5-2017

Low-resolution ADC receiver design, MIMO interference cancellation prototyping, and PHY secrecy analysis.

Chen Cao
University of Louisville

Recommended Citation

Cao, Chen, "Low-resolution ADC receiver design, MIMO interference cancellation prototyping, and PHY secrecy analysis." (2017). *Electronic Theses and Dissertations.* Paper 2697.
https://doi.org/10.18297/etd/2697

LOW-RESOLUTION ADC RECEIVER DESIGN, MIMO INTERFERENCE CANCELLATION PROTOPTYING, AND PHY SECRECY ANALYSIS

By

Chen Cao

A Dissertation
Submitted to the Faculty of the
J. B. Speed School of Engineering of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy in Electrical Engineering

Department of Electrical and Computer Engineering
University of Louisville
Louisville, Kentucky

May 2017

LOW-RESOLUTION ADC RECEIVER DESIGN, MIMO INTERFERENCE CANCELLATION PROTOPTYING, AND PHY SECRECY ANALYSIS

By

Chen Cao
B.S., Xi'an Jiaotong University, China, 2010
M.S., Xi'an Jiaotong University, China, 2013

A Dissertation Approved On

April 14, 2017

by the following Dissertation Committee:

Hongxiang Li, Dissertation Director

Andre Faul

Cindy Harnett

Huacheng Zeng

Lihui Bai

DEDICATION

To all I love.

ACKNOWLEDGMENTS

First of all, I would like to thank my advisor Dr. Hongxiang Li. He helped me a lot during my four year Ph.D. study and research. I could not have written this dissertation without his help. I would also like to thank Dr. Andre Faul, Dr. Cindy Harnett, Dr. Huacheng Zeng and Dr. Lihui Bai for their kind help. I really appreciate all the help and care for me. Thanks to my colleagues, Xiaohui Zhang, Wenqi Liu and Nadieh Moghadam for their support and advice on my research. I would also like to thank Tom Carroll, Lisa Bell, and any others who have helped me these years. Thank my family and friends, you are the hope.

ABSTRACT

LOW-RESOLUTION ADC RECEIVER DESIGN, MIMO INTERFERENCE CANCELLATION PROTOPTYING, AND PHY SECRECY ANALYSIS

Chen Cao

April 14, 2017

This dissertation studies three independent research topics in the general field of wireless communications.

The first topic focuses on new receiver design with low-resolution analog-to-digital converters (ADC). In future massive multiple-input-multiple-output (MIMO) systems, multiple high-speed high-resolution ADCs will become a bottleneck for practical applications because of the hardware complexity and power consumption. One solution to this problem is to adopt low-cost low-precision ADCs instead. In Chapter II, MU-MIMO-OFDM systems only equipped with low-precision ADCs are considered. A new turbo receiver structure is proposed to improve the overall system performance. Meanwhile, ultra-low-cost communication devices can enable massive deployment of disposable wireless relays. In Chapter III, the feasibility of using a one-bit relay cluster to help a power-constrained transmitter for distant communication is investigated. Nonlinear estimators are applied to enable effective decoding.

The second topic focuses prototyping and verification of a LTE and WiFi co-existence system, where the operation of LTE in unlicensed spectrum (LTE-U) is

discussed. LTE-U extends the benefits of LTE and LTE Advanced to unlicensed spectrum, enabling mobile operators to offload data traffic onto unlicensed frequencies more efficiently and effectively. With LTE-U, operators can offer consumers a more robust and seamless mobile broadband experience with better coverage and higher download speeds. As the coexistence leads to considerable performance instability of both LTE and WiFi transmissions, the LTE and WiFi receivers with MIMO interference canceller are designed and prototyped to support the coexistence in Chapter IV.

The third topic focuses on theoretical analysis of physical-layer secrecy with finite blocklength. Unlike upper layer security approaches, the physical-layer communication security can guarantee information-theoretic secrecy. Current studies on the physical-layer secrecy are all based on infinite blocklength. Nevertheless, these asymptotic studies are unrealistic and the finite blocklength effect is crucial for practical secrecy communication. In Chapter V, a practical analysis of secure lattice codes is provided.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

This dissertation studies three independent research topics in the general field of wireless communications. In particular, the studies of these three topics focus on different levels: 1) The first study is new system design and simulation for low-resolution ADC receivers. 2) The second study is system prototyping and verification of a new LTE and WiFi coexistence system based on MIMO interference cancellation. 3) The last study is theoretical analysis physical-layer secrecy.

A Low-Resolution ADC Receiver Design

Massive multiple-input-multiple-output (MIMO) is an enabling technology for next generation wireless communication systems such as 5G. In such systems, an indispensable component of any receiver architecture is the analog-to-digital converter (ADC), typically with a resolution of 8-12 bits. However, the high dimensionality of massive antennas considerably increases the hardware cost and power consumption of RF chains [1–3]. In particular, the power consumption of an ADC increases almost exponentially with the number of bits per sample. As larger and larger bandwidth is used, the corresponding sampling rate of an ADC scales up. As a result, massive high-speed high-resolution ADCs are either unavailable or too costly and power-hungry. Apparently, ADCs will become a major obstacle for next generation wireless transceivers. To overcome this problem, the use of low-resolution ADCs was proposed in [4], etc, which yields significantly quantized massive MIMO systems. The capacity of different quantized systems has been analyzed in [5–10]. Although low-resolution ADCs will incur performance loss, the quantized systems can benefit from the massive antennas while keeping the power consumption and hardware cost under practical constraints. ***In Chapter II, a new turbo receiver structure is proposed for uplink MU-MIMO-OFDM systems based on low-resolution ADCs only.***

The contribution:

1) Unlike the traditional independent quantization error model [11], the interdependency between the input analog signal and the output quantized signal of low-resolution ADCs is exploited and a turbo receiver structure is proposed to effectively reduce the large quantization error. Specifically, using the extrinsic information from the channel decoder, it can estimate the quantization error and enable de-quantization.

2) Based on the turbo receiver structure, both the Linear Minimum Mean Square Error (LMMSE) receiver and the Approximate Message Passing (AMP) receiver are designed. In particular, the turbo AMP receiver is preferred in massive MIMO applications because its computational complexity increases only linearly with the number of receiver antennas.

3) The simulation results demonstrate that the proposed turbo receiver structure can significantly improve the performance of quantized receivers. Specifically, even with highly inaccurate channel state information (CSI), the turbo LMMSE and AMP receiver are superior to the conventional receiver in terms of bit error rate (BER). Compared with the LMMSE receiver, the AMP receiver is less sensitive to CSI so it is more advantageous

under inaccurate CSI conditions.

Besides massive MIMO, another potential application for low-resolution ADC receiver is the Internet of Things (IoT) wireless network. In particular, Ubiquitous communications (UbiComm) that provide reliable connections for anything at anywhere will take center stage of future wireless networks. In such networks, traditional wireless communication transceivers may impede their massive deployment due to their high cost and power consumption. As a result, ultra-low-cost transceivers have become an intriguing alternative to realize future IoT. In particular, small wireless sensors with a hard cost/power constraint will benefit from the ultra-low-cost relays. Furthermore, the ultra-low-cost transceiver will enable massive deployment of disposable wireless relays. Note that these disposable relays can be arbitrarily scattered at places needed to provide a reliable connection, and the failure of some of these relays will not interrupt the connection. ***In Chapter III, the one-bit relay cluster scheme is studied.***

The contribution:

1) This is the first work to study the one-bit relay cluster for distant transmission.

2) A simple DC-offset method is proposed to support the M-QAM modulation.

3) A black-box solution is provided to avoid the unscalable channel estimation process.

4) Two nonlinear estimators are designed with significant performance gains over the traditional linear estimator.

B MIMO Interference Cancellation Prototyping

LTE-unlicensed (LTE-U) technology is initiated as part of LTE Release 13 to allow users to access both licensed and unlicensed spectrum under a unified LTE network infrastructure. LTE-U extends LTE to the unlicensed spectrum and aggregates the unlicensed spectrum with the licensed spectrum. It can provide better coverage and larger capacity than cellular/WiFi interworking while allowing seamless data flow between licensed and unlicensed spectrum through a single network. For operators, LTE-U means synchronized integrated network management, the same authentication procedures, more efficient resource utilization, and thus lower operational costs. The primary challenge of LTE-U is the coexistence with the unlicensed systems such as WiFi systems. The current solution for LTE and WiFi coexistence is to adopt an on/off time pattern in LTE. In on-state, the LTE basestation transmit based on standards. In off-state, the LTE basestation ceases all transmissions while WiFi systems can access these gaps for transmission. This is a simple mechanism to enable both LTE and WiFi systems to work in the unlicensed band. However, the LTE and WiFi performance in the unlicensed spectrum will inevitably fluctuate because of frequent back-off, leading to considerable performance instability. ***To address this problem, in Chapter IV, a prototype of the interference cancelling receivers for LTE and WiFi coexistence is demonstrated.***

The contribution:

1) Both the WiFi and LTE receivers with MIMO interference canceller are designed and implemented on Soft Defined Radios (SDR).

2) The prototype system is tested and evaluated in real world environment.

C PHY Secrecy Analysis

In traditional communication systems, physical-layer channel codes provide the data reliability, while upper layer encryption algorithms and key exchange protocols ensure the data secrecy. Unlike upper layer security approaches, physical-layer communication security can guarantee the information-theoretic secrecy, which is measured quantitatively in terms of the statistical dependence between transmitted messages and observations. Specifically, the state-of-the-art coding schemes are measured by the secrecy capacity, which is defined as the maximum coding rate that can be used by the sender to ensure that the legitimate receivers can decode successfully while the eavesdroppers cannot obtain any information. The basic assumption of the secrecy capacity analysis is the coding blocklength goes to infinity. However, this asymptotic assumption is unrealistic in practical systems. In fact, the effect of finite blocklength on secrecy performance can be quite significant in many practical communication scenarios. ***In Chapter V, a thorough analysis of the finite-blocklength secrecy performance of secure lattice codes is provided.***

The contribution:

1) Two secrecy performance metrics are provided: the leak probability L_P and the average leakage L_A.

2) Since the direct analysis based on L_A and L_P are computationally prohibitive, a practical approach is further proposed to approximate L_A and L_P.

3) The trade-off between the secrecy and the reliability is examined.

4) Finally, the secrecy performance of secure nested lattice codes with finite blocklengths is analyzed.

CHAPTER II

UPLINK MU-MIMO-OFDM RECEIVERS WITH LOW-RESOLUTION ADCS

In a multi-user MIMO (MU-MIMO) system with spatial multiplexing, all users simultaneously transmit their data using the same bandwidth. Conventional massive MU-MIMO systems based on perfect ADCs have been extensively studied. Recently, quantized MU-MIMO systems attract a lot of attention because of the benefits of low-resolution ADCs. In [12, 13], the authors studied the massive MIMO systems with one-bit ADCs. They adopted the least square (LS) channel estimation and use the maximum-ratio-combining (MRC) and zero-forcing (ZF) filter for the MIMO detector. The underlying assumption is that the quantization error is additive and uncorrelated to the analog input signal, which is the classic quantization model [11]. The results show that MIMO is resilient to the quantization error. Along the same line, [14] formulated a maximum likelihood (ML) detection problem, which was solved by a exhaustive search over all possible transmitted vectors. The authors further proposed an iterative algorithm by relaxing the constraints on the transmitted vector. The joint channel-and-data estimation is considered in [15]. The coarse quantization makes the acquisition of channel state information (CSI) more challenging in quantized MIMO systems. The requirement of a long pilot sequence motivates to adopt the joint channel-data estimation. In [15], it proposed a Bayes-optimal inference method based on minimum mean square error (MMSE). The result demonstrates the joint channel-and-data estimation can help the receiver to obtain more accurate CSI in quantized MIMO systems. In [16], the authors focused on the channel estimation with one-bit quantization. They utilized the characteristics of mmWave that the channel is sparse, and formulated the estimation problem as a one-bit compressed sensing problem. In [17], nonlinear estimators such as support vector machine and neural network were exploited to be used in the massive wireless relay transmission with only one-bit ADCs. However all these work [12–17] assume the MU-MIMO operation over flat-fading channel. Orthogonal-frequency-division-modulation (OFDM) is the most adopted technique for frequency-selective channel. An ideal OFDM process will convert the frequency-selective channel into a set of orthogonal flat-fading channels. However, in quantized systems, the linear discrete-Fourier-transform (DFT) operation at a receiver is not able to project the nonlinear-correlated quantization error into orthogonal components at each subcarrier. This is equivalent to incurring nonlinear interference at the subcarriers, which makes it more difficult to mitigate the quantization error compared to the systems only designed for flat-fading channel.

Recent work [18] developed a dedicated MU-MIMO-OFMD detector for the quantized case. Two quantization error models are adopted: the exact and mismatched model. The mismatch quantization model simply regards the quantization error as an additive uncorrelated noise, which is identical as the classic quantization error model [11]. The detector for the mismatch model is exactly the same as the conventional detector designed for infinite-resolution ADC receivers. The detector of the exact model works on solving optimal problems based on MAP or MMSE criterions. Obviously, these optimal problems require high computational complexities. To address the computational complexity issue, the authors suggested a technique named as forward-backward splitting (FBS). However, the reduction of the complexity still rely on exploiting fast transforms instead of matrix products and selecting an appropriate step-size. In [19], the authors exploited the massive MU-MIMO-OFDM system with mixed high-resolution and one-bit ADCs, where the nonlinearity of the one-bit quantization is modeled as a linear inter-symbol interference after the DFT operation. Accordingly, they proposed a linear frequency domain equalizer to maximize the transmission rate when there is only one user. For the multiuser scenario, the multiuser interference is regarded as an additional Gaussian noise so that the linear frequency domain equalizer is obtained the same as the single user case. The authors also exploited the error matrix structure to reduce the computational complexity to the cubic of the number of receiver antennas. [20–28] was first proposed as a channel equalizer to mitigate the inter-symbol interference over frequency-selective channel. In a turbo receiver, there exists a feedback loop from the channel decoder output to the equalizer. Specifically, the channel decoder will feed back the extrinsic information of the transmitted symbols into the equalizer as a priori symbol probabilities. Then, the equalizer combines this a priori information and the input signal to estimate the transmitted symbols. The turbo equalizer works in an iterative way until a stopping criterion is reached. Motivated by the success of the turbo equalizer, the turbo receiver has been extended to various applications such as MIMO detection [29–31], carrier-frequency offset estimation [32–34]. Belief propagation (BP) is a message-passing Bayesian inference framework applied to factor graphical probabilistic models, where messages are sent between factor nodes and variable nodes. The message computation rules are obtained from the stationary points of the Bethe free energy. When the factor graph is free of cycles, BP provides the exact marginal distributions. When the graph has cycles, BP outputs an approximation of the marginal distribution. BP with factor graph has been widely used in the channel decoding and turbo receiver design [35–39]. However, BP still involves high-dimensional matrix operations in large-scale systems. To address the dimensionality issue, approximate-message-passing (AMP) [40–44] is further proposed for large-scale systems to decouple the vector-valued estimation problem into a sequence of scalar problems and linear transforms. Meanwhile, AMP also enables parallel and distributed computing.

Throughout this chapter, the following notations are used: $\mathrm{Quan}\left(\cdot\right)$ denotes the quantization function; $\mathcal{Q}\left(\cdot\right)$ denotes the Q function; $\mathcal{CN}\left(x;\tau,\nu\right)$ denotes the complex Gaussian random variable x with the mean τ and variance ν; $\mathcal{N}\left(x;\tau,\nu\right)$ denotes the real Gaussian random variable x with the mean τ and variance ν; $\mathcal{R}\left(x\right)$ and $\mathcal{I}\left(x\right)$ denote the real part and image part of x, respectively. $U\left(x\right)$ and $L\left(x\right)$ denote the mapping from the quantized value x to its corresponding upper and lower continuous values before the quantization, respectively; $f\left(X\right)|_{X=x_2}^{X=x_1}$ is short for $f\left(x_1\right)-f\left(x_2\right)$; unbold letters (e.g., x) denote scalars and bold letters (e.g., x) denote vectors or matrices. In addition, some frequently used system model notations are summarized in Table I.

TABLE 1

System model notations

n_t	user (transmitter) index
N_t	number of users (transmitters)
n_r	receiver antenna index
N_r	number of receiver antennas
k	subcarrier index
K	number of subcarriers
t	time snapshot (tap) index
T	number of time snapshots (taps)

A System model

Consider a generic MU-MIMO-OFDM uplink system, where the receiver is equipped with multiple antennas while each user has only single antenna. All users intend to send their individual information bits to the receiver using the same time-frequency resource. The information bits are segmented and concatenated into transport blocks in media access (MAC) layer. As only focus on physical-layer, assume there is only one transport block in each user. This transport block first enters a channel encoder followed by a block interleaver. The interleaved-coded bits are then mapped into constellation symbols (QPSK, QAM, etc.). After a series-to-parallel converter, the constellation symbols are grouped into several OFDM symbols for the IDFT operation [1]. Finally, the OFDM symbols are sent to the antenna and transmitted sequentially by each user. To make concise expressions, focus on one specific OFDM symbol. The transmitted signal is

[1] The cyclic-prefix process is not specifically demonstrated in this paper because it does not affect the proposed algorithms.

$\mathbf{X}_{K\times N_t} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{N_t}]$, and the discrete channel frequency response is

$$\mathbf{H}_{N_r\times N_t\times K} = \begin{bmatrix} \mathbf{h}_{11} & \cdots & \mathbf{h}_{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{N_r1} & \cdots & \mathbf{h}_{N_rN_t} \end{bmatrix} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_K]\,. \tag{1}$$

The noise matrix is denoted as $\mathbf{W}_{T\times N_r} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{N_r}]$ and the elements in $\mathbf{W}_{T\times N_r}$ are i.i.d random variables with the Gaussian distribution $N\left(0, \sigma^{\mathrm{w}}\right)$. In Figure 1, the analog signal before the ADCs can be expressed as

$$\mathbf{y}_{n_r,T\times 1} = \frac{1}{K}\sum_{n_t=1}^{N} \mathbf{F}_{T\times K}\mathrm{diag}\left(\mathbf{h}_{n_rn_t,K\times 1}\right)\mathbf{x}_{n_t,K\times 1} + \mathbf{w}_{n_r,T\times 1}, \tag{2}$$

where $\mathbf{F}$ is the IDFT matrix and $\mathbf{Y}_{T\times N_r} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_{N_r}]$.

