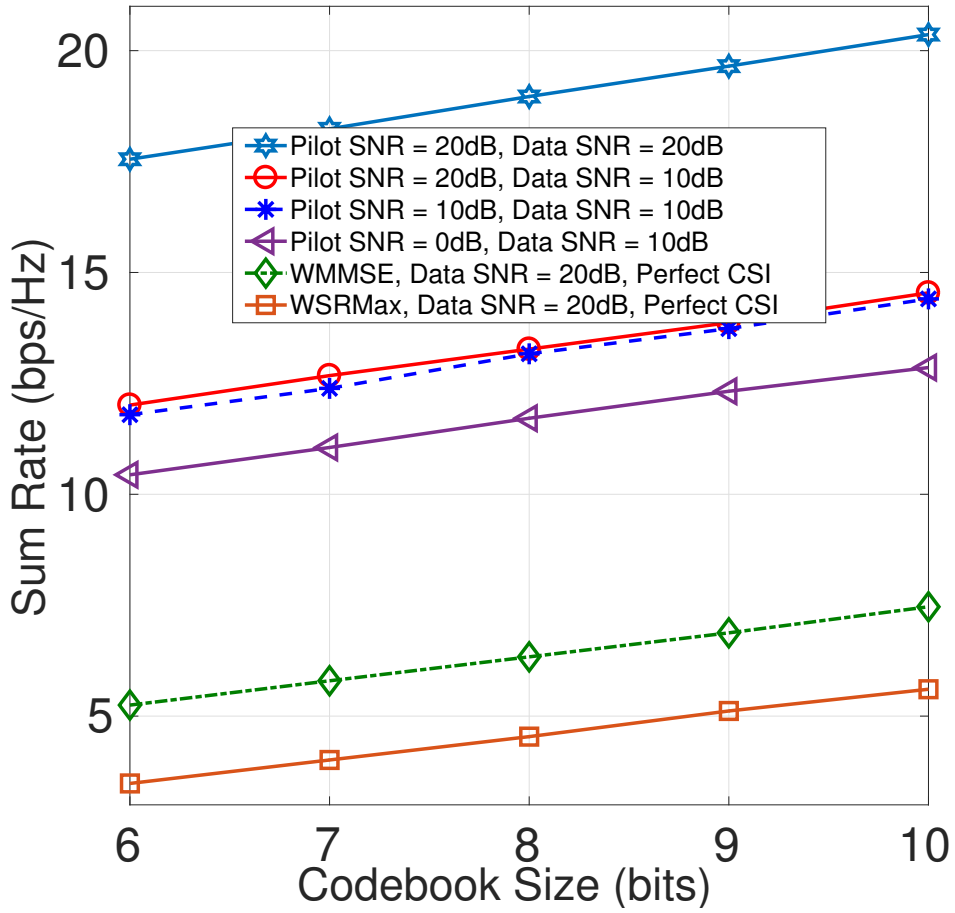


Figure 5.8: Sum rate vs. codebook size for $K = 10, N_r = 2, N_t = 16$, CSI quantized to 6 bits. We observe a monotonic increase in the sum rate with pilot and data SNR. As the codebook size increases, the AP has more number of beamforming vectors to choose resulting in an increase in the sum rate performance.

as the AP has more choices to select the beamforming vectors, which helps in canceling the multiuser interference. Note that, although the complexity of the problem increases with the codebook size, if the computational resources are limited, the IMM algorithm can be stopped at any iteration, resulting in a correspondingly effective solution. The figure also compares the the SMM and IMM algorithms with the WMMSE and WSRMax procedures. The sum rates achieved by WMMSE and WSRMax when the precoding vectors are quantized to the nearest vectors in the codebook are far inferior compared to that achieved by our algorithms. Once again, this illustrates the importance of accounting

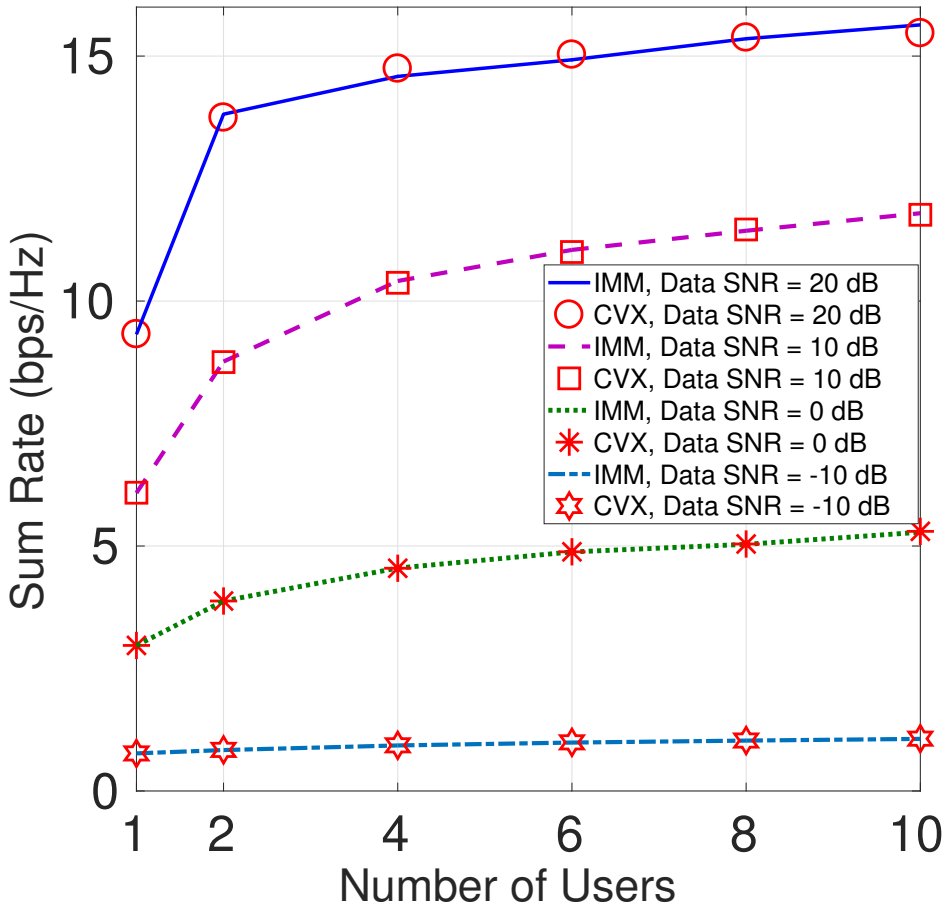


Figure 5.9: Sum Rate vs. $K$ with $N_r = 1$, $N_t = 32$, Pilot SNR=20 dB, $N = 1024$ for SMM, IMM and CVX. We see that the MM based algorithms and CVX converge to almost same sum rate for various values of data SNR.

for the codebook constraint while solving the sum rate optimization problem.

In Figure 5.9, we compare the sum rates achieved by directly solving (5.13) using CVX with the MM based algorithms. We see that both solutions converge to almost the same sum rates. However, the IMM algorithm exhibits faster run time, as shown in Figures 5.10 and 5.11. In these figures, we plot the ratio of the run times of the IMM and the convex solver CVX for various values of data SNRs and number of transmit antennas and number of users. The pilot SNR and the codebook size are set to 20 dB and $\{512, 1024\}$, respectively. We observe that as the data SNR increases, IMM converges much faster than the CVX package. This is thanks to the closed form expressions obtained for the