B Conventional receiver with uncorrelated quantization error model

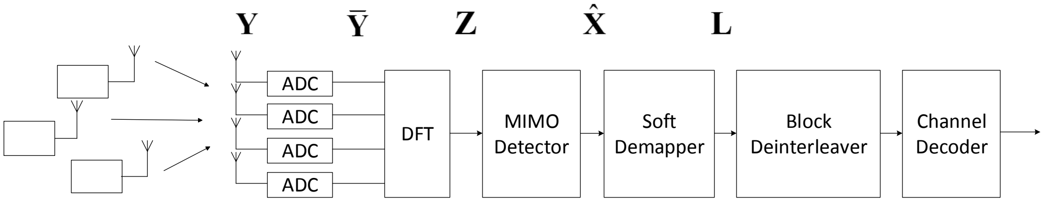


Figure 1: Conventional receiver structure for the MU-MIMO-OFDM system

In the conventional receiver design, the quantization error is approximated as a Gaussian noise and spread more or less uniformly over the Nyquist bandwidth [11]. The underlying assumption is that the quantization error is uncorrelated to the input analog signal, which is accurate enough when high-resolution ADCs are used. In this case, the quantization error can be treated as a channel noise so that it can adopt the conventional receiver structure shown in Figure1.

After the quantization, the digital signal is

$$\bar{\mathbf{Y}}_{T\times N_r} = \text{Quan}\left(\mathbf{Y}_{T\times N_r}\right) = \text{Quan}\left(\mathcal{R}\left(\mathbf{Y}_{T\times N_r}\right)\right) + \text{Quan}\left(\mathcal{I}\left(\mathbf{Y}_{T\times N_r}\right)\right). \tag{3}$$

The quantized signal is modeled as

$$\bar{\mathbf{Y}}_{T\times N_r} = \mathbf{Y}_{T\times N_r} + \boldsymbol{\Omega}_{T\times N_r}, \tag{4}$$

where the elements in $\boldsymbol{\Omega}_{T\times N_r}$ are i.i.d noises with the distribution $\mathcal{CN}\left(0, \sigma^{\omega}\right)$. Because the noise here is assumed additive and uncorrelated, the DFT operation can be applied directly to yield

$$\mathbf{Z}_{K\times N_r} = \mathbf{F}^{*}\bar{\mathbf{Y}}_{T\times N_r} = \mathbf{F}^{*}\mathbf{Y}_{T\times N_r} + \mathbf{F}^{*}\boldsymbol{\Omega}_{T\times N_r}. \tag{5}$$

Let $\tilde{\mathbf{W}}_{K\times N_r} = \mathbf{F}^{*}\tilde{\mathbf{W}}_{T\times N_r}$ and $\tilde{\mathbf{W}}^{d}_{K\times N_r} = \mathbf{F}^{*}\tilde{\mathbf{W}}^{d}_{T\times N_r}$, it gives

$$\mathbf{z}^{T}_{k,1\times N_r} = \mathbf{h}_{k,N_r\times N_t}\mathbf{x}^{T}_{k,1\times N_t} + \tilde{\mathbf{w}}^{T}_{k,1\times N_r} + \tilde{\boldsymbol{\omega}}^{T}_{k,1\times N_r}. \tag{6}$$

For a specific k, the elements in $\tilde{\mathbf{w}}_{k,1\times N_r}$ (resp. $\tilde{\boldsymbol{\omega}}_{k,1\times N_r}$) are i.i.d noises with the distribution $\mathcal{CN}\left(0, T\sigma^{\mathrm{w}}\right)$ (resp. $\mathcal{CN}\left(0, T\sigma^{\omega}\right)$).

1 Channel estimation

At the user (transmitter), the IDFT operation is directly applied to the pre-defined reference symbols $\mathbf{X}^{P}$ as pilot signals. To improve the channel estimation accuracy, each user transmits the pilot with time-division to avoid multiuser interference. At the receiver side, the pilot signal after the DFT operation is

$$\mathbf{Z}^{P}_{K\times N_r\times N_t} = \mathbf{H}_{K\times N_r\times N_t}\cdot\mathbf{X}^{P} + \tilde{\mathbf{W}}^{P}_{K\times N_r\times N_t} + \tilde{\boldsymbol{\Omega}}^{P}_{K\times N_r\times N_t}, \tag{7}$$

The channel estimation is based on maximum a posterior (MAP) with the Gaussian assumption:

$$\hat{\mathbf{H}} = \arg\max_{\mathbf{H}} P\left(\mathbf{Z}_1^P, \mathbf{Z}_2^P, \cdots |\mathbf{H}\right) = \arg\min_{\mathbf{H}} \sum_i \left\|\mathbf{H}\cdot\mathbf{X}_i^P - \mathbf{Z}_i^P\right\|^2, \tag{8}$$

and the solution is given by

$$\hat{\mathbf{H}} = \frac{\sum\limits_i \mathrm{conj}\left(\mathbf{X}_i^P\right)\cdot\mathbf{Z}_i^P}{\sum\limits_i \mathrm{conj}\left(\mathbf{X}_i^P\right)\cdot\mathbf{X}_i^P}, \tag{9}$$

where the dot-product and divide operations are both element-wise operations.

2 LMMSE MIMO detector

While the maximum a posterior (MAP) data signal detector is optimal, it is impractical for real applications because of its high computational complexity. Among suboptimal linear MIMO detectors, the minimum mean square error (MMSE) detector has the best performance. By solving the optimization problem

$$\min_{\hat{\mathbf{x}}_{k,1\times N_t}} E\left[\left\|\hat{\mathbf{x}}_{k,1\times N_t} - \mathbf{x}_{k,1\times N_t}\right\|^2\right], \tag{10}$$

the estimated $\mathbf{X}$ is given as

$$\hat{\mathbf{x}}_{k,1\times N_t}^T {=}^x\mathbf{C}_{N_t\times N_t}\hat{\mathbf{h}}_{k,N_r\times N_t}^H\left(\hat{\mathbf{h}}_{k,N_r\times N_t}{}^x\mathbf{C}_{N_t\times N_t}\hat{\mathbf{h}}_{k,N_r\times N_t}^H{+}^w\mathbf{C}_{N_r\times N_r}{+}^\omega\mathbf{C}_{N_r\times N_r}\right)^{-1}\mathbf{z}_{k,1\times N_r}^T, \tag{11}$$

where ${}^x\mathbf{C}$ is the auto-covariance matrix of $\mathbf{x}_k$, ${}^{\mathrm{w}}\mathbf{C}$ and ${}^\omega\mathbf{C}$ are the auto-covariances of the channel noise and the quantization error with ${}^{\mathrm{w}}\mathbf{C} \sim \mathrm{diag}\left(T\sigma^{\mathrm{w}}\right)$ and ${}^\omega\mathbf{C} \sim \mathrm{diag}\left(T\sigma^\omega\right)$, respectively. The quantization error covariance after the LMMSE detector is given as

$$\begin{aligned}&{}^e\mathbf{C}_{k,N_t\times N_t} =^x \mathbf{C}_{N_t\times N_t}- \\ &{}^x\mathbf{C}_{N_t\times N_t}\hat{\mathbf{h}}_{k,N_r\times N_t}^H\left(\hat{\mathbf{h}}_{k,N_r\times N_t}{}^x\mathbf{C}_{N_t\times N_t}\hat{\mathbf{h}}_{k,N_r\times N_t}^H{+}^w\mathbf{C}_{N_r\times N_r}{+}^\omega\mathbf{C}_{N_r\times N_r}\right)^{-1}\hat{\mathbf{h}}_{k,N_r\times N_t}{}^x\mathbf{C}_{N_t\times N_t}.\end{aligned} \tag{12}$$

With the input $\hat{\mathbf{x}}_k$ and error covariance ${}^e\mathbf{C}_k$, the soft demapper translates the results of the LMMSE detector into log likehood ratios (LLRs). With the estimated symbol $\hat{x}_{kn_t}$, the probability that the constellation symbol ϕ_i was transmitted is

$$P\left(\phi_i|\hat{x}_{kn_t}\right) \sim \exp\left(-\frac{\left(\hat{x}_{kn_t}-\phi_i\right)^2}{2\sigma^e_{kn_t}}\right), \tag{13}$$

where $\sigma^e_{kn_t}$ is the diagonal element of ${}^e\mathbf{C}_k$, $\boldsymbol{\sigma}^e_k = \mathrm{diag}\left({}^e\mathbf{C}_k\right)$, $\sigma^e_k = \left[\sigma^e_{k1},\cdots,\sigma^e_{kn_t},\cdots,\sigma^e_{kN_t}\right]$ and $\hat{\mathbf{x}}_k = \left[\hat{x}_{k1},\cdots,\hat{x}_{kn_t},\cdots,\hat{x}_{kN_t}\right]$. Therefore, the LLR of the jth bit of x_{kn_t} can be calculated as

$$l_{kn_tj} = \log\frac{\sum\limits_i P\left(\phi_i|\hat{x}_{kn_t}\right), i \in \{j\text{th bit of } \phi_i \text{ is } '0'\}}{\sum\limits_i P\left(\phi_i|\hat{x}_{kn_t}\right), i \in \{j\text{th bit of } \phi_i \text{ is } '1'\}}. \tag{14}$$

The LLR l_{kn_tj} is then sent to the block deinterleaver and the channel decoder to estimate the source bits.

C Turbo LMMSE receiver design

The conventional receiver design in Section III assumes the quantization error is independent of the analog input signal. However, this assumption is invalid and can cause significant performance loss when low-resolution ADCs are used. In this section, the interdependence of the quantization error and the analog input signal is fully considered, and a new turbo receiver structure is proposed to effectively reduce the large quantization error caused by low-resolution ADCs.

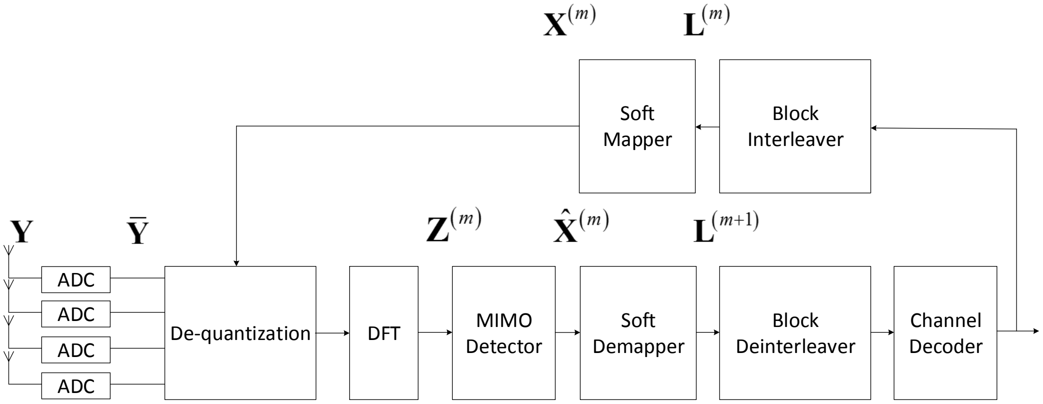


Figure 2: Turbo LMMSE receiver for de-quantization

1 Turbo LMMSE receiver design

The dependence of the quantization error $\mathbf{\Omega}$ on the input signal $\mathbf{Y}$ can be modeled as

$$\mathbf{\Omega}\left(\mathbf{Y}\right) = \mathrm{Quan}\left(\mathbf{Y}\right) - \mathbf{Y}. \tag{15}$$

A simplified turbo receiver structure based on Equation(15) is shown in Figure2. The basic idea is described as follows: using the extrinsic information of $\mathbf{Y}$ provided by the channel decoder, it can obtain an estimation $\hat{\mathbf{Y}}$ and thus the estimated quantization error $\mathbf{\Omega}\left(\hat{\mathbf{Y}}\right)$. Then $\mathbf{\Omega}\left(\hat{\mathbf{Y}}\right)$ will be subtracted from $\mathrm{Quan}\left(\mathbf{Y}\right)$ to reduce the quantization error. The whole detection and decoding process is realized in an iterative way with m as the iteration index.

Let $l_{kn_tj}^{(m)}$ denote the updated LLR from the channel decoder and block interleaver, the soft mapper produces the estimated constellation symbol as

$$\hat{x}_{kn_t}^{(m)} = \sum_{\phi\in\Phi} \phi P\left(\phi | l_{kn_t1}^{(m)}, l_{kn_t2}^{(m)}, \cdots\right), \tag{16}$$

$$\sigma_{kn_t}^{\hat{x}(m)} = \sum_{\phi\in\Phi} |\phi|^2 P\left(\phi | l_{kn_t1}^{(m)}, l_{kn_t2}^{(m)}, \cdots\right) - \left|\hat{x}_{kn_t}^{(m)}\right|^2. \tag{17}$$

Using the estimated symbol $\hat{x}_{kn_t}^{(m)}$, it can further estimate the analog signal before the quantization as

$$\hat{\mathbf{y}}_{n_r}^{(m)} = \frac{1}{K} \sum_{n_t=1}^{N_t} \mathbf{F} \mathrm{diag}\left(\hat{\mathbf{h}}_{n_r n_t}\right) \hat{\mathbf{x}}_{n_t}^{(m)}. \tag{18}$$

Therefore, the estimated quantization error is

$$\mathbf{\Omega}^{(m)} = \mathrm{Quan}\left(\hat{\mathbf{Y}}^{(m)}\right) - \hat{\mathbf{Y}}^{(m)}. \tag{19}$$

The residual error after the de-quantization is

$$\Omega\left(\mathbf{Y}\right) - \Omega\left(\hat{\mathbf{Y}}\right) = \mathrm{Quan}\left(\mathbf{Y}\right) - \mathrm{Quan}\left(\hat{\mathbf{Y}}\right) + \hat{\mathbf{Y}} - \mathbf{Y}. \tag{20}$$

Assume that the estimated $\hat{\mathbf{Y}}$ follows the Gaussian distribution so it is equivalent to an add-on Gaussian noise matrix $\mathbf{W}^{\hat{y}} \sim \mathcal{CN}\left(0, \sigma^{\hat{y}}\right)$. Thus, the residual error (20) can be revised as

$$\Omega\left(\mathbf{Y}\right) - \Omega\left(\hat{\mathbf{Y}}\right) = \mathrm{Quan}\left(\mathbf{Y}\right) - \mathrm{Quan}\left(\mathbf{Y} + \mathbf{W}^{\hat{y}}\right) + \mathbf{W}^{\hat{y}}. \tag{21}$$

When the elements of $\mathbf{W}^{\hat{y}}$ are much smaller than the quantization unit, the quantized signal stays the same as $\mathrm{Quan}\left(\mathbf{Y}\right) = \mathrm{Quan}\left(\mathbf{Y} + \mathbf{W}^{\hat{y}}\right)$. Therefore, the residual error is equal to the estimation error as

$$\Omega\left(\mathbf{Y}\right) - \Omega\left(\hat{\mathbf{Y}}\right) = \mathbf{W}^{\hat{y}}. \tag{22}$$

With the de-quantization process, the de-quantized signal after the DFT process is

$$\mathbf{Z}^{(m)} = \mathbf{F}^H \left(\bar{\mathbf{Y}} - \mathbf{\Omega}^{(m)}\right), \tag{23}$$

and the residual error power after the DFT operation can be estimated by

$${}^{re}\mathbf{C}_k^{(m)} = \hat{\mathbf{h}}_k \mathrm{diag}\left(\boldsymbol{\sigma}_k^{\hat{x}(m)}\right) \hat{\mathbf{h}}_k^H. \tag{24}$$

Because the overall noise power (i.e., ${}^{\mathrm{w}}\mathbf{C} + {}^{re}\mathbf{C}_k^{(m)}$) updates in each iteration, the LMMSE detector coefficients have to be re-calculated based on the updated noise covariance matrix as

$$\mathbf{\Sigma}_k^{(m)} = {}^{x}\mathbf{C}\hat{\mathbf{h}}_k^H \left(\hat{\mathbf{h}}_k {}^{x}\mathbf{C}\hat{\mathbf{h}}_k^H + {}^{\mathrm{w}}\mathbf{C} + {}^{re}\mathbf{C}_k^{(m)} \right)^{-1}. \tag{25}$$

Therefore, the updated estimated signal matrix and error matrix are given by

$$\hat{\mathbf{x}}_k^{(m)T} = \mathbf{\Sigma}_k^{(m)} \mathbf{z}_k^{(m)T}, \tag{26}$$

$${}^{e}\mathbf{C}_k^{(m)} = {}^{x}\mathbf{C} - \mathbf{\Sigma}_k^{(m)} \hat{\mathbf{h}}_k {}^{x}\mathbf{C}. \tag{27}$$

The soft demapper produces the updated LLRs based on the updated symbol estimation as

$$P\left(\phi | \hat{x}_{kn_t}^{(m)}\right) \sim \exp\left(-\frac{\left(\hat{x}_{kn_t}^{(m)} - \phi_i\right)^2}{2\sigma_{kn_t}^{e(m)}} \right), \tag{28}$$

$$l_{kn_tj}^{(m+1)} = \log \frac{\sum\limits_{\phi} P\left(\phi | \hat{x}_{kn_t}^{(m)}\right), i \in \{j\text{th bit of } \phi \text{is } '0'\}}{\sum\limits_{\phi} P\left(\phi | \hat{x}_{kn_t}^{(m)}\right), i \in \{j\text{th bit of } \phi \text{is } '1'\}}. \tag{29}$$

2 Complexity

For the turbo receiver structure, the computational complexity is a key factor for practical applications. Therefore, it is necessary to investigate the computational complexity of the receiver in one iteration between the softer mapper and demapper. The complexities for the matrix multiplication and inversion operation depend on specific algorithms. Assume that arithmetic with individual elements has complexity $O(1)$, as is the case with fixed-precision floating-point arithmetic. In this section, adopt the general setting: the complexity of the multiplication of one $n \times m$ matrix and one $m \times p$ matrix is $O(nmp)$; and the complexity of the inversion of one $n \times n$ matrix is $O(n^3)$. In one iteration, the complexity to calculate $\hat{\mathbf{y}}_{n_r}^{(m)}$ using Equation(18) is $O(N_r N_t K) + O(N_r N_t T) + O(N_r N_t T K)$. Then, the complexity to estimate the quantization error $\boldsymbol{\Omega}^{(m)}$ using Equation(19) is only $O(N_r T)$. The de-quantized signal $\mathbf{Z}^{(m)}$ is calculated using Equation(23) with complexity $O(N_r T) + O(N_r K T)$. In each iteration, it needs to re-calculate the residual error variance matrix ${}^{re}\mathbf{C}_k^{(m)}$ using Equation(24), and its complexity is $O(N_t N_r^2 K)$. Because ${}^{re}\mathbf{C}_k^{(m)}$ updates in each iteration, the LMMSE detector needs to be updated using Equation(25) with complexity $O(N_t N_r K) + O(N_t N_r^2 K) + O(N_r K) + O(N_r^3)$. Thus, the estimated symbol $\hat{\mathbf{x}}_k^{(m)}$ and error covariance ${}^{e}\mathbf{C}_k^{(m)}$ are updated using Equation(26) and Equation(27), and their complexities are $O(N_r N_t)$ and $O(N_r N_t^2) + O(N_t^2) + O(N_t)$, respectively. Overall, the complexity for the turbo LMMSE receiver is $O\left(N_r N_t K T\right) + O\left(N_r^3\right) + O\left(N_r^2 N_t K\right) + O\left(N_t^2\right)$. Apparently, for massive MIMO systems, the complexity could be overwhelming because of $O\left(N_r^3\right) + O\left(N_r^2 N_t K\right)$.

D Turbo AMP receiver design

Belief propagation (BP) is an iterative message-passing type estimator. To reduce the complexity of the turbo receiver, it can adopt approximate-message-passing (AMP) [40–43] as an alternative signal detector. Figure3 shows the AMP receiver structure. Instead of vector estimations, the AMP detector works on a decoupled-scalars factor graph. For the AMP detector design, it uses the factor graph shown in Figure4, in which the square and circle represents the factor node and variable node, respectively. The messages are passed between the factor nodes and variable nodes. Note that each message represents a specific probability density function (PDF).

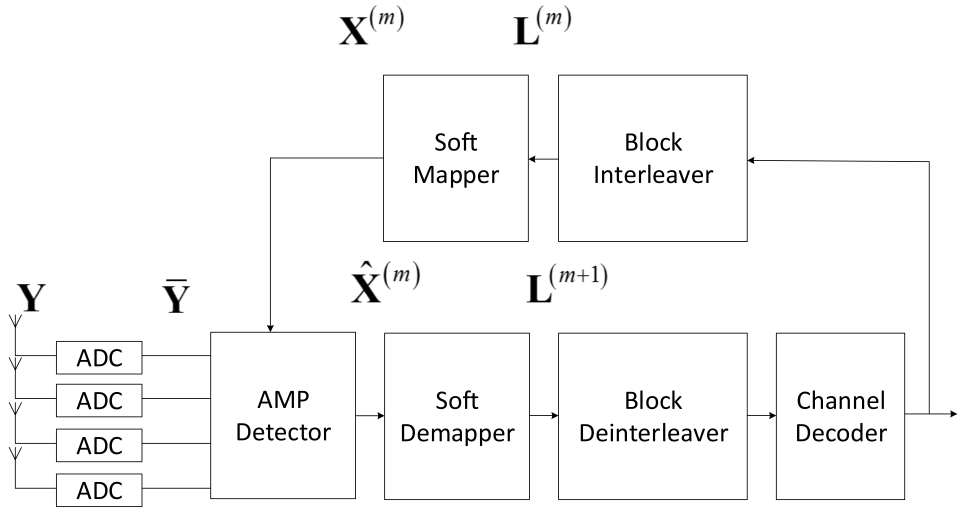


Figure 3: Turbo AMP receiver for de-quantization

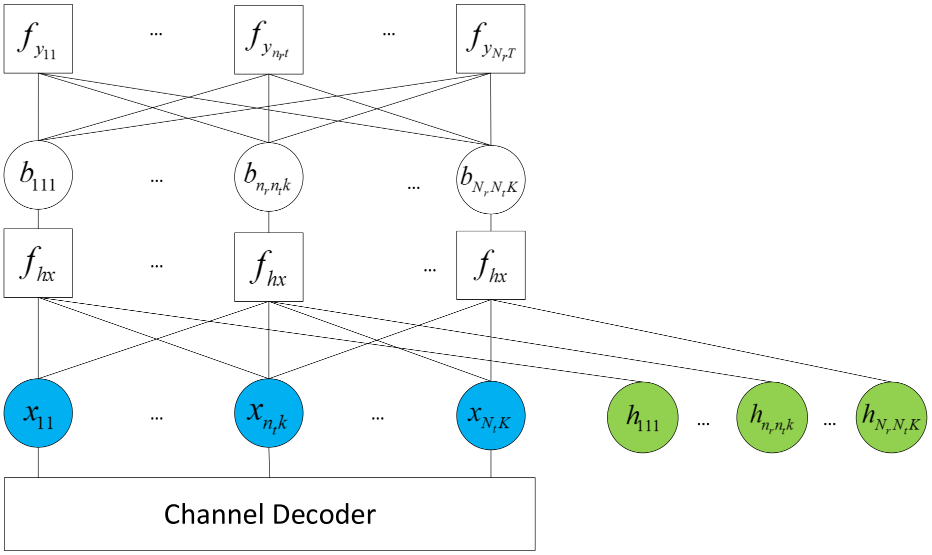


Figure 4: Factor graph for the AMP detector

1 Message-passing algorithm

1) calculate $\mu_{y_t \to b_{n_r n_t k}}$

As shown in Figure5, $\mu_{y_t \to b_{n_r n_t k}}$ is the message sent from the factor node $f_{y_{n_r t}}$ to

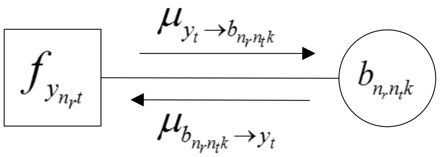


Figure 5: Messages between the factor node $f_{y_{n_rt}}$ and the variable node $b_{n_rn_tk}$

the variable node $b_{n_rn_tk}$. First of all, the conditional probability of y_{n_rt} is

$$f\left(y_{n_rt}|\mathbf{b}_{n_r}\right) = \int\limits_{L(y_{n_rt})}^{U(y_{n_rt})} \mathcal{CN}\left(y - \sum_{n_t=1}^{N_t}\sum_{k=1}^{K} b_{n_rn_tk} e^{j2\pi(t-1)(k-1)/K}; 0, \sigma^{\mathrm{w}}\right) dy. \tag{30}$$

Thus, the joint probability is further given as

$$f\left(y_{n_rt}, \mathbf{b}_{n_r}\right) = f\left(y_{n_rt}|\mathbf{b}_{n_r}\right) \prod_{n_r=1 k=1}^{N_rK} \mu_{b_{n_rn_tk\to y_t}}. \tag{31}$$

To decouple $\mathbf{b}_{n_r}$ into scalar variables, it could adopt the mean-field (MF) approximation and derive $\mu_{y_t\to b_{n_rn_tk}}$ as

$$\mu_{y_t\to b_{n_rn_tk}} \propto \frac{\exp\left(E[\ln f\left(y_{n_rt}, \mathbf{b}_{n_r}\right)]_{\prod\limits_{n_t'k'\neq n_tk} \mu_{b_{n_rn_t'k'\to y_t}}}\right)}{\mu_{b_{n_rn_tk\to y_t}}}. \tag{32}$$

However, different from the convectional systems without quantization, there exists an integral part in $f\left(y_{n_rt}, \mathbf{b}_{n_r}\right)$ so that the calculation of Equation(32) is non-trivial. Therefore, it goes back to the original sum-product (SP) algorithm with the assumption that $\mu_{b_{n_rn_tk\to y_t}}$ is a Gaussian distribution. Specifically, denote

$$\tau_{y_{n_rt}\backslash b_{n_rn_tk}} = \sum_{n_t'k'\neq n_tk} \tau_{b_{n_rn_t'k'}\to y_t} e^{j2\pi(t'-1)(k'-1)/K}, \tag{33}$$

$$\nu_{y_{n_rt}\backslash b_{n_rn_tk}} = \sum_{n_t'k'\neq n_tk} \nu_{b_{n_rn_t'k'}\to y_t}, \tag{34}$$

where $\tau_{b_{n_r n_t k} \to y_t}$ and $\nu_{b_{n_r n_t k} \to y_t}$ are the expectation and variance of the probability distribution $\mu_{b_{n_r n_t k} \to y_t}$. It can derive a new decoupled conditional probability as

$$f\left(y_{n_r t} | b_{n_r n_t k}\right) = \int\limits_{L(y_{n_r t})}^{U(y_{n_r t})} \mathcal{CN}\left(y; \tau_{y_{n_r t} \backslash b_{n_r n_t k}} + b_{n_r n_t k} e^{j2\pi(t-1)(k-1)/K}, \nu_{y_{n_r t} \backslash b_{n_r n_t k}} + \sigma^{\mathrm{w}}\right) dy. \tag{35}$$

The a posterior probability is further given by

$$f\left(b_{n_r n_t k} | y_{n_r t}\right) = \frac{f\left(y_{n_r t} | b_{n_r n_t k}\right) f^{\text{prior}}\left(b_{n_r n_t k}\right)}{f^{\text{prior}}\left(y_{n_r t}\right)}, \tag{36}$$

where $f^{\text{prior}}\left(b_{n_r n_t k}\right)$ and $f^{\text{prior}}\left(y_{n_r t}\right)$ are the prior probability for $b_{n_r n_t k}$ and $y_{n_r t}$. With the Gaussian approximation, the message $\mu_{y_t \to b_{n_r n_t k}}$ is

$$\mu_{y_t \to b_{n_r n_t k}} \approx \mathcal{CN}\left(b_{n_r n_t k}; \tau_{y_t \to b_{n_r n_t k}}, \nu_{y_t \to b_{n_r n_t k}}\right), \tag{37}$$

where

$$\tau_{y_t \to b_{n_r n_t k}} = \int b f\left(b | y_{n_r t}\right) db, \tag{38}$$

$$\nu_{y_t \to b_{n_r n_t k}} = \int b^2 f\left(b | y_{n_r t}\right) db - \tau^2_{y_t \to b_{n_r n_t k}}. \tag{39}$$

However, the calculation of Equation(38) and Equation(39) still requires the integral operation. To simplify the calculation, further assume the prior distribution of $b_{n_r n_t k}$ is also Gaussian. Let

$$f^{\text{prior}}\left(b_{n_r n_t k}\right) = \mathcal{CN}\left(b_{n_r n_t k}; 0, \sigma^b\right), \tag{40}$$

and define

$$\gamma_{y_t \to b_{n_r n_t k}} = \tau_{y_t \to b_{n_r n_t k}} e^{-j2\pi(t-1)(k-1)/K}, \tag{41}$$

It can derive

$$\mathcal{R}\left(\gamma_{y_t \to b_{n_r n_t k}}\right) = \frac{\sigma^b \mathcal{N}\left(y; \mathcal{R}\left(\tau_{y_{n_r t} \backslash b_{n_r n_t k}}\right), \nu_{y_{n_r t} \backslash b_{n_r n_t k}} + \sigma^{\mathrm{w}} + \sigma^b\right)\Big|_{y=U(\mathcal{R}(y_{n_r t}))}^{y=L(\mathcal{R}(y_{n_r t}))}}{\mathcal{Q}\left(\left(y - \mathcal{R}\left(\tau_{y_{n_r t} \backslash b_{n_r n_t k}}\right)\right) / \left(\nu_{y_{n_r t} \backslash b_{n_r n_t k}} + \sigma^{\mathrm{w}} + \sigma^b\right)\right)\Big|_{y=U(\mathcal{R}(y_{n_r t}))}^{y=L(\mathcal{R}(y_{n_r t}))}}, \tag{42}$$

$$\mathcal{I}\left(\gamma_{y_t\to b_{n_rn_tk}}\right) = \frac{\sigma^b\mathcal{N}\left(y;\mathcal{I}\left(\tau_{y_{n_rt}\backslash b_{n_rn_tk}}\right), \nu_{y_{n_rt}\backslash b_{n_rn_tk}} + \sigma^{\mathrm{w}} + \sigma^b\right)\Big|_{y=U(\mathcal{I}(y_{n_rt}))}^{y=L(\mathcal{I}(y_{n_rt}))}}{\mathcal{Q}\left(\left(y-\mathcal{I}\left(\tau_{y_{n_rt}\backslash b_{n_rn_tk}}\right)\right)/\left(\nu_{y_{n_rt}\backslash b_{n_rn_tk}} + \sigma^{\mathrm{w}} + \sigma^b\right)\right)\Big|_{y=U(\mathcal{I}(y_{n_rt}))}^{y=L(\mathcal{I}(y_{n_rt}))}}. \quad (43)$$

The detailed derivation of Equation(42) and Equation(43) is provided in Appendix. Thus, it gives

$$\tau_{y_t\to b_{n_rn_tk}} = \left(R\left(\gamma_{y_t\to b_{n_rn_tk}}\right) + jI\left(\gamma_{y_t\to b_{n_rn_tk}}\right)\right)e^{j2\pi(t-1)(k-1)/K}. \quad (44)$$

It can be easily verified that

$$\lim_{L(y_{smt})-U(y_{n_rt})\to 0} \tau_{y_{n_rt}\to b_{n_rn_tk}} = \frac{\sigma^{\mathrm{b}}\left(y_{n_rt} - \tau_{y_{n_rt}\backslash b_{n_rn_tk}}\right)e^{-j2\pi(t-1)(k-1)/K}}{\nu_{y_{n_rt}\backslash b_{n_rn_tk}} + \sigma^{\mathrm{w}} + \sigma^b} \triangleq \dot{\tau}_{y_{n_rt}\to b_{n_rn_tk}}, \quad (45)$$

and

$$\nu_{y_t\to b_{n_rn_tk}} = \frac{\nu_{y_{n_rt}\backslash b_{n_rn_tk}}\sigma^b + \sigma^{\mathrm{w}}\sigma^b}{\nu_{y_{n_rt}\backslash b_{n_rn_tk}} + \sigma^{\mathrm{w}} + \sigma^b} + \left|\tau_{y_{n_rt}\to b_{n_rn_tk}} - \dot{\tau}_{y_{n_rt}\to b_{n_rn_tk}}\right|^2. \quad (46)$$

Therefore, Equation(41)-(46) provide a feasible close-form expression to calculate the message $\mu_{y_t\to b_{n_rn_tk}}$.

2) calculate $\mu_{b_{n_rn_tk}\to y_t}$

To avoid the loop in the factor graph, only use the extrinsic information from the channel decoder. Thus, it adopts $\mu_{b_{n_rn_tk}\to y_t} = \mu_{hx\to b_{n_rn_tk}}$.

3) calculate $\mu_{b_{n_rn_tk}\to hx}$

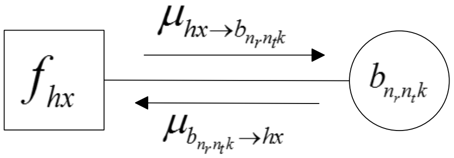


Figure 6: Messages between the factor node f_{hx} and the variable node $b_{n_rn_tk}$

In Figure6, $\mu_{b_{n_rn_tk}\to hx}$ is the message sent from the variable node $b_{n_rn_tk}$ to the factor node f_{hx}. According to the SP algorithm, it gives

$$\mu_{b_{n_rn_tk}\to hx} = \prod_{t=1}^{T} \mu_{y_t\to b_{n_rn_tk}} = \mathcal{CN}\left(b; \tau_{b_{n_rn_tk}\to y_t}, \nu_{b_{n_rn_tk}\to y_t}\right), \tag{47}$$

where $\tau_{b_{n_rn_tk}\to hx}$ and $\nu_{b_{n_rn_tk}\to hx}$ are given by

$$\tau_{b_{n_rn_tk}\to hx} = \nu_{b_{n_rn_tk}\to hx} \sum_{t=1}^{T} \frac{\tau_{y_t\to b_{n_rn_tk}}}{\nu_{y_t\to b_{n_rn_tk}}}, \tag{48}$$

$$\nu_{b_{n_rn_tk}\to hx} = \left(\sum_{t=1}^{T} \frac{1}{\nu_{y_t\to b_{n_rn_tk}}}\right)^{-1}. \tag{49}$$

4) calculate $\mu_{hx\to b_{n_rn_tk}}$

$\mu_{hx\to b_{n_rn_tk}}$ is the message sent from the factor node f_{hx} to the variable node $b_{n_rn_tk}$ and it is given by

$$\mu_{hx\to b_{n_rn_tk}} = \int\limits_{hx=b} \mu_{x_{n_tk}\to hx}\mu_{h_{n_rn_tk}\to hx} dxdh = \mathcal{CN}\left(b; \tau_{hx\to b_{n_rn_tk}}, \nu_{hx\to b_{n_rn_tk}}\right), \tag{50}$$

where $\tau_{hx\to b_{n_rn_tk}}$ and $\nu_{hx\to b_{n_rn_tk}}$ are calculated by

$$\tau_{hx\to b_{n_rn_tk}} = \tau_{x_{n_tk}\to hx}\tau_{h_{n_rn_tk}\to hx}, \tag{51}$$

$$\nu_{hx\to b_{n_rn_tk}} = \upsilon_{x_{n_tk}\to hx}\nu_{h_{n_rn_tk}\to hx} + \tau^2_{x_{n_tk}\to hx}\nu_{h_{n_rn_tk}\to hx} + \tau^2_{h_{n_rn_tk}\to hx}\nu_{x_{snk}\to hx}, \tag{52}$$

where $\tau_{x_{n_tk}\to hx}$, $\nu_{x_{n_tk}\to hx}$ are given in Section 1-6). $\tau_{h_{n_rn_tk}\to hx}$ and $\nu_{h_{n_rn_tk}\to hx}$ are obtained from the channel estimation:

$$\tau_{h_{n_rn_tk}\to hx} = \hat{h}_{n_rn_tk}, \tag{53}$$

$$\upsilon_{h_{n_rn_tk}\to hx} = \tilde{\mathrm{w}}^p_{n_rn_tk} + \tilde{\omega}^p_{n_rn_tk}, \tag{54}$$

where $\hat{h}_{n_rn_tk}$ is the element of $\hat{\mathbf{H}}$ in Equation(7), $\tilde{\mathrm{w}}^p_{n_rn_tk}$ and $\tilde{\omega}^p_{n_rn_tk}$ are the elements of $\tilde{\mathbf{W}}^P$ and $\tilde{\mathbf{\Omega}}^P$, respectively, in Equation(9).