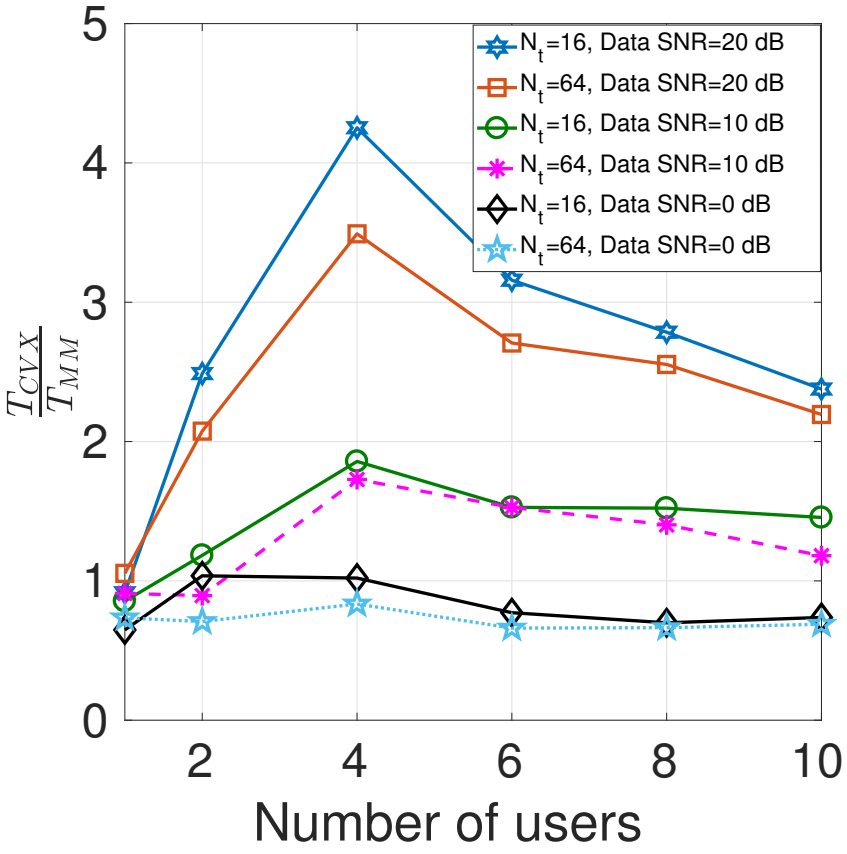


Figure 5.10: Ratio of run times of CVX and IMM with respect to $K$. Pilot SNR = 20 dB, $N_r = 1, N = 1024$. We see that IMM converges much faster than CVX in the interference limited regime.

final optimization problem. Also, as the number of transmit antennas increases, the ratio of the run times of CVX and IMM increases, which shows that IMM converges faster for large MIMO systems.

## 5.5 Conclusions

In this chapter, we presented two procedures, named square root MM (SMM) and inverse MM (IMM), to solve the problem of codebook based DL sum rate maximization in a MU-MIMO broadcast system. Both SMM and IMM procedures find a locally optimal allocation of beamforming vectors and data signal powers to each user, so as to maximize

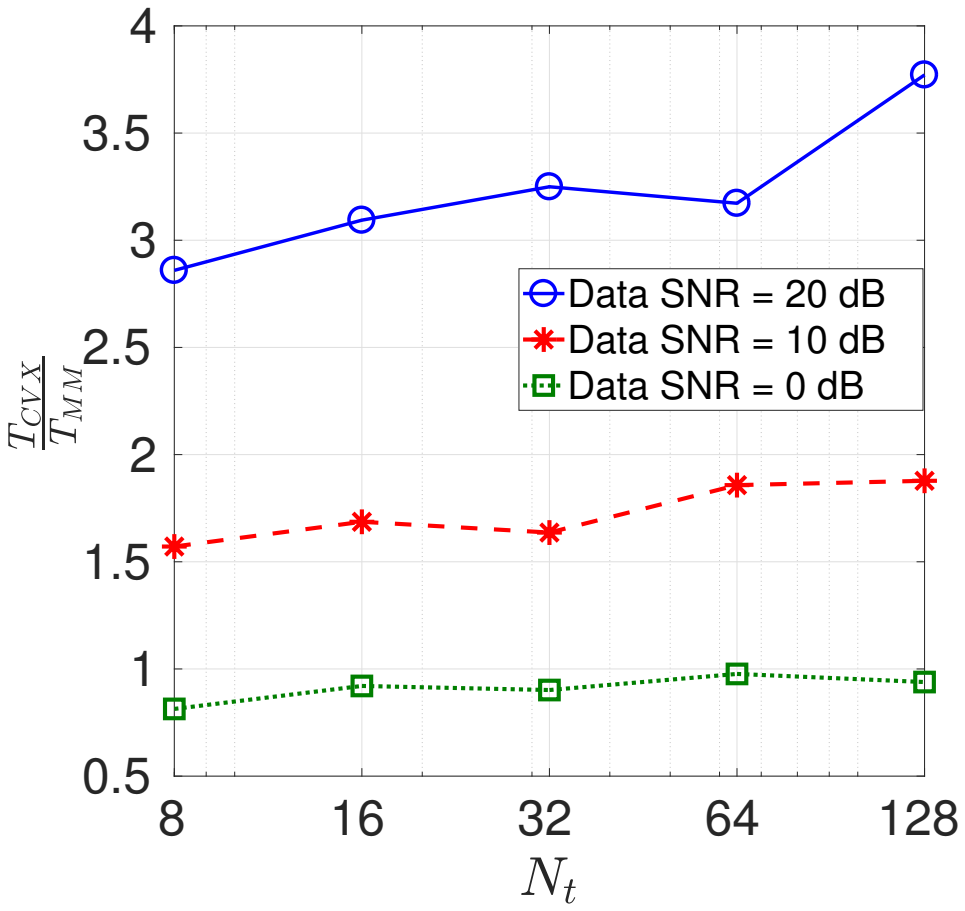


Figure 5.11: Ratio of run times of CVX and IMM with respect to $N_t$. Pilot SNR = 20 dB, $N_r = 1, N = 512, K = 8$.

the DL sum rate. These procedures are based on a nested application of the MM principle to lower bound the objective function, which is then maximized in an iterative fashion. The novelty of the algorithms lies in the choice of the surrogate functions used to bound the objective function. We also proved the optimality of the solutions to the surrogate optimization problems, as a consequence of which, the SMM and IMM algorithms reach a local optimum of the overall sum rate from any initialization.

We empirically illustrated the dependence of the achieved sum rate on the number of antennas, users, pilot power, and size of the codebook. We compared the sum rate of the SMM and IMM algorithms with that of WMMSE and WSRMax algorithms from the existing literature. The comparisons illustrated the importance of accounting for the

codebook constraint while solving the sum rate optimization problem. Future work could include a theoretical analysis of the dependence of the optimal sum rate on parameters such as the codebook size, pilot/data SNR, and number of antennas.

## 5.6 Derivations and Proofs

### 5.6.1 Derivation of Equation (5.8)

We simplify the covariance matrix of interest due to channel estimation error as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\widetilde{\mathbf{H}}_k \mathbf{x}\mathbf{x}^H \widetilde{\mathbf{H}}_k^H\right] \\
&= \mathbb{E}\left[\widetilde{\mathbf{H}}_k \mathbf{C} \sum_{j=1}^{K} \mathbf{\Phi}_j \mathbf{C}^H \widetilde{\mathbf{H}}_k^H\right] \\
&= \mathbb{E}\left[\widetilde{\mathbf{H}}_k \mathbf{C} \mathbf{\Phi}^{\frac{1}{2}} \mathbf{\Phi}^{\frac{1}{2}} \mathbf{C}^H \widetilde{\mathbf{H}}_k^H\right] \\
&= \mathbb{E}\left(\begin{bmatrix} \widetilde{\mathbf{h}_k^{(1)}}^T \\ \vdots \\ \widetilde{\mathbf{h}_k^{(N_r)}}^T \end{bmatrix} \left[\sqrt{P^{(1)}}\mathbf{c}_1, \ldots, \sqrt{P^{(N)}}\mathbf{c}_N\right] \times \begin{bmatrix} \sqrt{P^{(1)}}\mathbf{c}_1^H \\ \vdots \\ \sqrt{P^{(N)}}\mathbf{c}_N^H \end{bmatrix} \left[\widetilde{\mathbf{h}_k^{(1)}}^*, \ldots, \widetilde{\mathbf{h}_k^{(N_r)}}^*\right]\right) \\
&= \mathbb{E}\left(\begin{bmatrix} \sqrt{P^{(1)}}\mathbf{c}_1^T \widetilde{\mathbf{h}_k^{(1)}} & \ldots & \sqrt{P^{(N)}}\mathbf{c}_N^T \widetilde{\mathbf{h}_k^{(1)}} \\ \vdots & \vdots & \vdots \\ \sqrt{P^{(1)}}\mathbf{c}_1^T \widetilde{\mathbf{h}_k^{(N_r)}} & \ldots & \sqrt{P^{(N)}}\mathbf{c}_N^T \widetilde{\mathbf{h}_k^{(N_r)}} \end{bmatrix} \times \begin{bmatrix} \sqrt{P^{(1)}}\widetilde{\mathbf{h}_k^{(1)}}^H \mathbf{c}_1^* & \ldots & \sqrt{P^{(1)}}\widetilde{\mathbf{h}_k^{(N_r)}}^H \mathbf{c}_1^* \\ \vdots & \vdots & \vdots \\ \sqrt{P^{(N)}}\widetilde{\mathbf{h}_k^{(1)}}^H \mathbf{c}_N^* & \ldots & \sqrt{P^{(N)}}\widetilde{\mathbf{h}_k^{(N_r)}}^H \mathbf{c}_N^* \end{bmatrix}\right) \\
&= (\beta_k - \gamma_k) \sum_{n=1}^{N} P^{(n)} \mathbf{I}_{N_r} \\
&= (\beta_k - \gamma_k) \operatorname{tr}\left(\sum_{j=1}^{K} \mathbf{\Phi}_j\right) \mathbf{I}_{N_r} = (\beta_k - \gamma_k)\, \mathbf{I}_{N_r}. \qquad (5.40)
\end{aligned}
$$

In the above derivation, $P^{(n)} = \sum_{j=1}^{N} P_j(n)$, i.e., the sum of powers allocated to all the users in the $n^{\text{th}}$ beamforming vector. Also, $\widetilde{\mathbf{h}_k^{(l)}}$ denotes the $l^{\text{th}}$ row of the matrix $\widetilde{\mathbf{H}}_k$.

### 5.6.2 Proof of Lemma 9

Using the Woodbury identity for the inverse term in the function $f$ in (5.19), we get

$$\begin{aligned}
&-\mathrm{tr}\left(\mathbf{A}\left(\mathbf{B}^{-1}-\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\left(\mathbf{I}+\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\right)^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\right)\mathbf{A}^{H}\right)\\
&\quad=-\mathrm{tr}\left(\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^{H}\right)+\mathrm{tr}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}\right), \qquad (5.41)
\end{aligned}$$

where $\mathbf{X}\triangleq\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}$ and $\mathbf{Y}\triangleq\mathbf{I}+\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}$. The function $\mathrm{tr}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}\right)$ is jointly convex in $\mathbf{X}$ and $\mathbf{Y}$, and, can be minorized using a first order Taylor series. The complex matrix differential of $\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}$ is computed as follows [175]:

$$\mathrm{tr}\left(\mathrm{d}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}\right)\right)=\mathrm{tr}\left(\mathbf{Y}^{-1}\mathbf{X}^{H}\mathrm{d}\mathbf{X}-\mathbf{Y}^{-1}\mathbf{X}^{H}\mathbf{X}\mathbf{Y}^{-1}\mathrm{d}\mathbf{Y}+\mathbf{X}\mathbf{Y}^{-1}\mathrm{d}\mathbf{X}^{H}\right).$$

Thus, around the point $(\hat{\mathbf{X}},\hat{\mathbf{Y}})$, (5.41) can be lower bounded as

$$\begin{aligned}
&-\mathrm{tr}(\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^{H})+\mathrm{tr}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}\right)\\
&\quad\geq-\mathrm{tr}(\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^{H})+\mathrm{tr}\Big(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^{H}\left(\mathbf{X}-\hat{\mathbf{X}}\right)\\
&\qquad-\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^{H}\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\left(\mathbf{Y}-\hat{\mathbf{Y}}\right)+\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\left(\mathbf{X}-\hat{\mathbf{X}}\right)^{H}\Big)\\
&\quad=-\mathrm{tr}(\hat{\mathbf{K}})+\mathrm{tr}\Big(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^{H}\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}-\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^{H}\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\\
&\qquad+\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\left(\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\right)^{H}\Big)\\
&\quad=-\mathrm{tr}(\hat{\mathbf{K}})+\mathrm{tr}\Big(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^{H}\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}+\mathbf{C}^{H}\mathbf{B}^{-H}\mathbf{A}^{H}\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}}\\
&\qquad-\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^{H}\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\Big),
\end{aligned}$$

where $\hat{\mathbf{K}}$ is as defined in Lemma 9. Finally, grouping the constant matrices together, we get (5.20).

### 5.6.3 Proof of Lemma 10

We lower bound the function $f(\mathbf{R})$ using $\lambda$, the largest eigenvalue of the matrix $\mathbf{B}$, as follows:

$$\begin{aligned} f(\mathbf{R}) &= -\mathrm{tr}\left(\mathbf{AR}(\mathbf{B}-\lambda\mathbf{I})\mathbf{R}+\lambda\mathbf{AR}^2\right) \\ &= -\mathrm{tr}\left(\mathbf{ARCR}+\lambda\mathbf{AR}^2\right), \end{aligned} \tag{5.42}$$

where $\mathbf{C} \triangleq (\mathbf{B}-\lambda\mathbf{I})$. The complex matrix differential of the first term in (5.42) is [175]

$$\mathrm{tr}\left(\mathrm{d}(\mathbf{ARCR})\right) = \mathrm{tr}\left(\mathbf{CRA}(\mathrm{d}\mathbf{R}) + \mathbf{ARC}(\mathrm{d}\mathbf{R})\right).$$

Hence, around the previous iterate $\mathbf{R}^{(m)}$, a lower bound on $f$ can be written as

$$f(\mathbf{R}) \geq -\mathrm{tr}\left(\mathbf{AR}^{(m)}\mathbf{BR}^{(m)}\right) - \mathrm{tr}\left(\left(\mathbf{CR}^{(m)}\mathbf{A}+\mathbf{AR}^{(m)}\mathbf{C}\right)\left(\mathbf{R}-\mathbf{R}^{(m)}\right)\right) - \lambda\mathrm{tr}\left(\mathbf{AR}^2\right). \tag{5.43}$$

Grouping the constant terms in (5.43) together and substituting for $\mathbf{C}$, we get (5.24).

### 5.6.4 Proof of Lemma 11

Note that (5.28) is a separable convex optimization problem, which can be solved in closed form using the Lagrangian method. The Lagrangian for (5.28) is given by

$$\sum_{i=1}^{KN}\left([\mathbf{Q}^{(m)}]_{(i,i)}P(i) + [\mathbf{W}_A^{(m)}]_{(i,i)}P(i)^{\frac{1}{2}} + [\mathbf{W}_B^{(m)}]_{(i,i)}P(i)\right) + \eta\left(\sum_{i=1}^{KN}P(i)-1\right), \tag{5.44}$$

where $P(i), i = 1, 2, \ldots, KN$ denote the diagonal entries of $\mathbf{\Phi}$. Note that the sign of the objective function in (5.28) has been flipped while forming the Lagrangian, which makes it a minimization problem. By straightforward differentiation with respect to $P(i)$ in

(5.44), we obtain the closed form solution

$$P(i) = \left( \frac{\left[\mathbf{W}_A^{(m)}\right]_{(i,i)}}{2\left(\left[\mathbf{W}_B^{(m)}\right]_{(i,i)} + [\mathbf{Q}^{(m)}]_{(i,i)} + \eta\right)} \right)^2, \tag{5.45}$$

where $\eta$ is chosen to satisfy $\sum_{i=1}^{KN} P(i) = 1$. Since the $P(i)$ in (5.45) is strictly decreasing in $\eta$, it can be found using a simple line search or the bisection method [174].

Now, we show that the solution in (5.45) satisfies the second order sufficiency condition for optimality. That is, we show that the Hessian matrix of the Lagrangian in (5.44) is positive definite. The Hessian matrix of the Lagrangian in (5.44) is diagonal, with the $i^{\text{th}}$ diagonal entry $-[\mathbf{W}_A^{(m)}]_{(i,i)}/(4P^{3/2}(i))$, where $P(i)$ is given by (5.45). Thus, we need to show that the diagonal entries of $\mathbf{W}_A$ are strictly negative.