5) calculate $\mu_{hx\to x_{n_tk}}$ In Figure7, $\mu_{hx\to x_{n_tk}}$ is the message from the factor node f_{hx}

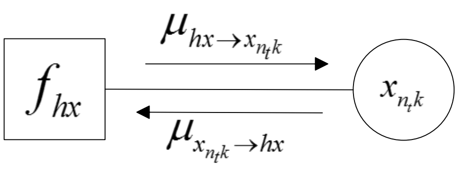


Figure 7: Messages between the factor node f_{hx} and the variable node $x_{n_t k}$

to the variable node $x_{n_t k}$:

$$\mu_{hx \to x_{n_t k}} = \int \mu_{b_{n_r n_t k} \to hx} \mu_{h_{n_r n_t k} \to hx} dh. \tag{55}$$

However, unlike Equation(50), Equation(55) is not a Gaussian distribution and too complicated for efficient message-passing, so use the following approximation:

$$\hat{\mu}_{hx \to x_{n_t k}} \approx \mathcal{CN}\left(x; \tau_{hx \to x_{n_t k}}, \nu_{hx \to x_{n_t k}}\right), \tag{56}$$

where

$$\tau_{hx \to x_{n_t k}} = \int x \mu_{hx \to x_{n_t k}} dx, \tag{57}$$

$$\nu_{hx \to x_{n_t k}} = \int x^2 \mu_{hx \to x_{n_t k}} dx - \tau^2_{hx \to x_{n_t k}}, \tag{58}$$

Similar to Equation(38)(39), Equation(57) has a high complexity, so adopt the MF approximation:

$$\begin{aligned} \mu_{hx \to x_{n_t k}} &\propto \exp\left(E\left[\ln \mu_{b_{n_r n_t k} \to hx} \mu_{h_{n_r n_t k} \to hx}\right]_{\mu_{h_{n_r n_t k} \to hx}}\right) \\ &= \mathcal{CN}\left(x; \tau_{hx \to x_{n_t k}}, \nu_{hx \to x_{n_t n_r k}}\right), \end{aligned} \tag{59}$$

and

$$\tau_{hx \to x_{n_t k}} = \frac{\tau^*_{h_{n_r n_t k} \to hx} \tau_{b_{n_r n_t k} \to hx}}{\nu_{h_{mnk} \to hx} + \tau^2_{h_{n_r n_t k} \to hx}}, \tag{60}$$

$$\nu_{hx \to x_{n_r n_t k}} = \frac{\nu_{b_{n_r n_t k} \to hx}}{\nu_{h_{n_r n_t k} \to hx} + \tau^2_{h_{n_r n_t k} \to hx}}. \tag{61}$$

6) calculate $\mu_{x_{n_t k} \to hx}$

$\mu_{x_{n_tk}\to hx}$ is the message from the variable node x_{n_tk} to the factor node f_{hx}, which is provided by the soft mapper as

$$\mu_{x_{n_tk}\to hx} \approx \mathcal{CN}\left(x; \tau_{x_{n_tk}\to hx}, \nu_{x_{n_tk}\to hx}\right), \tag{62}$$

with

$$\tau_{x_{n_tk}\to hx} = \sum_{\phi\in\Phi} \phi P\left(\phi | l_{kn_t1}^{(m)}, l_{kn_t2}^{(m)}, \cdots\right), \tag{63}$$

$$\nu_{x_{n_tk}\to hx} = \sum_{\phi\in\Phi} |\phi|^2 P\left(\phi | l_{kn_t1}^{(m)}, l_{kn_t2}^{(m)}, \cdots\right) - \left|\tau_{x_{n_tk}\to hx}\right|^2. \tag{64}$$

2 Complexity

In this section, the computational complexity of the AMP receiver in one iteration between the softer mapper and demapper is investigated. As all messages are calculated based on scalars, there is no matrix operation. For the message $\mu_{y_t\to b_{n_rn_tk}}$, it requires to calculate $\tau_{y_{n_rt}\to b_{n_rn_tk}}$ and $\nu_{y_{n_rt}\to b_{n_rn_tk}}$. Following the steps in Section 1-1), first it needs to calculate $\tau_{y_{n_rt\setminus b_{n_rn_tk}}}$ and $\nu_{y_{n_rt\setminus b_{n_rn_tk}}}$ using Equation(33) and Equation(34). Note that it can obtain the overall summation only once for all n_t and k with complexity $O(N_rN_tKT) + O(N_rKT)$. Then for each n_t and k, $\tau_{y_{n_rt\setminus b_{n_rn_tk}}}$ and $\nu_{y_{n_rt\setminus b_{n_rn_tk}}}$ are realized by a subtraction from the summation part with complexity $O(N_rN_tKT)$. For Equation(42) and Equation(43), the standard Gaussian function and Q function are realized by hash maps which does not require any computational operation; and the complexities for $\mathcal{R}\left(\gamma_{y_t\to b_{n_rn_tk}}\right)$ and $\mathcal{I}\left(\gamma_{y_t\to b_{n_rn_tk}}\right)$ are both $O(N_rN_tKT)$. Then, calculating $\tau_{y_{n_rt}\to b_{n_rn_tk}}$ using Equation(44) and $\nu_{y_{n_rt}\to b_{n_rn_tk}}$ using Equation(46) has the same complexity $O(N_rN_tKT)$. Therefore, the complexity to obtain $\mu_{y_t\to b_{n_rn_tk}}$ is $O(N_rN_tKT) + O(N_rKT)$.

For the message $\mu_{b_{n_rn_tk}\to hx}$, the complexity for $\nu_{b_{n_rn_tk}\to hx}$ using Equation(48) and $\tau_{b_{n_rn_tk}\to hx}$ using Equation(49) is $O(N_rN_rKT)+O(N_rN_tK)$. For the message $\mu_{hx\to b_{n_rn_tk}}$, the complexity for $\tau_{hx\to b_{n_rn_tk}}$ using Equation(51) and $\nu_{hx\to b_{n_rn_tk}}$ using (51) is $O(N_rN_tK)$. For the message $\mu_{hx\to x_{n_tk}}$, the complexity for $\mu_{hx\to x_{n_tk}}$ using Equation(60) and $\nu_{hx\to x_{n_tk}}$ using Equation(61) is $O(N_rN_tK)$. Overall, the complexity for the designed AMP detector is only $O(N_rN_tKT)+O(N_rKT)+O(N_rT)+O(N_rN_tK)$, which is linear with N_r and N_t. Therefore, compared with the turbo LMMSE receiver, the AMP receiver is advantageous in computational complexity for large-scale systems.

E Simulation results

TABLE 2

Parameter Settings for the Simulations

ADC resolution	$3/4/8$ bits
Pilot pattern	Random 16QAM
Number of training slots (TS)	1
Constellation alphabet	16QAM/QPSK
Number of information bits	324
Number of subcarriers	18
Number of data OFDM symbols	27
Number of users	3
Number of receiver antennas	12
Interleaver	Block
LDPC code	$H_{648,1/2,27}$
Number of iterations	5
Number of trials	300

To compare with the conventional receiver structure, define the signal-to-noise ratio (SNR) at each subcarrier as $E\left[|h|^2\right]E\left[|x|^2\right]/K\sigma^{\mathrm{w}}$ when $T=K$. During one training slot (TS), each user sends one pilot OFDM symbol to the receiver with time-division to avoid interference between the pilot symbols. Apparently, the receiver is not able to obtain the accurate CSI only from one TS because of the low-resolution ADCs. While increasing the number of TSs will help the receiver to better estimate the channel, it also means increased power consumption and overhead for the pilot transmission.

In the simulation setting, only one TS is used for the channel estimation and the pilot pattern is randomly selected 16QAM symbols. Adopt LDPC as the channel code because it shares the same parallel decoding process as the proposed algorithms. Use the LDPC code $H_{648,1/2,27}$ defined in IEEE 802.11 which is a $1/2$ rate code. The number of information bits for each user is 324, and after the $1/2$ code, there are 648 effective bits. Skip cyclic redundancy check (CRC) and assume there is no other duplication or redundancy added after the channel encoder. Therefore, 648 bits are directly mapped by 16QAM and modulated using OFDM with 18 subcarriers. At the receiver, 27 OFDM symbols (from 3 users) are decoded simultaneously. Overall, 972 information bits are the output from the physical-layer decoding process. Considering delay and complexity in practical applications, limit the maximum number of the iterations to 5. A total of 300 independent trials are conducted for each setting to average the numerical results. The system parameters are summarized in Table 2.

Figure8 shows the bit error rate (BER) performance of the 8-bit, 4-bit and 3-bit ADC receivers with 16QAM. As expected, the BER decreases with the resolution for all receivers and the turbo receiver structure can significantly improve the performance. More specifically, for the conventional LMMSE receiver, there are approximately 1.5 dB and 0.5 dB performance loss when the resolution changes from 8-bit to 4-bit and from 4-bit to 3-bit, respectively. For the turbo LMMSE receiver, it brings about 0.5dB performance

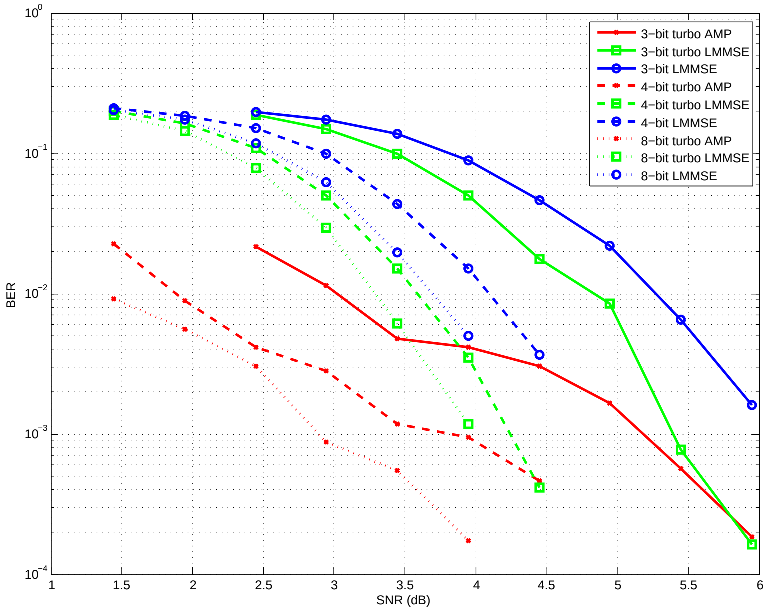


Figure 8: BER performance of the turbo receiver,16QAM

gain over the conventional LMMSE receiver. Generally speaking, the turbo AMP decoder has a better performance with up to 2.5dB gain over the conventional LMMSE receiver. In the high SNR region, the turbo LMMSE decoder has a better performance due to the more accurate CSI estimation. Figure 9 shows a similar BER performance with QPSK. The turbo LMMSE receiver and turbo AMP receiver bring approximately 0.5dB and 1.5dB gain over the conventional LMMSE receiver, respectively. It is worth noting that: 1) The performance of the turbo receiver depends on the accurate CSI estimation. In particular, the signal reconstruction in Equation(18) and the symbol estimation in Equation(25) require an accurate $\hat{\mathbf{H}}$. In the simulation setting, only one training slot is allocated for the channel estimation so that the CSI is highly inaccurate because of the low-resolution ADCs. 2) Compared to the turbo LMMSE receiver, the turbo AMP receiver is less sensitive to the CSI estimation. Because the AMP detector also takes the variance of the estimated CSI into consideration. 3) The turbo LMMSE receiver uses the extrinsic information from the

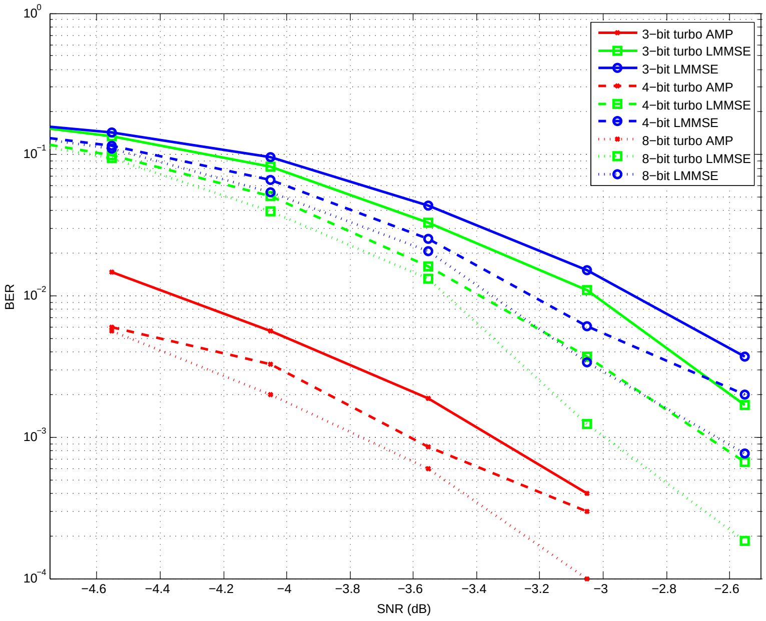


Figure 9: BER performance of the turbo receiver,QPSK

channel decoder only for de-quantization. On the other hand, the turbo AMP receiver also utilizes this extrinsic information for decoupling the multi-user interference.

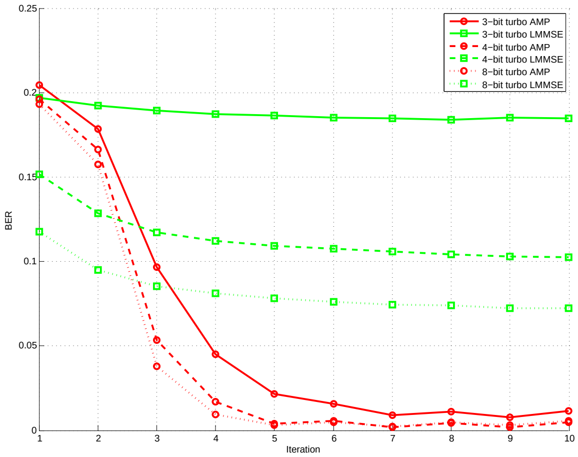


Figure 10: Convergence speed of the iterative decoding,16QAM,2.45dB

Another important factor for a turbo receiver is the convergence speed, which is closely related to the computational complexity. Figure10 and Figure11 show the convergence speeds of the turbo LMMSE and AMP receivers for 16QAM (SNR=2.45dB)

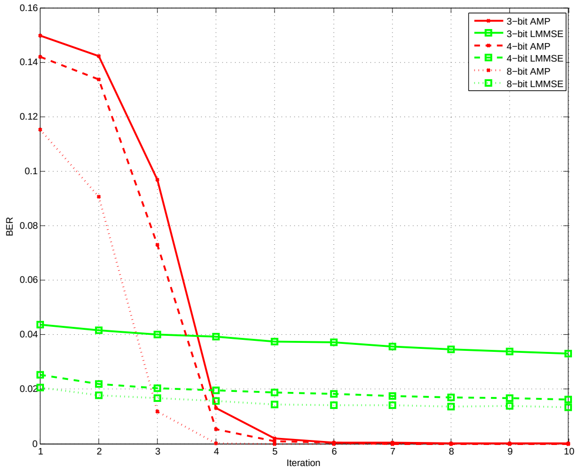


Figure 11: Convergence speed of the iterative decoding,QPSK,-3.55dB

and QPSK (SNR=-3.55dB), respectively. The turbo AMP decoder initially has a high BER because it initializes all estimated symbols as zero. Nevertheless, it only takes about 5-6 iterations for both the 16QAM and QPSK cases to converge. On the other hand, the turbo LMMSE receiver has slow but smooth convergence curves.

F Conclusion

In this Chapter, a new turbo receiver structure was proposed for the uplink MU-OFDM-MIMO systems with low-resolution ADCs. The interdependency between the input analog signal and the quantization error was exploited to realize an iterative turbo structure. Two specific turbo receivers were designed based on LMMSE and AMP, and the computational complexity was analyzed. The simulation results demonstrated that the turbo structure can effectively improve the performance in quantized MU-OFDM-MIMO systems.

CHAPTER III

ONE-BIT TRANSCEIVER CLUSTER FOR RELAY TRANSMISSION

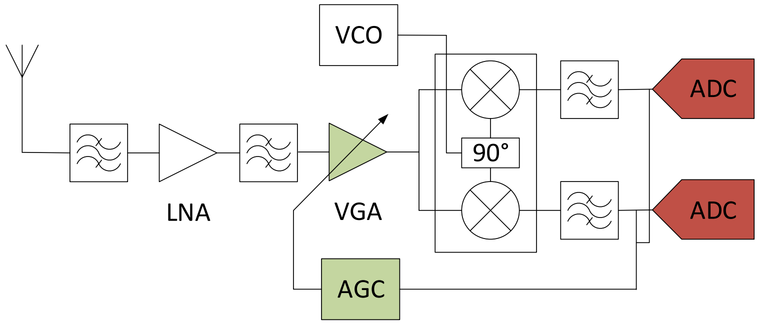


Figure 12: Direct-conversion receiver architecture

Figure 12 shows the architecture of conventional direct-conversion receivers, where a fundamental component is the analog to digital converter (ADC), converting the received signal into digital format with a typical precision of 8-12 bits. For example, a complete direct-conversion receiver RF front end in [45] consists of a frequency synthesizer, a quadrature demodulator (including a Variable Gain Amplifier (VGA) and an Automatic Gain Control (AGC)) and a 10-Bit ADC. The typical power dissipation in [45] for each IC is 40mW, 340mW and 565mW, respectively, which indicates the majority of the power is consumed in ADCs [1]. Therefore, it has been proposed to use comparators (one-bit ADCs) to replace the high resolution ADCs [4, 5, 12, 13, 16, 46], and the power consumption can be reduced by up to one order of magnitude. Meanwhile, since only two different phases are recognized after the comparator, the VGA and AGC become unnecessary, which further reduces the cost and the power dissipation. Finally, the one-bit comparator also reduces the performance requirement of power amplifiers for further efficiency improvement. The drawback of using the one-bit transceiver is obvious: the extreme quantization will cause information loss and failure of traditional DSP techniques. Thus, the key of one-bit transceiver design is to develop a feasible baseband to handle the one-bit quantization. In the literature, use of the one-bit ADCs has only been recently studied in MIMO systems [4, 5, 8, 12, 13, 16, 46].

In this chapter, the feasibility of deploying a one-bit relay cluster for a distant transmission is explored. The envisioned system model is shown in Figure 13, where there are one source node, one destination node and multiple one-bit relays, all equipped with single-antenna. It must be emphasized that the one-bit relays transmit simultaneously at the same frequency band, which completely eliminates the traditional multiple access requirement and thus greatly simplifies the transmission protocol (especially for a large number of relays). However, to make it work, the challenges are multi-fold: Unlike the one-bit MIMO systems in [8, 12–14, 16], in this case the destination node cannot separate the superimposed signal with single-antenna. Meanwhile, one-bit ADCs also impede high-order amplitude modulation schemes because the amplitude information will be totally lost at the destination. Furthermore, the equivalent channel between the source and

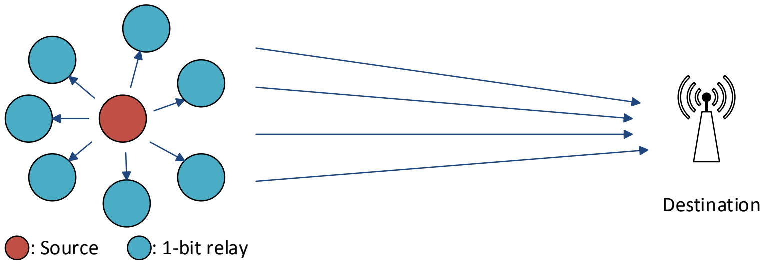


Figure 13: One-bit relays cluster transmission.

the destination is nonlinear so that traditional estimation techniques are not applicable any more.

A System model

Assume a power-restricted source with single-antenna needs to reach a distant destination. The signal strength at the destination is too low for detection. As shown in Figure 2, a number of one-bit communication nodes scattered around the source can help its transmission as intermediate relays. For low-cost reasons, these one-bit relays are equipped with single-antenna and the full connection is realized in physical layer (PHY) without any complex protocol. The only responsibility of these relays is to receive-store-transmit in a specific time period. All the relays work independently without any message exchange among themselves or control from the destination. Therefore, the information of each the relay is not necessarily required by the destination, and the existence of a specific relay does not need to be acknowledged. This highly increases the flexibility of the entire network. Note that the source node will periodically broadcast a preset synchronization signal to realize both time and symbol synchronization among the relays.