For simplicity and without loss of generality, we assume that $\sigma^2 = 1$. First, we simplify the first term, $\mathbf{W}_{1,k}$, of $\mathbf{W}_A$ in (5.26). Note that, $\mathbf{W}_{1,k}$, defined in (5.23), is a negative semidefinite matrix, as it is the sum of the matrix $\left(-\mathbf{Y}_k^{-1}\mathbf{X}_k^H\mathbf{F}_k\mathbf{\Psi}_k\right)$ and its conjugate transpose. The first term $\mathbf{Y}_k^{-1}\mathbf{X}_k^H\mathbf{F}_k\mathbf{\Psi}_k$ becomes

$$\begin{aligned}\left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{\Psi}^H\mathbf{F}_k^H\mathbf{F}_k\mathbf{\Psi} &= \left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{\Psi}^H\left(\mathbf{I}_{N_r} + \mathbf{\Psi}_k\mathbf{\Phi}\mathbf{\Psi}_k^H\right)\mathbf{\Psi}_k \\ &= \left(\mathbf{\Phi}^{-\frac{1}{2}} + \mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}\left(\mathbf{I}_{KN} + \mathbf{S}_k\mathbf{\Phi}\right)\mathbf{S}_k \\ &= \mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k.\end{aligned}$$

Thus, $\mathbf{W}_{1,k}$ can be written as

$$\mathbf{W}_{1,k} = -\left(\mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k + \mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right). \tag{5.46}$$

Define $\hat{\lambda}_k \triangleq \lambda_{\max}(\mathbf{S}_k)$. The term $\left(\mathbf{S}_k - \hat{\lambda}_k\mathbf{I}_{KN}\right)\mathbf{\Phi}^{\frac{1}{2}}\mathbf{W}_{2,k}$ in the expression for $\mathbf{W}_A$ in (5.26) becomes

$$\left(\mathbf{S}_k - \hat{\lambda}_k\mathbf{I}_{KN}\right)\mathbf{\Phi}^{\frac{1}{2}}\left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{\Psi}_k^H$$

$$\times \left(\mathbf{I}_{N_r} + \mathbf{\Psi}_k \mathbf{\Phi} \mathbf{\Psi}_k^H\right) \mathbf{\Psi}_k \mathbf{\Phi}^{\frac{1}{2}} \left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}$$
$$= \left(\mathbf{S}_k - \hat{\lambda}_k \mathbf{I}_{KN}\right) \mathbf{\Phi} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} \left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}. \quad (5.47)$$

Combining (5.46) and (5.47), we get the term inside the summation in the expression for $\mathbf{W}_A$ in (5.26) as follows:

$$-\mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k - \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} + \left(\mathbf{S}_k - \hat{\lambda}_k \mathbf{I}_{KN}\right) \mathbf{\Phi} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} \left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}$$
$$= -\left(\mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k + \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} + \hat{\lambda}_k \mathbf{\Phi} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right) \left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}$$
$$= -\mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k - \left(\mathbf{I}_{KN} + \hat{\lambda}_k \mathbf{\Phi}\right) \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} \left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}. \quad (5.48)$$

Now, we simplify the term $\left(\mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}} \left(\mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right)^{-1}\right)$ in (5.48). This becomes

$$\mathbf{S}_k \left(\mathbf{\Phi}^{-\frac{1}{2}} + \mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k\right)^{-1} = \mathbf{S}_k \left(\mathbf{\Phi}^{\frac{1}{2}} - \mathbf{\Phi} \left(\mathbf{I}_{KN} + \mathbf{S}_k \mathbf{\Phi}\right)^{-1} \mathbf{S}_k \mathbf{\Phi}^{\frac{1}{2}}\right) \quad (5.49)$$
$$= \left(\mathbf{S}_k - \mathbf{S}_k \left(\mathbf{\Phi}^{-1} + \mathbf{S}_k\right)^{-1} \mathbf{S}_k\right) \mathbf{\Phi}^{\frac{1}{2}}. \quad (5.50)$$

Here, the right hand side of (5.49) is obtained by applying the Woodbury matrix identity to the inverse term in the left hand side. Substituting (5.50) in (5.48), we get

$$\mathbf{W}_A = -\sum_{k=1}^{K} \left\{\mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k + \left(\mathbf{I}_{KN} + \lambda_{\max}\left(\mathbf{S}_k\right) \mathbf{\Phi}\right) \mathbf{W}_{A1,k} \mathbf{\Phi}^{\frac{1}{2}}\right\},$$

where $\mathbf{W}_{A1,k} \triangleq \mathbf{S}_k - \mathbf{S}_k \left(\mathbf{\Phi}^{-1} + \mathbf{S}_k\right)^{-1} \mathbf{S}_k$. Since $\mathbf{S}_k$ is symmetric and p.s.d., the diagonal entries of $\mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k$ are non-negative. Moreover, since $\mathbf{H}_k$ is drawn from a continuous valued distribution, the diagonal entries of $\mathbf{\Phi}^{\frac{1}{2}} \mathbf{S}_k$ are strictly positive with probability 1. Also, $\left(\mathbf{I}_{KN} + \lambda_{\max}\left(\mathbf{S}_k\right) \mathbf{\Phi}\right)$, and $\mathbf{\Phi}^{\frac{1}{2}}$ are diagonal matrices with strictly positive entries on the diagonal. Finally, it is easy to show that the diagonal entries of $\mathbf{W}_{A1,k}$ are also non-negative. For this, it suffices to show that the diagonal entries of $\left\{\mathbf{S}_k - \mathbf{S}_k \left(\frac{1}{\lambda_{\max}(\mathbf{\Phi})} \mathbf{I}_{KN} + \mathbf{S}_k\right)^{-1} \mathbf{S}_k\right\}$ are non-negative, where $\lambda_{\max}(\mathbf{\Phi})$ is the largest eigenvalue of the matrix $\mathbf{\Phi}$. The eigenvalues of this matrix are given by

$\lambda(\mathbf{S}_k)/\left(1+\lambda_{\max}\left(\mathbf{\Phi}\right)\lambda\left(\mathbf{S}_k\right)\right)$, where $\lambda(\mathbf{S}_k)$ is an eigenvalue of $\mathbf{S}_k$. These eigenvalues are non-negative, and hence the eigenvalues of $\mathbf{W}_{A1,k}$ are also non-negative. Moreover, since $\mathbf{H}_k$ is drawn from a continuous valued distribution, the diagonal entries of $\mathbf{W}_{A1,k}$ are strictly positive with probability 1. Therefore, the diagonal entries of $\mathbf{W}_A$ are strictly negative, thus satisfying the second order sufficient conditions for optimality of the surrogate optimization problem in (5.28).

### 5.6.5 Proof of Lemma 12

The objective function is quadratic in $\mathbf{\Phi}$, and is therefore amenable to optimization. Further, the problem in (5.37) is separable in the optimization variables, and can be solved using the Lagrangian method to obtain a closed form solution. The Lagrangian is given by

$$\sum_{i=1}^{KN}\left(\left[\mathbf{Q}^{(m)}\right]_{(i,i)}P(i)+\left[\mathbf{Z}^{(m)}\right]_{(i,i)}\frac{1}{P(i)}\right)+\eta\left(\sum_{i=1}^{KN}P(i)-1\right). \tag{5.51}$$

Note that, the negative sign in the optimization problem in (5.37) is removed while forming the Lagrangian, making it a minimization problem. By differentiating (5.51) with respect to $P(i)$, we obtain (5.38).

Now, we show that the solution to the surrogate convex optimization problem in (5.38) satisfies the second order sufficiency condition for optimality. The Hessian matrix of the Lagrangian in (5.51) is diagonal, with the $i^{\text{th}}$ diagonal entry $\{2[\mathbf{Z}^{(m)}]_{(i,i)}/P(i)^3\}$, where $P(i)$ is given by (5.38). Thus, we need to show that the diagonal entries of $\mathbf{Z}$ are strictly positive.

Without loss of generality, we assume that $\sigma^2=1$. We can simplify each term inside the summation in (5.36) as follows:

$$\widetilde{\mathbf{\Phi}}\left(\widetilde{\mathbf{S}}_k-\widetilde{\mathbf{S}}_k\left(\widetilde{\mathbf{\Phi}}^{-1}+\widetilde{\mathbf{S}}_k\right)^{-1}\widetilde{\mathbf{S}}_k\right)\widetilde{\mathbf{\Phi}},$$

where $\widetilde{\mathbf{S}}_k \triangleq \widetilde{\mathbf{\Psi}}_k^H\widetilde{\mathbf{\Psi}}_k$. By following a similar procedure as done for showing the optimality

of the SMM solution in (5.45), we can show that the diagonal entries of $\mathbf{Z}$ are strictly positive with probability 1, and hence it satisfies the second order sufficient conditions for optimality of the solution (5.38) to the surrogate optimization problem (5.37).

## 5.7 SMM and IMM Pseudocodes

The pseudo codes for the SMM and IMM algorithms are provided in Algorithms 9 and 10, respectively.

**Algorithm 9:** SMM

**Input:** $\widehat{\mathbf{H}}_1, \ldots, \widehat{\mathbf{H}}_K, \mathbf{C}, K, \rho_{dl}$

**Output:** $P_1(1), \ldots, P_1(N), \ldots, P_K(1), \ldots, P_K(N)$

Initialize $P_1(1), \ldots, P_1(N), \ldots, P_K(1), \ldots, P_K(N)$ to satisfy the total power constraint.
Compute $\widehat{\widehat{\mathbf{H}}}_k = \sqrt{\rho_{dl}}\widehat{\mathbf{H}}_k\mathbf{C}$, $k = 1, 2, \ldots, K$.
Compute $\mathbf{\Psi}_1, \ldots, \mathbf{\Psi}_K$ using (5.15).
**repeat**
    Compute $\mathbf{\Phi}$ using (5.14).
    Compute $\mathbf{Q}$ using (5.16).
    Compute $\mathbf{W}_A$, $\mathbf{W}_B$ using (5.26), (5.27).
    Calculate Lagrange multiplier $\eta$ using line search to satisfy maximum power constraint.
    Compute $P(i)$ using (5.29), $i = 1, 2, \ldots, KN$.
    **for** $k = 1$ to $K$ **do**
        **for** $i = 1$ to $N$ **do**
            Compute $P_k(i)$ using (5.39).
        **end for**
    **end for**
**until** convergence

**Algorithm 10:** IMM

**Input:** $\widehat{\mathbf{H}}_1, \ldots, \widehat{\mathbf{H}}_K, \mathbf{C}, K, \rho_{dl}$

**Output:** $P_1(1), \ldots, P_1(N), \ldots, P_K(1), \ldots, P_K(N)$

1: Initialize $P_1(1), \ldots, P_1(N), \ldots, P_K(1), \ldots, P_K(N)$ to satisfy the total power constraint.
2: Compute $\widehat{\widehat{\mathbf{H}}}_k = \sqrt{\rho_{dl}}\widehat{\mathbf{H}}_k\mathbf{C}$, $k = 1, 2, \ldots, K$.
3: Compute $\widetilde{\mathbf{\Psi}}_1, \ldots, \widetilde{\mathbf{\Psi}}_K$ using (5.31).
4: **repeat**
5: Compute $\mathbf{\Phi}, \widetilde{\mathbf{\Phi}}$ using (5.14), (5.30), respectively.
6: Compute $\mathbf{Q}$ and $\mathbf{Z}$ using (5.16) and (5.36) respectively.
7: Calculate Lagrange multiplier $\eta$ using line search to satisfy maximum power constraint.
8: Compute $P(i)$ using (5.38), $i = 1, 2, \ldots, KN$.
9: **for** $k = 1$ to $K$ **do**
10: **for** $i = 1$ to $N$ **do**
11: Compute $P_k(i)$ using (5.39).
12: **end for**
13: **end for**
14: **until** convergence

# Chapter 6

# Conclusions and Future Work

In this thesis, we investigated the application of VB techniques in the channel estimation and data decoding problems in massive MIMO-SC and MIMO-OFDM wireless communication systems with low-resolution ADCs. Further, we derived different Cramér-Rao type lower bounds to benchmark the performance of estimators of jointly compressible signals with quantized lower dimensional measurements. We also studied the model and data driven mmWave sparse channel estimation using greedy and Bayesian approaches. Finally, we used the minorization-maximization framework to maximize the downlink sum rate in a codebook based multi-user MIMO broadcast system. We summarize the main contributions of this thesis below.

## 6.1 Summary of the Thesis

In chapter 2, we developed channel estimation and data detection algorithms for uplink massive MIMO-SC communication systems with low-resolution ADCs. We formulated the channel estimation and data detection as a statistical inference problem to derive the posterior distributions of the channel and the data. The computational intractability involved in calculating the partition function in a high dimensional statistical inference problem necessitated alternate techniques to compute the posterior distributions of the latent variables and/or parameters. We used an approximate inference framework called

VB, which is a principled approach to obtain the posterior distributions of the latent variables. We judiciously introduced latent variables, which led to a computationally and analytically tractable solution. We imposed a factorized structure on the posterior distribution and developed a low complexity channel estimation and soft symbol decoding algorithm. We benchmarked the NMSE and the coded BER performance of the VB receiver with the state-of-the-art. We integrated the VB based joint channel estimator and soft symbol detector with the channel decoder, and showed that it is possible to utilize the extrinsic information to further improve the system performance.

In chapter 3, we derived the Bayesian CRLB on the MSE of an estimator of jointly compressible signals using quantized noisy measurements. We imposed a two-stage hierarchical complex normal prior on the compressible signal, with a conjugate Gamma hyperprior on the precision matrix. Due to the analytically intractability in obtaining a closed form expression for the Bayesian information matrix, we computed it via simulations. Then, we extended the joint channel estimation and soft symbol problems to a massive MIMO-OFDM system with low-resolution ADCs. We exploited the delay domain sparse nature of the wireless channels to formulate the channel estimation as a sparse signal recovery problem. We imposed a two stage hierarchical complex normal prior on the sparse channel, with a conjugate Gamma hyperprior on the precision matrix. We then used a factorized posterior structure on the users' channels and the unquantized received signals to obtain a VB based sparse channel estimation algorithm. Next, we utilized the estimated channels to compute the posterior distributions of the data symbols based on the VB framework. Finally, we computed the posterior means of the detected data symbols and use them as virtual pilots to re-estimate the channels in an iterative fashion. We integrated the iterative VB channel estimator and data decoder with a channel decoder to further improve the performance. We also discussed the per-iteration computational complexities of the iterative sparse channel estimation and soft symbol decoding algorithm. Finally, we benchmarked the Bit Error Rate (BER) and NMSE performances of our algorithm with the state-of-the-art.

In chapter 4, we considered a hybrid mmWave MIMO-OFDM system deployed in

a specific site. We utilized the beamspace sparse nature of the mmWave channels to formulate the channel estimation as a joint row sparse recovery problem. We integrated a fast greedy algorithm for dimensionality reduction with a high performing Bayesian learning procedure to estimate the mmWave channel. Then, we developed a multi-level greedy search algorithm to learn the sparsifying beamspace dictionary. We utilized the learnt dictionary to estimate the channel using a sparse Bayesian learning algorithm. We exploited the training dataset to learn the number of multipaths with significant energy, and used them for denoising the estimated channel in the frequency domain. Finally, we utilized the lag domain sparsity information of the channels from the dataset, to denoise the channels in the lag domain. We benchmarked the developed algorithm with the state-of-the-art.

In chapter 5, we considered a single cell downlink multiuser MIMO system where the beamforming vectors are chosen from a predefined codebook. We derived the downlink achievable sum-rate of this codebook based multiuser MIMO system when the channels are estimated using MMSE approach. We then solved the non-convex sum-rate maximization problem with a total transmit power constraint. We adopted the MM framework to bound the objective function, and maximize it in an iterative manner. We utilized the mathematical tools from complex matrix derivatives and inequalities, to lower bound the objective function multiple times, to obtain a tractable surrogate convex optimization problem. We also analytically proved that the SMM and IMM algorithms converged to a global optimum of the surrogate optimization problems. We also discussed the computational complexities of the algorithms, and shown empirically that a judicious choice of bounding the non-convex objective function is essential for faster convergence of the solution. Finally, we compared the sum-rate performance of the MM algorithms with the state-of-the-art.

## 6.2 Future Work

We list a few interesting research directions which can be pursued in future.

1. In the massive MIMO-OFDM system, we exploited the joint row sparse structure of the lag domain channel to design a VB based channel estimation algorithm. Along with the row sparsity, each column of the estimand also has a piecewise sparse structure, which can be potentially utilized to improve the system performance. By piecewise sparsity, we mean that each sub-block of every column has a joint row sparse structure. Utilizing the piecewise sparsity may assist in reducing the computational complexity and also improve the system performance.

2. The VB channel estimation algorithms for both the massive MIMO single carrier and OFDM systems require the computation of matrix inverses to obtain the covariance matrices of the posterior distributions of the users' channels. This results in a cubic complexity of the algorithms. To overcome this, we can replace the channel matrix computation part by a message passing based algorithm, which results in a linear complexity in terms of the number of users and the channel lengths. Exploring this option will be an interesting extension to the channel estimation problem.

3. One can employ data-driven approaches to implement the VB based channel estimation and soft symbol decoding algorithms. The posterior distributions of the data symbols in the VB algorithm follow a Boltzmann distribution, which is the same as the softmax operation that is used as activation functions in deep neural networks (DNN). Therefore, unfolding the VB algorithm as a DNN is an interesting research direction to explore in the future.