The Pseudo Noise (PN) sequence is well suited for the one-bit relay synchronization. Specifically, one PN sequence is transmitted only using I path (equivalent to BPSK). Since the PN sequence is binary, a one-bit relay can detect the sequence directly and synchronize its own transmission.

1 Random-static DC-offset for the M-QAM modulation

Using the one-bit relays, phase information of the source signal is preserved by the relay cluster, but amplitude levels are not distinguishable. To see that, let's see an example with three transmitted baseband signals: $signal_1 = 1$, $signal_2 = 2$ and $signal_3 = \exp(\pi/3j)$. There are 30 valid one-bit relays. The channel realizations are sampled based on the complex Gaussian distribution, and the SNR at each relay is $20dB$. In Figure 14, $signal_1$ and $signal_2$ have exactly the same histogram (the number of the relays with received signal 1 and -1), implying these two signals can not be distinguished by the relay cluster. On the contrary, the different histograms for $signal_1$ and $signal_3$ indicate that the phase difference is acknowledged by the cluster. Therefore, the M-Phase Shift Keying (M-PSK) is naturally supported by the one-bit relay cluster. To further support amplitude modulations such as the M-Quadrature Amplitude Modulation (M-QAM), a new receiver architecture is proposed with a random-static DC-offset (i.e., the value of this DC-offset is constant based on some probability density function).

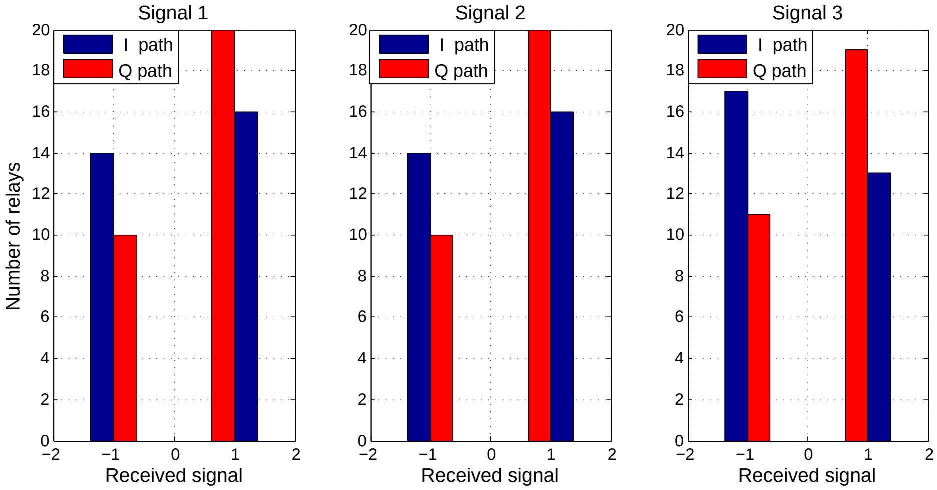


Figure 14: Histogram of the relays on received signals: Rayleigh fading channel and SNR=20dB, number of relays =30

Specifically, DC-offsets are added at each the relay before the comparator. This can be simply realized by coupling the local oscillator (LO) with the RF signal before the mixer, as showed in Figure 15. This artificial interference provides the needed diversity across the relays to distinguish amplitudes at the destination. Figure 16 shows the histograms of the relays with the DC-offsets. Compared to Figure 14, $signal_1$ and $signal_2$ yield the different histograms that carry the amplitude information. Another advantage of the random-static DC-offset solution is that it evolves naturally from the direct-conversion receiver architecture. Specifically, it is well known that a direct-conversion Rx suffers

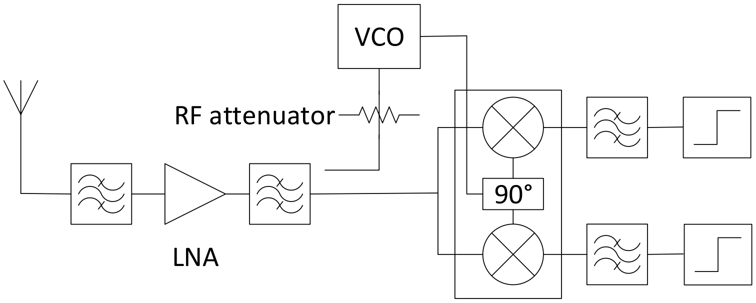


Figure 15: Relay Rx design with the LO coupling.

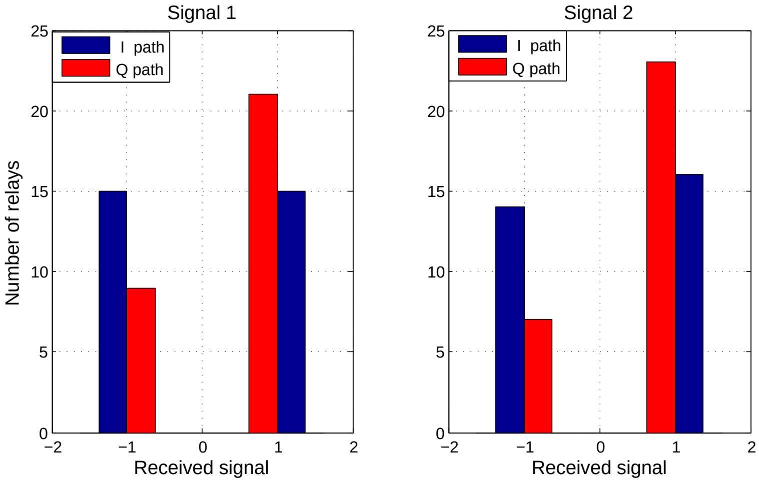


Figure 16: Histogram of the relays on received signal with the DC-offsets: Rayleigh fading channel and SNR=20dB, number of relays =30

from the LO leakage: the LO may be conducted or radiated through an unintended path to the mixer's input port. As a by-product, the LO signal effectively mixes with itself, producing a DC component at the mixer output.

In particular, due to manufacturing process and environment, the amount of LO leakage varies across the relays, which naturally results in a random-static DC-offset at baseband.

2 Black-box solution for the one-bit relays cluster system

For the single-input-single-output (SISO) system shown in Figure 2, assume flat uncorrelated Rayleigh slow fading channels. Denote the number of the relays in a cluster as n, the source symbol as $x_{1\times 1}$, the channel matrices as $\mathbf{h}_{1,n\times 1}$ (from the source to the relays) and $\mathbf{h}_{2,1\times n}$ (from the relays to the destination), the DC-offset as $\mathbf{d}_{1\times n}$ (constant) and the circularly symmetric complex Gaussian noise as $\mathbf{w}_{1,1\times n}$, $w_{2,1\times 1}$. The received symbol $y_{1\times 1}$ at the destination node is given by:

$$y = \mathbf{h}_2 f_q\left(\mathbf{h}_1 x + \mathbf{d} + \mathbf{w}_1\right) + w_2, \tag{65}$$

where $f_q(\cdot)$ is the one-bit quantization function. For massive relay deployment, n can be very large. Apparently, it is impractical to estimate $\mathbf{h}_1$ and $\mathbf{h}_2$. Estimating $\mathbf{h}_1$ and $\mathbf{h}_2$ will cost unscalable overhead and power. Therefore, to solve this problem, instead of using the traditional system model in (1), the destination node will treat the equivalent channel $y = \hat{f}^{-1}\left(x\right)$ between the source and the destination as a black box shown in Figure 17, and try to find out the relationship between x and y through training data. With the black-box assumption, decoding in the destination node is independent of the number of the relays.

In other words, increasing the number of the relays doesn't increase the system complexity, and failure of multiple the relays has little impact on decoding process[1].

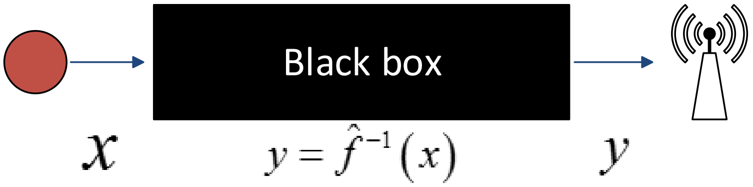


Figure 17: Black box view of the system.

B Estimator design for one-bit relays decoding

At the destination, the key is to estimate x based on y using the function $\hat{x} = \hat{f}\,(y)$, where parameters of $\hat{f}$ are determined by minimizing the mean square error (MSE) of the pilot samples $\left((\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2), \cdots, (\bar{x}_n, \bar{y}_n)\right)$ as

$$\min_{\hat{f}} \sum_i \left|\hat{\bar{x}}_i - \bar{x}_i\right|^2. \tag{66}$$

Consider the following three estimators.

1 Linear estimator (LE)

The traditional estimator is linear based:

$$\hat{x} = ay + b. \tag{67}$$

Plug (3) in (2). The optimization problem is quadratic and it can obtain the optimal solution as

$$a = \left(\sum_i |\tilde{\bar{y}}_i|^2\right)^{-1} \sum_i \tilde{\bar{y}}_i{}^* \tilde{\bar{x}}_i, b = \frac{1}{n}\sum_i \bar{x}_i - \frac{a}{n}\sum_i \bar{y}_i, \tag{68}$$

[1] The destination node will periodically re-estimate the black-box system, the failure of multiple relays only exhibits as a fluctuation of the equivalent channel.

where $\tilde{\bar{y}}_i = \bar{y}_i - \frac{1}{n}\sum\limits_i \bar{y}_i$ and $\tilde{\bar{x}}_i = \bar{x}_i - \frac{1}{n}\sum\limits_i \bar{x}_i$.

2 Support vector machine (SVM)

Linear parametric models can be re-cast into an equivalent ‘dual representation’ in which the estimations are also based on linear combinations of a kernel function evaluated at the training data points. For models based on a fixed nonlinear feature space mapping $\phi\left(y\right)$, the kernel function is given by $\kappa\left(y, y'\right) = \phi(y)^*\phi\left(y'\right)$. For example, the Gaussian kernel is given by $\kappa\left(y, y'\right) = \exp\left(-\left(y - y'\right)^2/2\sigma^2\right)$. By utilizing the nonlinear feature mapping, the black-box system can be modeled as $\hat{x} = \hat{f}\left(y\right) = w\phi\left(y\right)$, where w is the undetermined coefficient. The Representer theorem shows that it can adopt $w = \mathbf{a}^H\phi\left(\bar{\mathbf{y}}\right) = \sum\limits_i a_i\phi\left(\bar{y}_i\right)$, where a_i is the model parameter. Therefore, the nonlinear estimator based on the training samples is given by

$$\hat{x} = \mathbf{a}^H\phi\left(\bar{\mathbf{y}}\right)\phi\left(y\right). \tag{69}$$

Plug (5) into (2), it can solve the quadratic optimization problem with a analytical solution as

$$\mathbf{a} = \left(\phi\left(\bar{\mathbf{y}}\right)\phi(\bar{\mathbf{y}})^H\right)^{-1}\bar{\mathbf{x}}. \tag{70}$$

Thus, the nonlinear estimator is given by

$$\begin{aligned}\hat{x} &= \bar{\mathbf{x}}^H\left(\phi\left(\bar{\mathbf{y}}\right)\phi(\bar{\mathbf{y}})^H\right)^{-1}\phi\left(\bar{\mathbf{y}}\right)\phi\left(y\right) \\ &= \bar{\mathbf{x}}^H K^{-1}\left(\bar{\mathbf{y}}, \bar{\mathbf{y}}\right)K\left(\bar{\mathbf{y}}, y\right),\end{aligned} \tag{71}$$

where $K\left(\mathbf{y}, \mathbf{y}'\right) = \begin{pmatrix} \kappa\left(y_1, {y_1}'\right) & \cdots & \kappa\left(y_1, {y_n}'\right) \\ \vdots & \ddots & \vdots \\ \kappa\left(y_n, {y_1}'\right) & \cdots & \kappa\left(y_n, {y_n}'\right) \end{pmatrix}$.

3 Neural network (NN)

In the SVM, every single estimation requires re-computing of all samples because of $K(\bar{\mathbf{y}}, y)$. To make the model more compact, it can use other nonlinear models such as the neural network. Adopt a 2-layer neural network given as

$$\hat{x} = \mathbf{a}_2^H f_a \left(\mathbf{a}_1^H y + \mathbf{b}_1 \right) + b_2, \tag{72}$$

where $f_a(\cdot)$ is the nonlinear activation function (e.g., the sigmoid function $f_a(y) = (1 + \exp(-y))^{-1}$); $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{b}_1$ and b_2 are the model parameters; and the dimension of these parameters can be arbitrarily chosen by investigating the estimation performance. Plug (8) into (2), it gives

$$\min_{\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}_1, b_2} \sum_i \left| \mathbf{a}_2^H f_a \left(\mathbf{a}_1^H \bar{y}_i + \mathbf{b}_1 \right) + b_2 - \bar{x}_i \right|^2. \tag{73}$$

Equation(9) is not a convex optimization problem and there is no analytical solution. It can use the Newton algorithm to find a local optimal point and the parameters are updated by:

$$\left[\mathbf{a}_1^{(\beta+1)}; \mathbf{a}_2^{(\beta+1)}; \mathbf{b}_1^{(\beta+1)}; b_2^{(\beta+1)} \right] = \left[\mathbf{a}_1^{(\beta)}; \mathbf{a}_2^{(\beta)}; \mathbf{b}_1^{(\beta)}; b_2^{(\beta)} \right] - \Delta$$

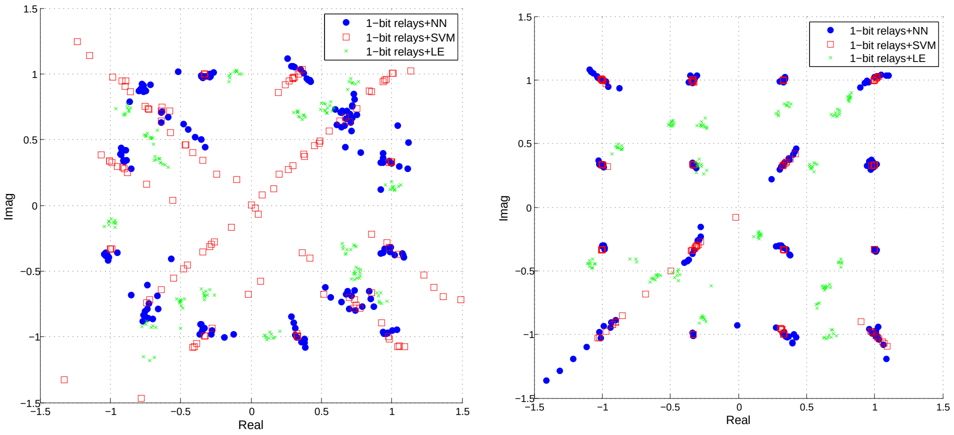
where $\Delta = \mathbf{H}^{-1}\mathbf{J}\mathbf{e}$, $\mathbf{e} = \left(\hat{\bar{x}}_1 - \bar{x}_1, \hat{\bar{x}}_2 - \bar{x}_2, \cdots, \hat{\bar{x}}_n - \bar{x}_n \right)^T$, $\mathbf{H}$ and $\mathbf{J}$ respectively are the Hessian and the Jacobian matrix of $\mathbf{e}$. To reduce the complexity of calculating the Hessian matrix, adopt the Levenberg-Marquardt algorithm that approximates the Hessian matrix using the Jacobian matrix as

$$\mathbf{H} \approx \left(\mathbf{J}^H \mathbf{J} + \lambda \mathrm{diag} \left(\mathbf{J}^H \mathbf{J} \right) \right) \tag{74}$$

where λ is the damping parameter which can be arbitrary chosen by investigating the estimation performance.

C Numerical Results

The simulation has the following setup: the modulations are 16-QAM and 16-PSK respectively; the number of relays is 30; the number of pilot symbols is 128; the number of data symbols is 1280; Since all relays are scattered near the source nodes, assume the SNR of these relays $E\left[|h_1|^2|x|^2/|w_1|^2\right]$ is 20dB. The DC-offset $\mathbf{d}$ is sampled from $\mathcal{N}(0,1)$. The kernel for the SVM is the Gaussian kernel; the number of hidden nodes for the NN is 10. Take 200 independent trials of the Rayleigh fading channel $\mathbf{h}_1$, $\mathbf{h}_2$ and the DC-offset $\mathbf{d}$. To make the results comparable, assume that the transmit power of a single relay is identical to that of the source node. First of all, Figure 18 compares the 16-QAM decoding constellations without and with the DC-offsets. In Figure 7(a), without the DC-offset, the pairs $(1+j, 1/3+j1/3)$, $(-1+j, -1/3+j1/3)$, $(1-j, 1/3-j1/3)$ and $(-1-j, -1/3-j1/3)$ are not distinguishable by any the estimator. This problem is caused by the one-bit relays rather than the estimator design. With the proposed DC-offset solution, the entire 16-QAM constellation can be recovered as showed in Figure 7(b), which proves the effectiveness of the DC-offset solution for the M-QAM. Next, compare the performance of the direct-link transmission (w/o relays) and the one-bit relay transmission with the three estimators. Since the direct-link channel is linear, the linear estimator is sufficient for decoding (i.e., the other two estimators cannot provider additional gains). After the symbol estimator, use the hard detection to get the bit error rate (BER) directly. Figure 19 shows the performance of the estimators based on SNR ($E\left[|h_2|^2/|w_2|^2\right]$). From Figure 19, it gives the following results: (1) The performance of the one-bit relay with the LE stays the same in all the conditions as expected. This is



(a) without the DC-offsets (b) with the DC-offsets

Figure 18: 16-QAM decoding constellation.

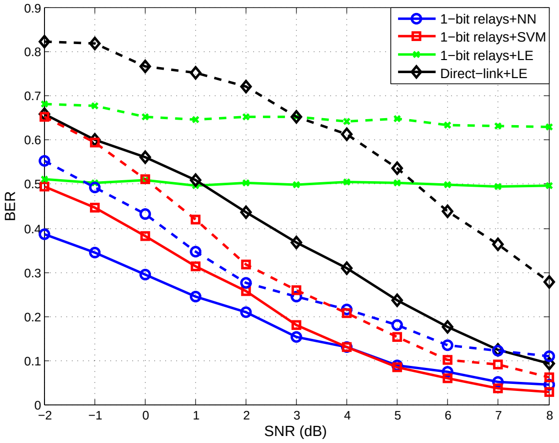


Figure 19: Performance of the nonlinear estimators. (Solid line: 16-QAM, Dash line: 16-PSK)

because that the equivalent channel itself is nonlinear and the LE is not able to handle this channel at all. The fluctuation appearing in the LE performance curve is caused by the finite set of test data. (2) Using the designed nonlinear estimators, 30 one-bit relays can approximately reduce the BER to a half, comparing to the direct-link without the relays. (2) In the low SNR region (SNR<4dB), the NN has a lower BER than the SVM. However, as the SNR increases, the performance of the SVM exceeds that of the NN.

D Conclusion

In this Chapter, the feasibility to massively deploy a one-bit relay cluster for a distant transmission was discussed. The simulation results verified that the solution is practical and effective for the one-bit relay transmission.

CHAPTER IV

MIMO INTERFERENCE CANCELLATION FOR LTE AND WIFI COEXISTENCE

LTE-U is a radio access technology that has been proposed for providing carrier-grade wireless service in the 5GHz unlicensed band. Until today, WiFi (WLAN that uses the IEEE 802.11 standard) has been the most popular choice for radio access in the unlicensed space. However, recent studies have highlighted that LTE technology, originally envisioned for cellular operation in licensed bands, has significant performance gains over Wi-Fi when operating in the unlicensed band. The main advantages for LTE-U over WiFi include better link performance, medium access control, mobility management, and excellent coverage. These benefits combined with the vast amount of available spectrum (>400MHz) in the 5GHz band make LTE-U a promising radio access technology in the unlicensed arena.

Since WiFi devices are already widespread in the 5GHz unlicensed band, there is a need for newly deployed LTE-U Small Cell (SC) to coexist with the WiFi ecosystem. Moreover, different LTE-U operators may occupy the same spectrum in the unlicensed band to provide data services to their users. Such an unplanned and unmanaged deployment of LTE-U SCs (femtocells, picocells) may result in excessive RF interference to the existing co-channel WiFi and other operator LTE-U nodes in the vicinity. It is therefore critical for

LTE-U SCs to choose the best operating channel while minimizing the interference caused to nearby WiFi and LTE-U networks. However, there are scenarios where all available channels are occupied by WiFi devices which forces LTE-U SC to operate on the same channel as WiFi. WiFi devices do not back off to LTE-U unless its interference level is above the energy detection threshold (-62dBm over 20MHz). Without proper coexistence mechanisms, LTE-U transmissions could cause considerable interference on WiFi network relative to WiFi transmissions.

The LTE-U Forum[1] was created by Verizon, in conjunction with Alcatel-Lucent, Ericsson, Qualcomm, and Samsung as members. The forum collaborates and creates technical specifications for base stations and consumer devices passing LTE-U on the unlicensed 5 GHz band, as well as coexistence specs to handle traffic contention with existing WiFi devices. In LTE-U SDL Coexistence Specifications V1.3 (2015-10) proposed by the LTE-U Forum, it requires the base station to be able to create an ON/OFF time pattern on the cell using the carrier sensing adaptive transmission (CSAT) procedure when other WiFi or other operator LTE-U co-channel nodes are sensed with energy level above -62 dBm. In ON-state, the cell is transmitting according to 3GPP LTE Rel-10 or later releases specification or the cell is transmitting LTE-U Discovery Signal. In OFF-state: the cell ceases all transmissions, including sync signal, SI signals, CRS, and etc. The CSAT duty cycle can change over time, for instance based on channel usage. With the exception of periodic transmissions for the MIB and LTE-U Discovery Signal (LDS), the base station shall put the cell in OFF-state when the cell is not needed such as no user in the cell coverage or there is no data in buffer for users in the cell coverage.

[1] http://www.lteuforum.org/

While the ON/OFF time pattern of LTE-U can release the spectrum periodically, In ON-state, WiFi devices will be fully blocked leading to considerable performance instability. Therefore, to support the LTE and WiFi coexistence, MIMO interference cancelling receivers are designed and prototyped in this work to enable the LTE and WiFi devices to work effectively even in ON-state of the cell.

A 802.11 and LTE PHY specifications

1 802.11 PHY

IEEE 802.11 is a set of medium access control (MAC) and physical layer (PHY) specifications for implementing Wireless Local Area Network (WLAN) communication using ISM band. The 802.11 PHY uses burst transmissions. The 802.11 standards define frame types for use in transmission of data as well as management and control of wireless links. In MAC layer, these frames are divided into three functions. Management frames[2] allow for the maintenance of communication. Control frames[3] facilitate in the exchange of data frames. Data frames carry higher-level protocol data in the frame body.

All the frames in MAC layer are transmitted using the same packet structure in PHY. Each packet contains a preamble and payload data. The preamble allows the receiver to obtain time and frequency synchronization and estimate channel characteristics for equalization. It is a sequence that receivers watch for to lock onto the rest of the transmission. For the legacy preamble, it is composed of three fields as shown in Figure

[2] including: Authentication Frame, Association Request Frame, Association Response Frame, Beacon Frame, Deauthentication Frame, Disassociation Frame, Probe Request Frame, Probe Response Frame, Reassociation Request Frame, Reassociation Response Frame.

[3] including: Acknowledgement (ACK) Frame, Request to Send (RTS) Frame, Clear to Send (CTS) Frame.

20.

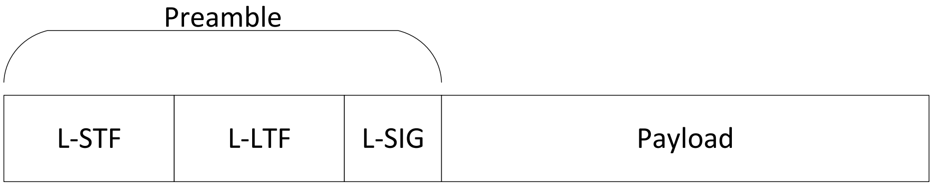


Figure 20: 802.11 packet structure.

The legacy short training field (L-STF) is the first field of the 802.11 OFDM legacy preamble. The L-LTF is composed of two identical OFDM symbols with a 1/4 CP each. The sequence uses 12 of the 52 subcarriers (every fourth) that are available per OFDM symbol. Because the sequence has good correlation properties, it is used for start-of-packet detection, for coarse frequency correction, and for setting the AGC. The L-STF duration is $8\mu s$. The legacy long training field (L-LTF) is the second field. Channel estimation, fine frequency offset estimation, and fine symbol timing offset estimation rely on the L-LTF. The L-LTF is composed of a CP followed by two identical OFDM symbols occupying all 52 subcarriers. The CP consists of the second half of the OFDM symbol. The L-LTF duration is also $8\mu s$. The legacy signal (L-SIG) field is the third field. The L-SIG is one OFDM symbol with its CP. It consists of 24 bits that contain rate, length, and parity information. It is transmitted using BPSK modulation with rate 1/2 binary convolutional coding (BCC). The L-SIG duration is $4\mu s$.

2 LTE PHY

Long Term Evolution (LTE) is a series of standards that define the entire network structure for implementing cellular network communication. LTE PHY operates on the licensed bands with dedicated resources. The LTE PHY transmission is fully synchronized and controlled by Base Station (BS).

There are five time units defined in LTE MAC and PHY: frame ($10ms$), half-frame ($5ms$), subframe ($5ms$), slot ($1ms$), symbol. A resource block (RB) is the smallest unit of resources that can be allocated to a user. The resource block is 180 kHz wide in frequency and 1 slot long in time. In frequency, resource blocks are 12 x 15 kHz subcarriers wide. The bandwidths defined by the LTE standard are 1.4, 3, 5, 10, 15, and 20 MHz. For full-duplex FDD, uplink and downlink frames are separated by frequency and are transmitted continuously and synchronously. Figure 21 show an example of one FDD downlink frame structure[4].

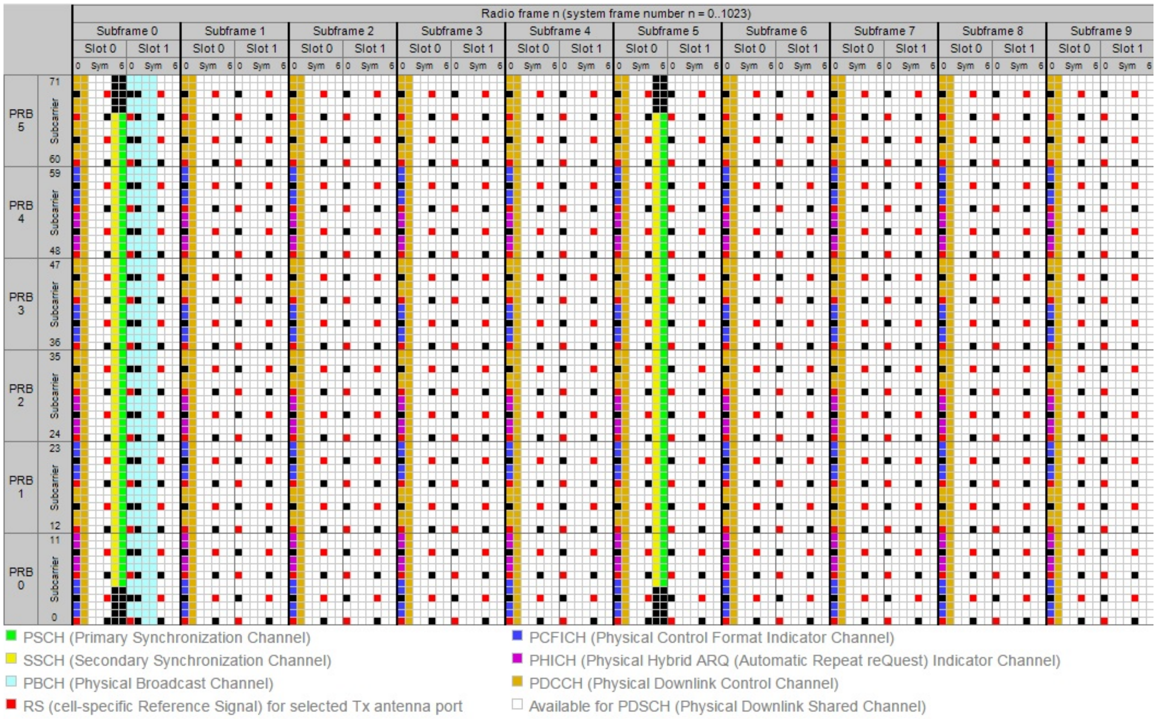

Radio frame n (system frame number n = 0..1023)
Subframe 0
Subframe 1
Subframe 2
Subframe 3
Subframe 4
Subframe 5
Subframe 6
Subframe 7
Subframe 8
Subframe 9
Slot 0
Slot 1
0 Sym 6
PRB 5
PRB 4
PRB 3
PRB 2
PRB 1
PRB 0
Subcarrier
71
60
59
48
47
36
35
24
23
12
11
0
🟩 PSCH (Primary Synchronization Channel)
🟨 SSCH (Secondary Synchronization Channel)
PBCH (Physical Broadcast Channel)
🟥 RS (cell-specific Reference Signal) for selected Tx antenna port
🟦 PCFICH (Physical Control Format Indicator Channel)
🟪 PHICH (Physical Hybrid ARQ (Automatic Repeat reQuest) Indicator Channel)
🟧 PDCCH (Physical Downlink Control Channel)
□ Available for PDSCH (Physical Downlink Shared Channel)

Figure 21: LTE resource grid, 1.4MHz FDD downlink.

LTE physical channels correspond to sets of time-frequency resources used for

[4] http://niviuk.free.fr/lte_resource_grid.html

transmission of particular transport channel data, control information, or indicator information. The PHY Downlink Shared Channel (PDSCH) carries user data and paging information to the terminal. The PHY Downlink Control Channel (PDCCH) conveys control information, scheduling decisions for PDSCH reception, and for scheduling grants enabling transmission on the PHY Uplink Shared Channel (PUSCH). The PHY BroadCast Channel (PBCH) carries part of the system information required for terminals to access the network. The PHY Hybrid-ARQ Indicator Channel (PHICH) conveys the hybrid-ARQ acknowledgement indicating to the terminal whether or not to retransmission of a transport block is required. The PHY Control Format Indicator Channel (PCFICH) provides terminals with information to decode the set of PDCCHs.

LTE physical signals are used for synchronization and channel estimation. The cell-specific reference signal (CRS) is transmitted on resource elements spread throughout the frame in specific locations as defined by the standard. The CRS are transmitted in every downlink subframe and in every resource block in the frequency domain, thus covering the entire cell bandwidth. The CRS can be used by user equipments for channel estimation and coherent demodulation of any downlink physical channel. The primary synchronization signal (PSS) are used by users for cell search and time synchronization. The secondary synchronization signal (SSS) are used to obtain the start of a frame. For FDD, the PSS is presented in the last symbol, and the SSS is presented in the second-to-last symbol of slots 0 and 10 in every frame. The PSS is mapped into the first 31 subcarriers either side of the DC subcarrier. Therefore, the PSS uses six resource blocks with five reserved subcarriers each side.

B Time and frequency synchronization

Time and frequency synchronization at receivers is required by all OFDM based communication systems.

1 802.11 synchronization

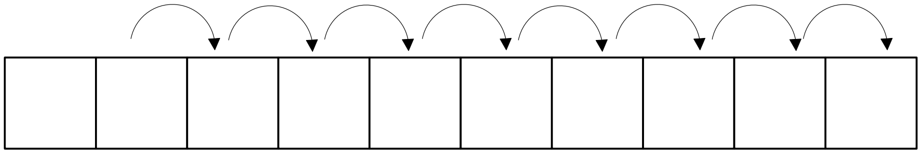


Figure 22: L-STF auto-correlation.

Time and frequency synchronization in 802.11 utilizes the structure of the L-STF and L-LTF. In time domain, the L-STF is a 16-points (samples) signal repeated 10 times. Conduct auto-correlation based on the L-STF signal as demonstrated by Figure 22:

$$P\left(d\right)=\sum_{n=1}^{8}\sum_{m=0}^{15}\left(r^{*}_{d+m+n\times 16}r_{d+m+16+n\times 16}\right), \tag{75}$$

$$R\left(d\right)=\sum_{n=1}^{8}\sum_{m=0}^{15}\left|r_{d+m+n\times 16}\right|\left|r_{d+m+16+n\times 16}\right|, \tag{76}$$

$$M\left(d\right)=\frac{\left|P\left(d\right)\right|^{2}}{R^{2}\left(d\right)}, \tag{77}$$

where r_m is the received discrete signal at index m. The auto-correlation method can be implemented with low-complex iterative formula as

$$P\left(d+1\right)=P\left(d\right)-\left(r^{*}_{d+16}r_{d+32}\right)+\left(r^{*}_{d+144}r_{d+160}\right), \tag{78}$$

$$R\left(d+1\right)=P\left(d\right)-\left|r_{d+16}\right|\left|r_{d+32}\right|+\left|r_{d+144}\right|\left|r_{d+160}\right|. \tag{79}$$

With a threshold λ, the packet start point is estimated as

$$\hat{d} = \arg\left(M\left(d\right) > \lambda\right). \tag{80}$$

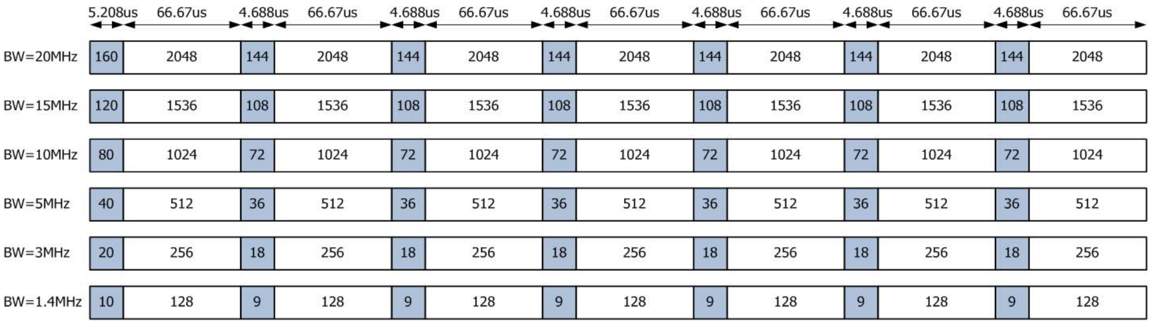
The advantage of the auto-correlation method is twofold: 1) it is robust to multipath channel, the auto-correlation value is independent from frequency-selective channel. 2) its complexity is constant $O\left(1\right)$ using the iterative method. However, the auto-correlation method usually leads to a wide mainlobe width. To increase the time accuracy, the cross-correlation method can be further adopted as a complementary. The cross-correlation method is straight-forward, it calculates the correlation of a pre-known signal with the received signal. The cross-correlation value is affected by the specific channel and with a linear complexity. To take advantage of both the auto and cross correlation methods, the receivers can use the cross-correlation method on the L-STF to find a candidate start point region $\left[\hat{d}_L, \hat{d}_R\right]$. Then search for the maximal cross-correlation value on the L-LTF in $\left[\hat{d}_L, \hat{d}_R\right]$ as the estimated packet start point $\hat{d}$. This method provides an accuracy estimation while remaining the constant complexity.

After finding the frame start point $\hat{d}$, the fractional frequency (phase) offset is estimated as

$$\hat{\phi} = \text{angle}\left(P\left(\hat{d}\right)\right)/16. \tag{81}$$

2 LTE sychronization

Different from burst transmission, the major synchronization of LTE is only required when a user first time accesses the network. The frame synchronization (i.e., MAC layer synchronization) is based on the SSS. The symbol time and frequency synchronization (i.e., PHY synchronization) of LTE users rely on the CPs of OFDM symbols. The PSS can also be used to help time synchronization. The length and position of the CPs vary from different bandwidths. Figure 23 shows the CP length in a slot of different bandwidths.



	5.208us	66.67us	4.688us	66.67us	4.688us	66.67us	4.688us	66.67us	4.688us	66.67us	4.688us	66.67us	4.688us	66.67us
BW=20MHz	160	2048	144	2048	144	2048	144	2048	144	2048	144	2048	144	2048
BW=15MHz	120	1536	108	1536	108	1536	108	1536	108	1536	108	1536	108	1536
BW=10MHz	80	1024	72	1024	72	1024	72	1024	72	1024	72	1024	72	1024
BW=5MHz	40	512	36	512	36	512	36	512	36	512	36	512	36	512
BW=3MHz	20	256	18	256	18	256	18	256	18	256	18	256	18	256
BW=1.4MHz	10	128	9	128	9	128	9	128	9	128	9	128	9	128

Figure 23: CP length in a slot.