# Bibliography


[1] N. K. Chavali and A. R. Yalla, "A soft-demapper for coded MIMO-OFDM system," in *Proc. Int. Conf. on Contemp. Comput. and Inform. (IC3I)*, 2014, pp. 451–457.

[2] E. G. Larsson, "Massive MIMO for 5G: Overview and the road ahead," in *Proc. Conf. on Inform. Sci. and Syst. (CISS)*, 2017, pp. 1–1.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[5] E. Björnson, J. Hoydis, and L. Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*. Now Foundations and Trends, 2017.

[6] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[7] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[8] Z. Jiayi, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Comm. Mag.*, vol. 56, no. 7, pp. 205–211, Jul. 2018.

[9] I. F. Akyildiz, A. Kak, and S. Nie, "6G and beyond: The future of wireless communications systems," *IEEE Access.*, vol. 8, pp. 133 995–134 030, Jul. 2020.

[10] X. Zhang and Q. Zhu, "NOMA and user-centric based cell-free massive MIMO over 6G big-data mobile wireless networks," in *Proc. Globecom*, 2020, pp. 1–6.

[11] M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, "The road to 6G: Ten physical layer challenges for communications engineers," *IEEE Comm. Mag.*, vol. 59, no. 1, pp. 64–69, Jan. 2021.

[12] B. Ji, Y. Han, S. Liu, F. Tao, G. Zhang, Z. Fu, and C. Li, "Several key technologies for 6G: Challenges and opportunities," *IEEE Comm. Std. Mag.*, vol. 5, no. 2, pp. 44–51, Jun. 2021.

[13] O. T. Demir, E. Björnson, and L. Sanguinetti, *Foundations of User-Centric Cell-Free Massive MIMO.* Now Foundations and Trends, 2021.

[14] H. L. Van Trees and K. L. Bell, *Detection, Estimation, and Modulation Theory.* Wiley, 2013.

[15] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory.* Prentice Hall, New Jersey, 1993.

[16] J. J. Bussgang, *Crosscorrelation functions of amplitude-distorded Gaussian signals.* Research Laboratory of Electronics, Massachussets Institute of Technology, Tech. Rep. 216, 1952.

[17] O. T. Demir and E. Bjornson, "The Bussgang decomposition of nonlinear systems: Basic theory and MIMO extensions [lecture notes]," *IEEE Signal Process. Mag.*, vol. 38, no. 1, pp. 131–136, Jan. 2021.

[18] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *Proc. IEEE Int. Symp. Inf. Theory*, 2007, pp. 1286–1289.

[19] ——, "Analysis of Rayleigh-fading channels with 1-bit quantized output," in *Proc. IEEE Int. Symp. Inf. Theory*, 2008, pp. 260–264.

[20] J. Singh, O. Dabeer, and U. Madhow, "Capacity of the discrete-time AWGN channel under output quantization," in *Proc. IEEE Int. Symp. Inf. Theory*, 2008, pp. 1218–1222.

[21] ——, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.

[22] J. Mo and R. W. Heath, "High SNR capacity of millimeter wave MIMO systems with one-bit quantization," in *Proc. ITA*, 2014, pp. 1–5.

[23] ——, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[24] "Evolved universal terrestrial radio access (E-UTRA) 3GPP TS 36.201 V16.0.0 (2020-06); LTE physical layer; general description," 2020.

[25] "3GPP TS 38.201 V16.0.0 (2019-12); NR; Physical layer; General description," 2019.

[26] A. Maltsev *et al.*, "Channel models for 60 GHz WLAN systems," IEEE, Piscataway, New Jersey, Tech. Rep. 802.11-09/0334r8, 2010.

[27] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint approximately sparse channel estimation and data detection in OFDM systems using sparse Bayesian learning," *IEEE Trans. Signal Process.*, vol. 62, no. 14, pp. 3591–3603, Jul 2014.

[28] ——, "Joint channel estimation and data detection in MIMO-OFDM systems: A sparse Bayesian learning approach," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5369–5382, Oct. 2015.

[29] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[30] A. Zymnis, S. Boyd, and E. Candes, "Compressed sensing with quantized measurements," *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 149–152, Feb. 2010.

[31] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. New York, NY, USA: Birkhäuser, 2013.

[32] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.

[33] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, 2007.

[34] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[35] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[36] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC Press, 2015.

[37] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *Proc. IEEE Inf. Theory Workshop*, 2010, pp. 1–5.

[38] ——, "Message passing algorithms for compressed sensing: II. analysis and validation," in *Proc. IEEE Inf. Theory Workshop*, 2010, pp. 1–5.

[39] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *IEEE Comm. Surveys and Tuts.*, vol. 19, no. 3, pp. 1527–1550, 2017.

[40] S. Khanna and C. R. Murthy, "Communication-efficient decentralized sparse Bayesian learning of joint sparse signals," *IEEE Trans. Signal and Inf. Process. over Netw.*, vol. 3, no. 3, pp. 617–630, Sep. 2017.

[41] ——, "Sparse recovery from multiple measurement vectors using exponentiated gradient updates," *IEEE Signal Process. Lett.*, vol. 25, no. 10, pp. 1485–1489, Oct. 2018.

[42] J. T. Parker, P. Schniter, and V. Cevher, "Bilinear generalized approximate message passing—part i: Derivation," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5839–5853, Nov. 2014.

[43] ——, "Bilinear generalized approximate message passing—part ii: Applications," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5854–5867, Nov. 2014.

[44] M. Al-Shoukairi, P. Schniter, and B. D. Rao, "A GAMP-based low complexity sparse Bayesian learning algorithm," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 294–308, Jan. 2018.

[45] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[46] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[47] R. C. Daniels and R. W. Heath, "60 GHz wireless communications: Emerging requirements and design recommendations," *IEEE Veh. Tech. Mag.*, vol. 2, no. 3, pp. 41–50, Sep. 2007.

[48] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for

fusion of wireless communications and radar sensing," *Proc. IEEE*, vol. 99, no. 7, pp. 1236–1259, Jul. 2011.

[49] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. Murdock, *Millimeter Wave Wireless Communications*. Englewood Cliffs, NJ, USA: Prentice-Hall, Sep. 2014.

[50] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

[51] J. Mo, P. Schniter, and R. W. Heath, "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, Mar. 2018.

[52] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[53] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Comm. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.

[54] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[55] 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 Comm. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[56] O. E. 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.

[57] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, 2018.

[58] S. S. Thoota and C. R. Murthy, "Variational Bayesian inference based soft-symbol decoding for uplink massive MIMO systems with low resolution ADCs," in *Proc. Asilomar Conf. on Signals, Syst., and Comput.*, 2019.

[59] S. S. Thoota, C. R. Murthy, and R. Annavajjala, "Quantized variational Bayesian joint channel estimation and data detection for uplink massive MIMO systems with low resolution ADCs," in *Proc. MLSP*, 2019, pp. 1–6.

[60] S. S. Thoota and C. R. Murthy, "Variational Bayes' joint channel estimation and soft symbol decoding for uplink massive MIMO systems with low resolution ADCs," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3467–3481, May 2021.

[61] ——, "Cramér-Rao lower bound for Bayesian estimation of quantized MMV sparse signals," in *Proc. SPAWC*, 2021.

[62] ——, "Massive MIMO-OFDM systems with low resolution ADCs: Cramér-Rao bound, sparse channel estimation, and soft symbol decoding," *IEEE Trans. Signal Process.*, 2022.

[63] S. S. Thoota, D. G. Marti, O. T. Demir, R. Mundlamuri, J. Palacios, C. M. Yetis, C. K. Thomas, S. H. Bharadwaja, E. Björnson, P. Giselsson, M. Kountouris, C. R. Murthy, N. González-Prelcic, and J. Widmer, "Site-specific millimeter-wave compressive channel estimation algorithms with hybrid MIMO architectures," *ITU Journal on Future and Evolving Technologies (ITU J-FET) special issue on AI and machine learning solutions in 5G and future networks*, vol. 2, no. 4, Jul. 2021.

[64] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, no. 1, 2004.

[65] K. Lange, *MM Optimization Algorithms*. Society for Industrial and Applied Mathematics, 2016.

[66] S. S. Thoota, P. Babu, and C. R. Murthy, "Codebook based precoding for multiuser MIMO broadcast systems: An MM approach," in *Proc. Nat. Conf. Commun. (NCC)*, Feb. 2019, pp. 1–6.

[67] ——, "Codebook-based precoding and power allocation for MU-MIMO systems for sum rate maximization," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8290–8302, Dec. 2019.