Conduct auto-correlation on the CPs in a slot as shown in Figure 24:

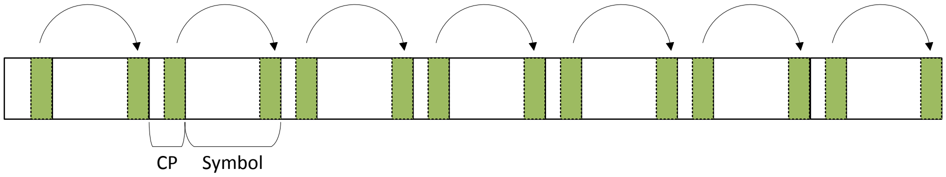


Figure 24: L-LTF auto-correlation.

$$P\left(d\right)=\sum_{n=0}^{6}\sum_{m=0}^{L_{cp}}\left(r^{*}_{d+m+n\times(L_s+L_{cp})}r_{d+m+L_s+n\times(L_s+L_{cp})}\right),\tag{82}$$

$$R\left(d\right)=\sum_{n=0}^{6}\sum_{M=0}^{L_{cp}}\left|r_{d+m+n\times(L_s+L_{cp})}\right|\left|r_{d+m+L_s+n\times(L_s+L_{cp})}\right|,\tag{83}$$

where L_{cp} is a length smaller than the length of one CP, L_s is the length of one OFDM symbol. Equation(82)(83) can also be calculated in a iterative way with a constant complexity. The symbol start point region is estimated as

$$\left[\hat{d}_L,\hat{d}_R\right]=\arg\left(\frac{\left|P\left(d\right)\right|^2}{R^2\left(d\right)}>\lambda\right).\tag{84}$$

Then conduct cross-correlation on the PSS time domain signal to find the start of the symbol $\hat{d}$. The fractional frequency (phase) offset is estimated as

$$\hat{\phi} = \text{angle}\left(P\left(\hat{d}\right)\right)/L_S. \tag{85}$$

C MIMO interference cancellation

In the LTE-U and WiFi coexistence network, the LTE (resp. WiFi) receiver needs to cancel WiFi (resp. LTE) signal to improve desired SNR. Different from the scenarios that the receiver is able to estimate all the interference channel, the LTE (resp. WiFi) receiver has no information about the WiFi (resp. LTE) system. Therefore the receiver has to implement MIMO interference canceller only with the knowledge of its own PHY structure.

1 802.11 interference canceller

For a WiFi receiver, the MIMO interference canceller is designed to minimize the square error of the data in the L-LTF.

$$\min_{\mathbf{g}_k} \left|\mathbf{g}_k\mathbf{y}_k - c_{LTF,k}\right|^2, \tag{86}$$

where $c_{LTF,k}$ is the L-LTF signal on the kth subcarrier, $\mathbf{y}_k$ is the received signal vector on the kth subcarrier from multiple receiving antennas, $\mathbf{g}_k$ is the receiving antenna combining coefficient vector on the kth subcarrier. The optimal solution is

$$\mathbf{g}_k^* = \left(\mathbf{y}_k^H\mathbf{y}_k\right)^{-1}\mathbf{y}_k^H c_{LTF,k}. \tag{87}$$

As the channels of the neighboring subcarriers are highly correlated, it can take average over neighboring $\mathbf{g}_k^*$ to combat the noise as

$$\bar{\mathbf{g}}_k = \frac{1}{\delta + 1} \sum_{|\Delta| \leq \delta} \mathbf{g}_{k+\Delta}^*, \tag{88}$$

where δ is the number of neighboring $\mathbf{g}_k^*$ used. $\bar{\mathbf{g}}_k$ will be used by the equalizer in this packet to recover the signal. However, if we assume $\mathbf{g_k}$ of the neighboring subcarriers are the same, the optimal solution should be given by

$$\min_{\mathbf{g}_k} \sum_{|\Delta| \leq \delta} \left| \mathbf{g}_k \mathbf{y}_{k+\Delta} - c_{LTF,k+\Delta} \right|^2 \tag{89}$$

$$\mathbf{g}_k^* = \left(\sum_{|\Delta| \leq \delta} \mathbf{y}_{k+\Delta}^H \mathbf{y}_{k+\Delta} \right)^{-1} \sum_{|\Delta| \leq \delta} \mathbf{y}_{k+\Delta}^H c_{LTF,k+\Delta} \tag{90}$$

2 LTE interference canceller

For a LTE receiver, the MIMO interference canceller is designed to minimize the square error of the data in the CRS.

$$\min_{\mathbf{g}_{k,t}} \left| \mathbf{g}_{k,t} \mathbf{y}_{k,t} - c_{RS,(k,t)} \right|^2, \tag{91}$$

where $c_{RS,(k,t)}$ is the CRS signal on the kth subcarrier of tth OFDM symbol, $\mathbf{y}_{k,t}$ is the received signal vector from multiple receiving antennas on the kth subcarrier of the tth OFDM symbol, $\mathbf{g}_{k,t}$ is the receiving antenna combining coefficient vector on the kth subcarrier of the tth OFDM symbol. The optimal solution is given by

$$\mathbf{g}_{k,t}^* = \left(\mathbf{y}_{k,t}^H \mathbf{y}_{k,t} \right)^{-1} \mathbf{y}_{k,t}^H c_{CR,(k,t)}, \tag{92}$$

The channels in a resource block are highly correlated. Therefore, take average of $\mathbf{g}_{k,t}^{*}$ in a resource block as

$$\bar{\mathbf{g}}_{\Omega_{k,t}} = \frac{1}{|\Omega_{k,t}|} \sum_{(k',t') \in \Omega_{k,t}} \mathbf{g}_{k',t'}^{*}, \tag{93}$$

where $\Omega_{k,t}$ is the set of the CRS positions. $\bar{\mathbf{g}}_{k,t}$ will be used by the equalizer in this resource block to recover the signal. Similarly, if we assume $\mathbf{g}_{\Omega_{k,t}}$ is the same in a resource block, the optimal solution is given by

$$\min_{\mathbf{g}_{\Omega}} \sum_{(k',t') \in \Omega} \left| \mathbf{g}_{\Omega} \mathbf{y}_{k',t'} - c_{RS,(k',t')} \right|^{2} \tag{94}$$

$$\mathbf{g}_{\Omega}^{*} = \left(\sum_{(k',t') \in \Omega} \mathbf{y}_{k',t'}^{H} \mathbf{y}_{k',t'} \right)^{-1} \sum_{(k',t') \in \Omega} \mathbf{y}_{k',t'}^{H} c_{RS,(k',t')} \tag{95}$$

3 Interference mitigation for synchronization

In section C.1 and C.2, both the WiFi and LTE interference cancellers are obtained following an accurate time and frequency synchronization. However, with a high power interference, the synchronization usually has very poor performance. To increase the synchronization performance, the receiver requires a pre-process filter to mitigate the interference without any prior information.

From the received signal stream $\mathbf{Y}_{N_r \times L}$ (N_r is the number of receiving antenna, L is the number of received discrete signal samples), the receiver needs to separate two uncorrelated signal streams $\mathbf{v}_{1,1\times L}$ and $\mathbf{v}_{2,1\times L}$ using a linear spatial filter $\mathbf{u}_{1,1\times N_r}$ and $\mathbf{u}_{2,1\times N_r}$, respectively (i.e., $\mathbf{v}_1 = \mathbf{u}_1\mathbf{Y}$, $\mathbf{v}_2 = \mathbf{u}_2\mathbf{Y}$). As the LTE and WiFi transmitters are totally independent, it can assume the separated signal stream $\mathbf{v}_1$ (or $\mathbf{v}_2$) only contains LTE signal or WiFi signal. Therefore, the synchronization can be realized based on just $\mathbf{v}_1$

or $\mathbf{v}_2$. The uncorrelation condition of $\mathbf{v}_1$ and $\mathbf{v}_2$ requires

$$\mathbf{v}_1 \mathbf{v}_2^H = 0, \tag{96}$$

which is equivalent to

$$\mathbf{u}_1 \mathbf{Y} \mathbf{Y}^H \mathbf{u}_2^H = 0. \tag{97}$$

Equation (97) is the definition for the eigendecomposition of $\mathbf{Y}\mathbf{Y}^H$, where $\mathbf{u}_1$ and $\mathbf{u}_2$ are two distinct eigenvectors. Therefore, the pre-process spatial filter can be obtained just from the decomposition of $\mathbf{Y}\mathbf{Y}^H$.

D Receiver design and prototyping

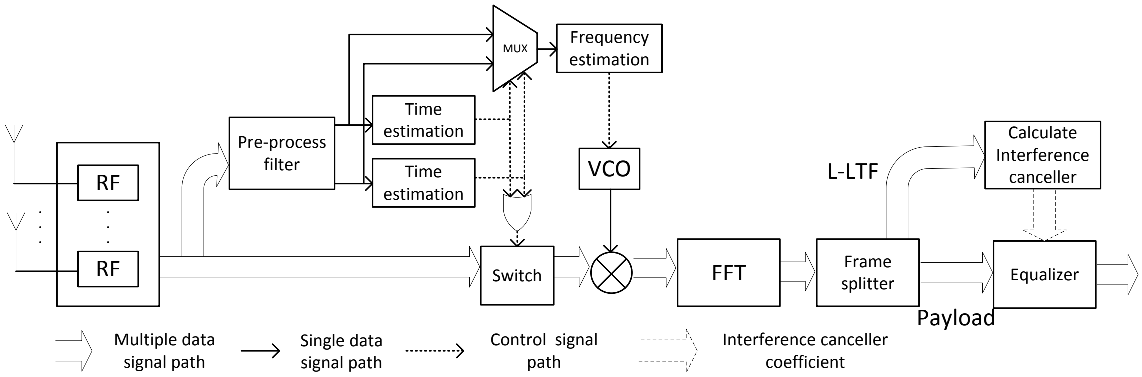


Figure 25: WiFi interference cancelling receiver.

The designed WiFi receiver structure is demonstrated in Figure 25. The pre-process filter separate the multiple signal paths from different antennas into two signal paths $\mathbf{v}_1$ and $\mathbf{v}_2$ based on the method in section C.3. The time estimation method introduced in section B.1 is conducted on both $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively. The time estimation block will generate a pulse if the correlation value is greater than the threshold. The MUX selects $\mathbf{v}_1$ or $\mathbf{v}_2$ based on the result of the time estimation block. Then the frequency estimation will be conducted on the signal path selected by the MUX. The frequency estimation result drives a VCO to generate a sine waveform to compensate the frequency offset on the signal path. For burst transmission, the pulse generated by the time estimation block will turn on a switch at the signal path. The switch will be turned off after reading the whole packet based on the length information in the L-SIG. After the time and frequency synchronization, the signals are converted into frequency domain by FFT. The frame splitter picks the L-LTF signal, and the interference canceller is obtained based on the L-LTF. Finally, the interference canceller coefficient $\bar{\mathbf{g}}$ is sent to the equalizer recovering constellation symbols. The LTE receiver structure is similar as demonstrated in Figure 26. The key difference is that the LTE transmission is not burst based. Therefore, a delay block is used instead of the switch to align the signal path to the pulse generated by the time estimation block.

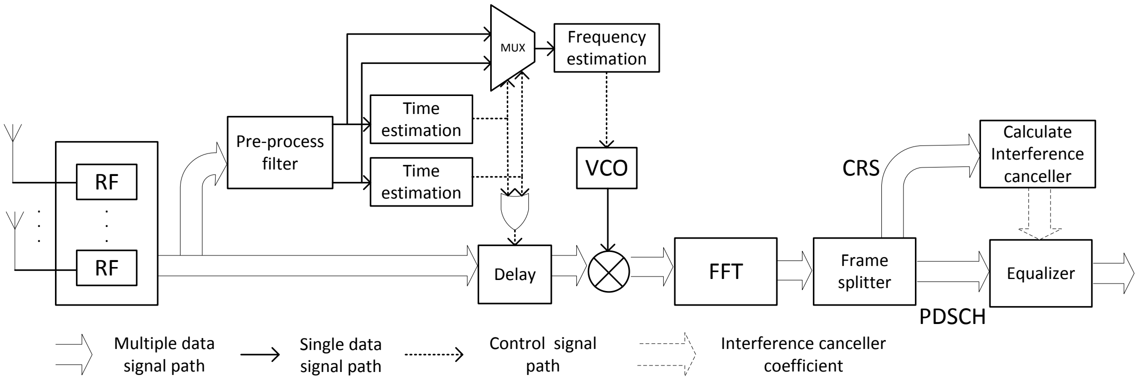


Figure 26: LTE interference cancelling receiver.

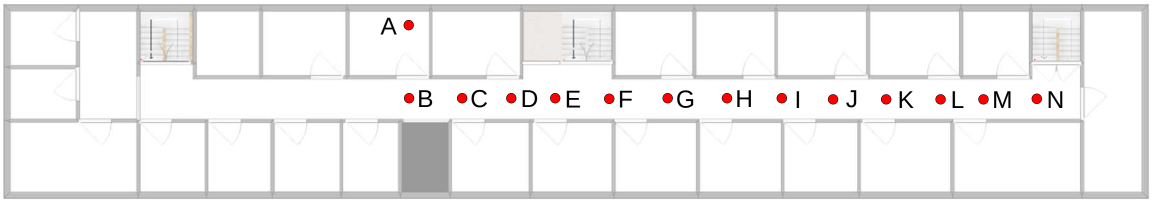


Figure 27: Test scenario.

The hardware used for prototyping is USRP N210. USRP N210 is one of the highest performing class of the USRP (Universal Software Radio Peripheral) family of products, which enables rapidly design and implement powerful, flexible software radio systems. N210 architecture includes a Xilinx Spartan FPGA, 100 MS/s dual ADC, 400 MS/s dual DAC and Gigabit Ethernet connectivity to stream data to and from host processors. USRP N210 can operate from DC to 2 GHz, while an expansion port allows multiple USRP N210 series devices to be synchronized and used in a MIMO configuration. An optional GPSDO module can also be used to discipline the USRP N210 reference clock to within 0.01 ppm of the worldwide GPS standard. The USRP N210 can stream up to 50 MS/s to and from host applications.

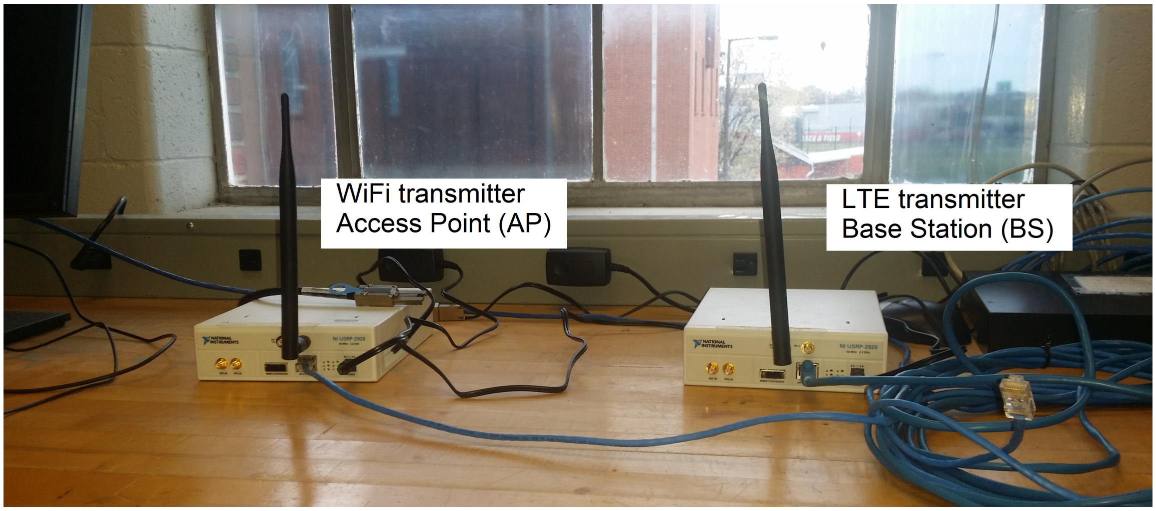


Figure 28: Transmitter location.

In the current stage, the sampled data from USRP N210 is processed off-line. The benefit of the off-line receiver is easier to measure the performance quantitatively. Furthermore, it is more convenient to take the advantage of GUI softwares such as MATLAB to demonstrate performance. For the next stage of prototyping, we will realize the real-time receiver with GNU-Radio. Two USRP N210 are used to implement a two-antenna WiFi or LTE receiver. The WiFi transmitter (i.e., AP) and LTE transmitter (i.e., BS) is realized by only one USRP N210 each. The transmitted signal waveform is generated from 801.11 and LTE PHY standards. The test scenario (shown in Figure 27) is the 2nd floor of W.S. Speed Hall, University of Louisville. Both the WiFi and LTE transmitters are placed at location (A) as shown in Figure 28.

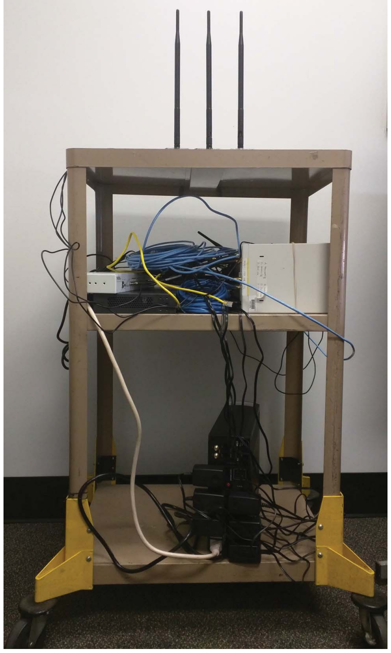
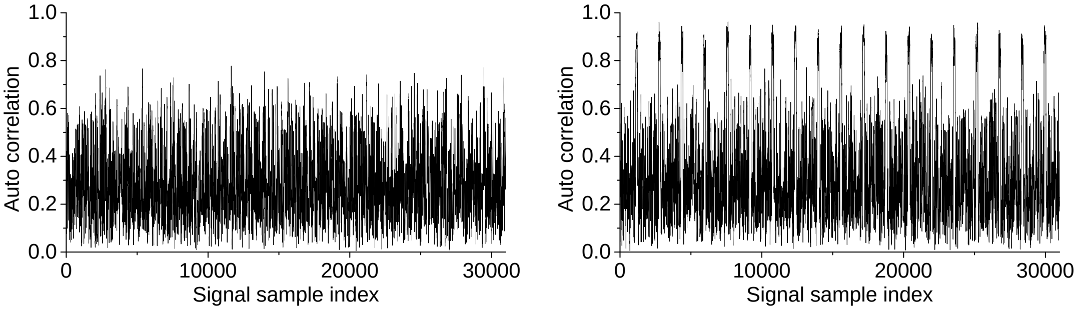


Figure 29: Receiver platform.

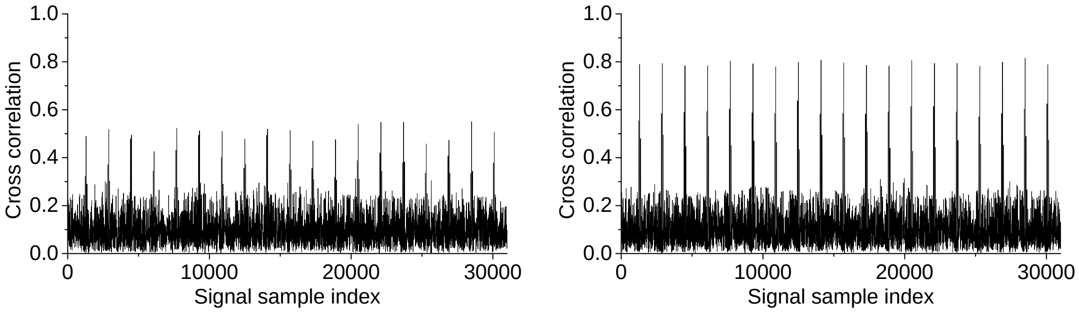


(a) without the pre-process filter

(b) with the pre-process filter

Figure 30: L-STF auto-correlation result.