[68] J. Liu, Z. Luo, and X. Xiong, "Low-resolution ADCs for wireless communication: A comprehensive survey," *IEEE Access*, vol. 7, pp. 91 291–91 324, Jul. 2019.

[69] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC," *ArXiv*, p. arXiv:1404.7736, Apr. 2014.

[70] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3111–3124, Apr. 2014.

[71] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. Wiley, 1995.

[72] A. P. Kannu and P. Schniter, "On communication over unknown sparse frequency-selective block-fading channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6619–6632, Oct. 2011.

[73] *Raymobtime Channel Generation Tool*, 2020. [Online]. Available: https://www.lasse.ufpa.br/raymobtime/

[74] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *Proc. Asilomar Conf. on Signals, Syst., and Comput.* IEEE, 2014, pp. 273–277.

[75] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[76] "IEEE Std 802.15.3c-2009, Part 15.3: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPANs) Amendment 2: Millimeterwave-based Alternative Physical Layer Extension," *IEEE Computer Society*, 2009.

[77] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[78] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[79] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[80] B. Murmann, "Energy limits in A/D converters," in *2013 IEEE Faible Tension Faible Consommation*. IEEE, 2013, pp. 1–4.

[81] *ADS54J20 Dual-Channel, 12-Bit, 1.0-GSPS, Analog-to-Digital Converter (Rev. B)*, Texas Instruments, 2011, rev. B. [Online]. Available: https://www.ti.com/document-viewer/ADS54J20/datasheet

[82] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[83] Y. Li, C. Tao, A. L. Swindlehurst, A. Mezghani, and L. Liu, "Downlink achievable rate analysis in massive MIMO systems with one-bit DACs," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1669–1672, Jul. 2017.

[84] A. Mezghani and A. L. Swindlehurst, "Blind estimation of sparse broadband massive MIMO channels with ideal and one-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2972–2983, Jun. 2018.

[85] H. Kim and J. Choi, "Channel estimation for spatially/temporally correlated massive MIMO systems with one-bit ADCs," *EURASIP J. on Wireless Commun. and Netw.*, vol. 2019, no. 1, p. 267, Dec. 2019.

[86] J. A. Nossek and M. T. Ivrlač, "Capacity and coding for quantized MIMO systems," in *Proc. Int. Conf. Wireless. Commun., and Mobile Comput. Conf. (IWCMC)*, 2006, pp. 1387–1392.

[87] S.-N. Hong, S. Kim, and N. Lee, "A weighted minimum distance decoding for uplink multiuser MIMO systems with low-resolution ADCs," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 1912–1924, May 2018.

[88] S.-N. Hong and N. Lee, "Soft-output detector for uplink MU-MIMO systems with one-bit ADCs," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 930–933, May 2018.

[89] Y.-S. Jeon, N. Lee, S.-N. Hong, and R. W. Heath, "One-bit sphere decoding for uplink massive MIMO systems with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4509–4521, Jul. 2018.

[90] S. Wang, Y. Li, and J. Wang, "Multiuser detection for uplink large-scale MIMO under one-bit quantization," in *Proc. ICC*, 2014, pp. 4460–4465.

[91] H. He, C.-K. Wen, and S. Jin, "Bayesian optimal data detector for hybrid mmwave MIMO-OFDM systems with low-resolution ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 469–483, Mar. 2018.

[92] H. Wang, W.-T. Shih, C.-K. Wen, and S. Jin, "Reliable OFDM receiver with ultra-low resolution ADC," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3566–3579, Jan. 2019.

[93] Z. Zhang, X. Cai, C. Li, C. Zhong, and H. Dai, "One-bit quantized massive MIMO detection based on variational approximate message passing," *IEEE Trans. Signal Process.*, vol. 66, no. 9, pp. 2358–2373, May 2018.

[94] J. Choi, J. Mo, and R. W. Heath, "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, May 2016.

[95] C.-K. Wen, C.-J. Wang, S. Jin, K.-K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, pp. 2541–2556, May 2016.

[96] P. Sun, Z. Wang, R. W. Heather, and P. Schniter, "Joint channel-estimation/decoding with frequency-selective channels and few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 899–914, Feb. 2019.

[97] Y. Cho and S.-N. Hong, "One-bit successive-cancellation soft-output (OSS) detector for uplink MU-MIMO systems with one-bit ADCs," *IEEE Access.*, vol. 7, pp. 27 172–27 182, Feb. 2019.

[98] Z. Shao, R. C. de Lamare, and L. T. N. Landau, "Iterative detection and decoding for large-scale multiple-antenna systems with 1-bit ADCs," *IEEE Wireless Comm. Letters*, vol. 7, no. 3, pp. 476–479, Jun. 2018.

[99] A. Mezghani and J. A. Nossek, "Belief propagation based MIMO detection operating on quantized channel output," in *Proc. IEEE Int. Symp. Inf. Theory*, 2010, pp. 2113–2117.

[100] M.-A. Badiu, G. E. Kirkelund, C. N. Manchón, E. Riegler, and B. H. Fleury, "Message-passing algorithms for channel estimation and decoding using approximate inference," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 2376–2380.

[101] G. E. Kirkelund, C. N. Manchón, L. P. Christensen, E. Riegler, and B. H. Fleury,

"Variational message-passing for joint channel estimation and decoding in MIMO-OFDM," in *Proc. Globecom*, 2010, pp. 1–6.

[102] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer New York, 2006.

[103] "3GPP TS 38.212 V15.7.0 (2019-09); NR; Multiplexing and channel coding," 2019.

[104] L. Schumacher, K. I. Pedersen, and P. E. Mogensen, "From antenna spacings to theoretical capacities-guidelines for simulating MIMO systems," in *Proc. PIMRC*, vol. 2, 2002, pp. 587–592.

[105] S. Wu, L. Kuang, Z. Ni, D. Huang, Q. Guo, and J. Lu, "Message-passing receiver for joint channel estimation and decoding in 3d massive MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8122–8138, Dec. 2016.

[106] D. Ying, F. W. Vook, T. A. Thomas, D. J. Love, and A. Ghosh, "Kronecker product correlation model and limited feedback codebook design in a 3D channel model," in *Proc. ICC*, Jun. 2014, pp. 5865–5870.

[107] S. Weber, J. G. Andrews, and N. Jindal, "The effect of fading, channel inversion, and threshold scheduling on ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4127–4149, Feb. 2007.

[108] P. Herath, C. Tellambura, and W. A. Krzymień, "Coverage probability analysis of three uplink power control schemes: Stochastic geometry approach," *EURASIP J. on Wireless Commun. and Netw.*, vol. 2018, no. 1, p. 141, 2018.

[109] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, Jun. 2016.

[110] H. Wang, C.-K. Wen, and S. Jin, "Bayesian optimal data detector for mmWave OFDM system with low-resolution ADC," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1962–1979, Sep. 2017.

[111] H. He, C.-K. Wen, and S. Jin, "Bayesian optimal data detector for hybrid mmwave MIMO-OFDM systems with low-resolution ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 469–483, Jun. 2018.

[112] Y.-S. Jeon, H. Do, S.-N. Hong, and N. Lee, "Soft-output detection methods for sparse millimeter-wave MIMO systems with low-precision ADCs," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 2822–2836, Apr. 2019.

[113] H. Wang, W. Shih, C. Wen, and S. Jin, "Reliable OFDM receiver with ultra-low resolution ADC," *IEEE Trans. Wireless Commun.*, vol. 67, no. 5, pp. 3566–3579, May 2019.

[114] Y. Ding, S.-E. Chiu, and B. D. Rao, "Bayesian channel estimation algorithms for massive MIMO systems with hybrid analog-digital processing and low-resolution ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 499–513, Jun. 2018.

[115] L. V. Nguyen, A. L. Swindlehurst, and D. H. N. Nguyen, "Svm-based channel estimation and data detection for one-bit massive mimo systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2086–2099, 2021.

[116] X. Cheng, B. Xia, K. Xu, and S. Li, "Bayesian channel estimation and data detection in oversampled OFDM receiver with low-resolution ADC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5558–5571, Sep. 2021.

[117] H. L. V. Trees, K. L. Bell, and Z. Tian, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory.* John Wiley and Sons, Inc., Hoboken, New Jersey, 2013.

[118] M. P. Shutzenberger, "A generalization of the Fréchet-Cramér inequality in the case of Bayes estimation," *Bull. Am. Math. Soc.*, vol. 63 (142), 1957.

[119] R. Prasad and C. R. Murthy, "Cramér-Rao-Type bounds for sparse Bayesian learning," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 622–632, Feb. 2013.

[120] J. García, J. Munir, K. Roth, and J. A. Nossek, "Channel estimation and data equalization in frequency-selective MIMO systems with one-bit quantization," *ArXiv*, p. arXiv:1609.04536, Mar. 2021.

[121] Z. Shao, L. T. N. Landau, and R. C. De Lamare, "Channel estimation for large-scale multiple-antenna systems using 1-bit ADCs and oversampling," *IEEE Access*, vol. 8, pp. 85 243–85 256, May 2020.

[122] R. Gribonval, V. Cevher, and M. E. Davies, "Compressible distributions for high-dimensional statistics," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5016–5034, Aug. 2012.

[123] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

[124] Y. Ding, S.-E. Chiu, and B. D. Rao, "Sparse recovery with quantized multiple measurement vectors," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2017, pp. 845–849.

[125] J. H. Dauwels, *On graphical models for communications and machine learning: Algorithms, bounds, and analog implementation*. ETH Zurich, 2006, vol. 17.

[126] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.

[127] E. Karami, "Tracking performance of least squares MIMO channel estimation algorithm," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2201–2209, 2007.

[128] Z. Marzi, D. Ramasamy, and U. Madhow, "Compressive channel estimation and tracking for large arrays in mm-wave picocells," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 514–527, 2016.

[129] S. Liu, F. Yang, W. Ding, X. Wang, and J. Song, "Two-dimensional structured-compressed-sensing-based NBI cancelation exploiting spatial and temporal correlations in MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 65, no. 11, pp. 9020–9028, 2016.

[130] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, 2017.

[131] E. Vlachos, G. C. Alexandropoulos, and J. Thompson, "Wideband MIMO channel estimation for hybrid beamforming millimeter wave systems via random spatial sampling," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1136–1150, 2019.

[132] Y. Wang, Y. Zhang, Z. Tian, G. Leus, and G. Zhang, "Super-resolution channel estimation for arbitrary arrays in hybrid millimeter-wave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 947–960, 2019.

[133] F. Dong, W. Wang, Z. Huang, and P. Huang, "High-resolution angle-of-arrival and channel estimation for mmWave massive MIMO systems with lens antenna array," *IEEE Trans. Veh. Tech.*, vol. 69, no. 11, pp. 12 963–12 973, 2020.

[134] H. Xie and N. González-Prelcic, "Dictionary learning for channel estimation in hybrid frequency-selective mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7407–7422, 2020.

[135] J. P. González-Coma, J. Rodríguez-Fernández, N. González-Prelcic, L. Castedo, and R. W. Heath, "Channel estimation and hybrid precoding for frequency selective multiuser mmWave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 353–367, 2018.

[136] S. Gao, X. Cheng, and L. Yang, "Estimating doubly-selective channels for hybrid mmWave massive MIMO systems: A doubly-sparse approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5703–5715, 2020.

[137] M. Jian, F. Gao, Z. Tian, S. Jin, and S. Ma, "Angle-domain aided UL/DL channel estimation for wideband mmWave massive MIMO systems with beam squint," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3515–3527, 2019.

[138] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and S. Li, "Channel estimation for millimeter-wave multiuser MIMO systems via PARAFAC decomposition," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7501–7516, 2016.

[139] J. Yang, C. Wen, S. Jin, and F. Gao, "Beamspace channel estimation in mmWave systems via cosparse image reconstruction technique," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4767–4782, 2018.

[140] L. Lian and V. K. N. Lau, "Configuration optimization and channel estimation in hybrid beamforming mmWave systems with channel support side information," *IEEE J. Sel. Topics Signal Process.*, vol. 68, pp. 6026–6039, 2020.

[141] C. K. Anjinappa, A. C. Gürbüz, Y. Yapıcı, and İ. Güvenç, "Off-grid aware channel and covariance estimation in mmWave networks," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3908–3921, 2020.

[142] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, 2017.

[143] B. Wang, M. Jian, F. Gao, G. Y. Li, and H. Lin, "Beam squint and channel estimation for wideband mmWave massive MIMO-OFDM systems," *IEEE J. Sel. Topics Signal Process.*, vol. 67, no. 23, pp. 5893–5908, 2019.

[144] F. Bellili, F. Sohrabi, and W. Yu, "Generalized approximate message passing for massive MIMO mmWave channel estimation with Laplacian prior," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3205–3219, 2019.

[145] F. Talaei and X. Dong, "Hybrid mmWave MIMO-OFDM channel estimation based

on the multi-band sparse structure of channel," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1018–1030, 2019.

[146] A. Klautau, P. Batista, N. González-Prelcic, Y. Wang, and R. W. Heath, "5G MIMO data for machine learning: Application to beam-selection using deep learning," in *Proc. ITA*. IEEE, 2018, pp. 1–9.

[147] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 912–926, Sep. 2011.

[148] J. Park and B. Clerckx, "Multi-user linear precoding for multi-polarized massive MIMO system under imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2532–2547, May 2015.

[149] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.

[150] E. G. Larsson and H. V. Poor, "Joint beamforming and broadcasting in massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 3058–3070, Apr. 2016.

[151] M. Schubert and H. Boche, "Solution of the multi-user downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Tech.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[152] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. Allerton Conf. on Commun., Control and Comput.*, Sep. 1999, pp. 987–996.

[153] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.

[154] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1313–1324, Oct. 1998.

[155] "IEEE standard for information technology–telecommunications and information exchange between systems—local and metropolitan area networks–specific requirements–part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications–amendment 4: Enhancements for very high throughput for operation in bands below 6 GHz." *IEEE Std 802.11ac(TM)-2013 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012, IEEE Std 802.11aa-2012, and IEEE Std 802.11ad-2012)*, pp. 1–425, Dec 2013.

[156] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sumrate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[157] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[158] L.-N. Tran, M. Juntti, M. Bengtsson, and B. Ottersten, "Weighted sum rate maximization for MIMO broadcast channels using dirty paper coding and zero-forcing methods," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2362–2373, Jun. 2013.

[159] S. Zarei, W. Gerstacker, R. Schober, "A low-complexity linear precoding and power allocation scheme for downlink massive MIMO systems," in *Proc. Asilomar Conf. on Signals, Syst., and Comput.*, 2013, pp. 285–290.

[160] D. H. N. Nguyen and T. Le-Ngoc, "Sum-rate maximization in the multicell MIMO broadcast channel with interference coordination," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1501–1513, Mar. 2014.

[161] J. Kaleva, A. Tölli, and M. Juntti, "Decentralized sum rate maximization with QoS

constraints for interfering broadcast channel via successive convex approximation," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2788–2802, Jun. 2016.

[162] S. Hu, Y. Zhang, X. Wang, and G. B. Giannakis, "Weighted sum-rate maximization for MIMO downlink systems powered by renewables," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5615–5625, Aug. 2016.

[163] H. V. Nguyen, V. D. Nguyen, and O. S. Shin, "Low-complexity precoding for sum rate maximization in downlink massive MIMO systems," *IEEE Wireless Comm. Letters*, vol. 6, no. 2, pp. 186–189, 2017.

[164] Y. Cheng and M. Pesavento, "Robust codebook-based downlink beamforming using mixed integer conic programming," in *Proc. ICASSP*, 2013.

[165] Y. Cheng, and M. Pesavento, "An optimal iterative algorithm for codebook-based downlink beamforming," *IEEE Signal Process. Lett.*, vol. 20, no. 8, pp. 775–778, Aug. 2013.

[166] H. Joudeh and B. Clerckx, "Robust transmission in downlink multiuser MISO systems: A rate-splitting approach," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6227–6242, Dec. 2016.

[167] ——, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4847–4861, Nov. 2016.

[168] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A rate splitting strategy for massive MIMO with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4611–4624, Jul. 2016.

[169] E. Piovano and B. Clerckx, "Optimal DoF region of the $k$-user MISO BC with partial CSIT," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2368–2371, Nov. 2017.

[170] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[171] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.

[172] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[173] Y. Sun, P. Babu and D. P. Palomar, "Robust estimation of structured covariance matrix for heavy-tailed elliptical distributions," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3576–3590, Jul. 2016.

[174] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[175] Are Hjørungnes, *Complex-Valued Matrix Derivatives*. Cambridge University Press, 2011.