The receiver (Figure 29) is first placed at location (G). The WiFi and LTE

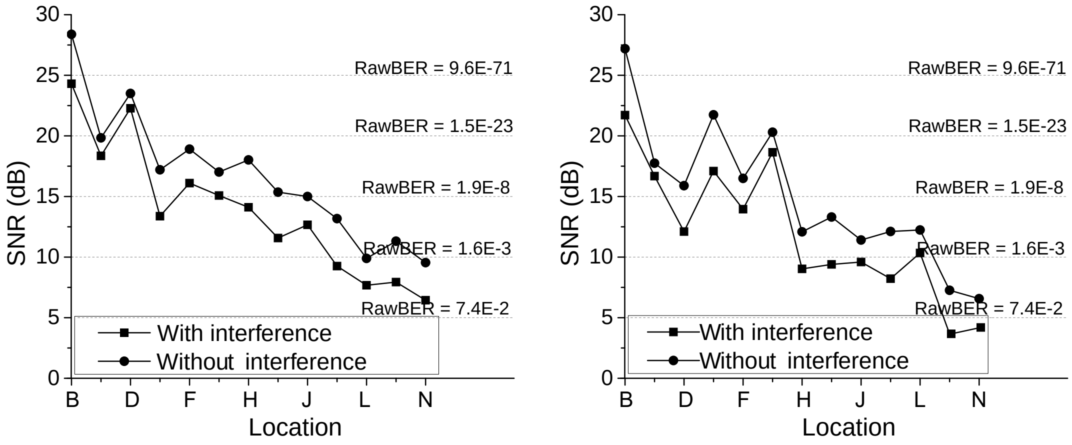


(a) without the pre-process filter

(b) with the pre-process filter

Figure 31: L-LTF cross-correlation result.

transmitters use the same transmit power 20 dBm. Figure 30 and Figure 31 present the correlation results without and with the pre-process filter. From the results, it shows that the pre-process filter effectively improves the correlation result for both the auto and cross correlation methods. The pre-process filter enables an accurate time synchronization even under high interference power.



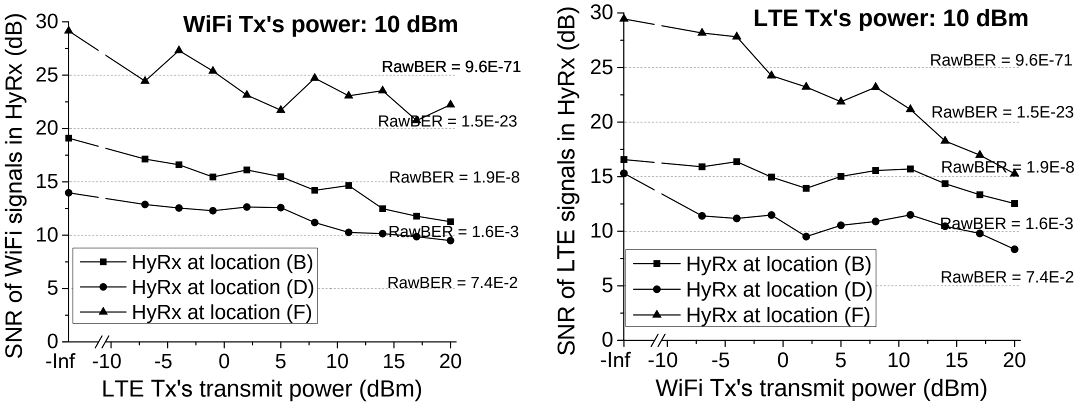
(a) WiFi receiver

(b) LTE receiver

Figure 32: Performance in different locations.

Then the receiver is tested at different locations. The received SNR is defined as $10\log_{10}(\mathbb{E}(|X - \hat{X}|^2)/\mathbb{E}(|X|^2))$. The QPSK system's Raw-BER (before channel decoding) can be evaluated based on the received SNR by

Raw-BER $= 2Q(\sqrt{\text{SNR}}) - Q^2(\sqrt{\text{SNR}})$, where $Q(\cdot)$ is the Q-function. The receiver is placed from location (B) to (N) and measured with the received SNR. The received SNR with interference is compared with the received SNR without interference, in order to demonstrate the efficiency of the designed interference canceller. Specifically, Figure 32 shows that the WiFi receiver has an average 2.9 dB received SNR degradation compared to the case without interference, the LTE receiver has an average 3.1 dB received SNR degradation compared to the case without interference.



(a) WiFi receiver

(b) LTE receiver

Figure 33: Impact of interference power.

The impact of interference power on the receiver is also studied. The receiver is placed at location (B), (C), and (D), respectively. At each location, set the WiFi (resp. LTE) transmitter's power to 10 dBm and change the LTE (resp. WiFi) transmitter's power from -7 dBm to 20 dBm. Figure 33 presents the received SNR of the WiFi and LTE receivers at different locations. The results show that, even if the WiFi (resp. LTE) transmit power is 10 dB smaller than the LTE (resp. LTE) power, the receiver can still successfully decode the signal.

E Conclusion

In this Chapter, a MIMO interference cancellation LTE and WiFi receiver was designed and prototyped to support the coexistence of the LTE-U and WiFi network. The platform was tested in real scenario demonstrating the performance. The test results indicated that the MIMO interference canceller receivers can effectively support the coexistence.

CHAPTER V

FINITE-BLOCKLENGTH SECRECY PERFORMANCE OF LATTICE CODES

The theoretical foundation of the physical-layer security can be traced back to the early works of Wyner, Csiszar and Korner. In [47], Wyner introduced the notion of the wiretap channel, which was further generalized by Csiszar and Korner[48]. The wiretap channel model is shown in Figure34, where there are two legitimate communicators: Alice and Bob. Alice wishes to send a secret message U, $u \in \{1, 2 \cdots, M_0\}$, to Bob through the main channel $\left(\mathcal{X}, W_{\mathbf{Y}|\mathbf{X}}, \mathcal{Y}\right)$. However, this transmission also reaches a passive eavesdropper Eve through the wiretap channel $\left(\mathcal{X}, W_{\mathbf{Z}|\mathbf{X}}, \mathcal{Z}\right)$. Alice and Bob have agreed on an encoding/decoding mechanism, while this information is also acknowledged at Eve. Furthermore, Eve does not suffer from any limitation on computational complexity and time consumption. Meanwhile, Alice also has an auxiliary message A, $a \in \{1, 2 \cdots, M_1\}$, which is unknown to both Bob and Eve. This auxiliary message is used to 'randomize' and 'cover' the secret message. Alice maps U and A to a transmitted word $\mathbf{X}$ and broadcasts it as $U \times A \to \mathbf{X} \to \mathbf{YZ}$, where $\mathbf{x} \in \mathcal{X}^n$, $\mathbf{y} \in \mathcal{Y}^n$ and $\mathbf{z} \in \mathcal{Z}^n$.

At the receiver side, Bob maps $\mathbf{Y}$ into an estimation $\hat{U}$. The reliability is measured

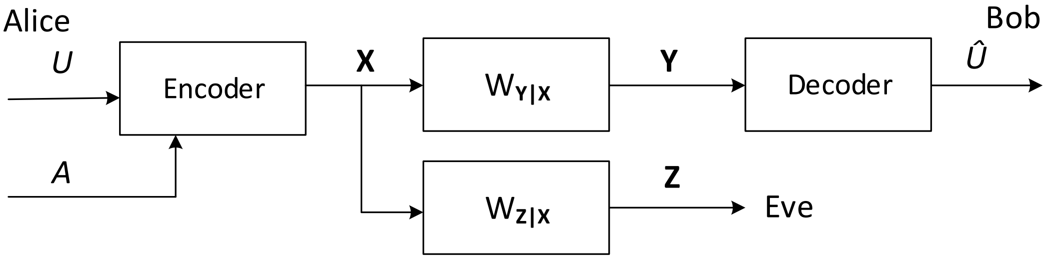


Figure 34: Wiretap channel model.

in terms of the error probability to recover U. Specifically, the asymptotic reliability is:

$$\lim_{n \to \infty} \mathbb{P}\left\{\hat{U} \neq U\right\} = 0. \tag{98}$$

And the secrecy is measured by the dependence between the observation $\mathbf{Z}$ and the secret message U. The weak secrecy is introduced by Wyner in [47] as:

$$\lim_{n \to \infty} \frac{I\left(\mathbf{Z}; U\right)}{n} = 0. \tag{99}$$

However, [49–51] showed that (99) is much too weak as a large portion of the message can still be recovered at Eve. Thus, Maurer further introduced the strong secrecy in [49] as

$$\lim_{n \to \infty} I\left(\mathbf{Z}; U\right) = 0, \tag{100}$$

The secrecy capacity is defined as the maximal secrecy rate satisfying both the reliability (98) and the secrecy (99)(100). In particular, Theorem 1 is obtained by Cisizar [51], Maurer [49] and Bloch [50] with different techniques.

Theorem 1. *The secrecy capacity of the memoryless channel* $\left(\mathcal{X}, \mathcal{Y}, P_{YZ|X}, \mathcal{Z}\right)$ *is given by*

$$C_s = \sup\left(I\left(Y; X\right) - I\left(Z; X\right)\right), \tag{101}$$

In the literature, [52–54] showed that the coding schemes with the auxiliary messages approaching the capacity of the wiretap channel can achieve the weak secrecy.

Nevertheless, it is not true when applied to the strong secrecy. Extensive studies have been conducted to find the coding schemes achieving the secrecy capacity. For example, lattice codes have long been known as one of the capacity-achieving coding schemes. When applied to secure transmissions, it has been proven that nested lattice codes can achieve the secrecy capacity [55, 56]. For linear codes, LDPC codes are proposed for the erasure channel to achieve the strong secrecy [57–60]. Along the same line, polar codes are another type of linear block codes introduced by Arikan [61]. Polar codes are structure-specified and provably capacity-achieving over any binary-input symmetric channel. The secure polar codes have been explored to achieve the strong secrecy [54, 62]. Other secure coding schemes can be found in [63–66]. The wiretap channel model is also studied with various configurations: multi-antenna wiretap channel[67–74], wiretap channel with feedback [75–77], wiretap channel with side information[78, 79] and multiple access wiretap channel[80]. However, these coding schemes and systems are all measured by the secrecy capacity, in which the secrecy are analyzed asymptotically with infinite blocklength.

Assume a finite blocklength n ($n < \infty$) and a erasure channel with a erasure probability δ for Eve. Name the case that no erasure happens as *error-free realization*. The probability of the error-free realization is $(1-\delta)^n > 0$. For any coding scheme with a coding rate strictly less than 1, Eve can definitely recover this secret message when the error-free realization happens. Similarly, Bob is not able to achieve any error-free transmission with the finite blocklength. It is worth noting that this finite blocklength penalty has little impact on the reliable communication, in which the channel coding is combined with the auto-repeat-request (ARQ) technique to achieve a desired reliability.

However, the finite-blocklength penalty is crucial for the secrecy. As the information leak is irreversible, in the worst case (i.e., the error-free realization happens), Eve will recover the whole secret message. Because there is no limitation on Eve, it can assume that Eve is powerful enough to simultaneously obtain a large number of independent channels, and have at least one error-free realization after searching all these wiretap channels. Theoretically, Eve is always able to recover the secret message regardless of coding schemes. In summary, the example tells us: (1) The secret information leak is inevitable with finite blocklengths; (2) The secret information leak is irreversible. Therefore, it is essential to quantify how difficult for Eve to recover this secret message from an information-theoretic perspective.

The literature that focuses on the finite-blocklength secrecy can only be found recently. [81] investigates the maximal secrecy rate over a wiretap channel at a given blocklength. The maximal secrecy rate in [81] is defined by the constraint of a given amount of secret information leak probability, which is similar to the definition of the leak probability in this work. The achievability and converse bounds of this maximal secrecy rate are derived using the uniform-partition codes in [81]. These bounds lead to the tightest second-order coding rate, stated as the theorem 9 in [81]. It can be easily proved that, even through different approaches, the tightest second-order coding rate in [81] is identical to the results stated as Theorem 8 in this work. Compared to [81], this work otherwise focuses on computationally trackable analysis. The average secret information leakage is also adopted as one metric apart from the leak probability. With different emphases, this work demonstrates how secure lattice codes can be analyzed practically at a given blocklength. [82] provide a new secrecy metric for finite blocklengths — the bit error rate cumulative distribution function (BER-CDF). Different from traditional impractical information-theoretic secrecy measures, the BER-CDF is simple and computationally trackable. Therefore, the BER-CDF provide a practical approach for the analysis of the finite-blocklength secrecy. However, the BER-CDF can only be used in binary input channels and is not able to guarantee the information-theoretic secrecy. Compared to [82], this work strive to provide a practical analysis of the finite-blocklength secrecy but based on the information-theoretic secrecy. This approach is not limited to binary input channels, as shown by the analysis of secure lattice codes.

In this chapter, upper case letters and the corresponding lower case letters denote random variables (e.g., X) and their realizations (e.g., x), respectively; bold letters denote arrays of random variables (e.g., $\mathbf{X}$) or realizations (e.g., $\mathbf{x}$); $P_X(\cdot)$ (resp. $f_X(\cdot)$) denotes the probability mass (density) function of the discrete (continuous) random variable X; $\mathbb{E}[g(X)]_{f_X(\cdot)}$ denotes the expectation of $g(X)$ in regard to the random variable X with the probability density function (PDF) $f_X(\cdot)$; $\log(x)$ and $\ln(x)$ refer to the binary and nature logarithm respectively. Note that only memoryless channels are considered in this paper.

A Finite-blocklength secrecy analysis

To facilitate the discussion, define the entropy density and the mutual information density as:

$$h(\mathbf{x}) = \log \frac{1}{P_{\mathbf{X}}(\mathbf{x})},\ i(\mathbf{z};\mathbf{x}) = \log \frac{P_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x})}{P_{\mathbf{Z}}(\mathbf{z})},$$

where $h(\mathbf{X})$ and $i(\mathbf{Z};\mathbf{X})$ are the corresponding random variables.

For a secure code, the auxiliary message A is chosen to ‘fully cover’ the secret message U. For example, in a transmission $u \times a \to \mathbf{x} \to \mathbf{z}$, the receiver is not able to infer the message u based on the observation $\mathbf{z}$ without the message a as

$$\int\limits_{\bar{a} \in A} P\left(u|\mathbf{z}\bar{a}\right) P\left(\bar{a}\right) d\bar{a} = P\left(u\right) \tag{102}$$

With the correct message a at the receiver, the receiver is able to further infer the message u as

$$P\left(u|\mathbf{z}a\right) \geq P\left(u\right) \tag{103}$$

However, with an incorrect message a' ($a' \neq a$), the receiver will make a false inference on the message u as

$$P\left(u|\mathbf{z}a'\right) \leq P\left(u\right) \tag{104}$$

Apparently, the maximal mutual information density is obtained when the receiver correctly decode a as

$$\underset{P(a|\mathbf{z})}{arg\max}\, i\left(\mathbf{z};u\right) = \underset{P(a|\mathbf{z})}{arg\max} \int\limits_{\bar{a} \in A} P\left(u|\mathbf{z}\bar{a}\right) P\left(\bar{a}|\mathbf{z}\right) d\bar{a} = 1 \tag{105}$$

Therefore, the maximal $i\left(\mathbf{z};u\right)$ is

$$\max i\left(\mathbf{z};u\right) = \max \log \frac{P\left(u|\mathbf{z}\right)}{P\left(u\right)} = \log \frac{P\left(ua|\mathbf{z}\right) P\left(a\right)}{P\left(u\right) P\left(a\right)} \tag{106}$$

As $\mathbf{x}$ is fully determined by u and a (i.e., $P\left(u\right) P\left(a\right) = P\left(\mathbf{x}\right)$), it gives

$$\max i\left(\mathbf{z};u\right) = \log \frac{P\left(\mathbf{x}|\mathbf{z}\right) P\left(a\right)}{P\left(\mathbf{x}\right)} = i\left(\mathbf{z};\mathbf{x}\right) - h\left(a\right) \tag{107}$$

Note that $P\left(u|\mathbf{z}a\right) \geq P\left(u\right)$, the maximal $i\left(\mathbf{z};u\right)$ is non-negative.

1 Finite-blocklength secrecy performance

The finite-blocklength secrecy performance for different channels can be measured by Definition 1.

Definition 1. *i)* Leak Probability*:*

$$L_P = \mathbb{P}\left[\max i\left(\mathbf{Z}; U\right) > 0\right]; \tag{108}$$

ii) Average Leakage*:*

$$L_A = \mathbb{E}\left[\max i\left(\mathbf{Z}; U\right)\right]. \tag{109}$$

For the convenience of illustrations, adopt the following notations:

Expectation:

$$I_Z = \frac{1}{n}\mathbb{E}\left[i\left(\mathbf{Z}; \mathbf{X}\right)\right] = \sum_{x\in X}\sum_{z\in Z} P_X\left(x\right) P_{Z|X}\left(z|x\right) \log\frac{P_{Z|X}\left(z|x\right)}{P_Z\left(z\right)}; \tag{110}$$

Variance:

$$V_Z = \frac{1}{n}\mathbb{E}\left[\left(i\left(\mathbf{Z}; \mathbf{X}\right) - nI_Z\right)^2\right] = \sum_{x\in X}\sum_{z\in Z} P_X\left(x\right) P_{Z|X}\left(z|x\right) \log^2\frac{P_{Z|X}\left(z|x\right)}{P_Z\left(z\right)} - I_Z^2; \tag{111}$$

Skewness:

$$T_Z = \sqrt{n}\mathbb{E}\left[\left|\frac{i\left(\mathbf{Z}; \mathbf{X}\right) - nI_Z}{\sqrt{nV_Z}}\right|^3\right] = \sum_{x\in X}\sum_{z\in Z} P_X\left(x\right) P_{Z|X}\left(z|x\right) \left|\frac{\log\frac{P_{Z|X}\left(z|x\right)}{P_Z\left(z\right)} - I_Z}{\sqrt{V_Z}}\right|^3. \tag{112}$$

Theorem 2. *The blocklength-n leak probability L_P is bounded by*

$$L_P \leq Q\left(\frac{\log M_1 - nI_Z}{\sqrt{nV_Z}}\right) + \frac{T_Z}{2\sqrt{n}}, \tag{113}$$

$$L_P \geq Q\left(\frac{\log M_1 - nI_Z}{\sqrt{nV_Z}}\right) - \frac{T_Z}{2\sqrt{n}}. \tag{114}$$

Proof. Theorem 2 can be easily proved with the Berry-Esseen Theorem.

Theorem 3. *(Berry-Esseen) Suppose $X_1, X_2, \ldots, X_n$ are i.i.d random variables with $\mu = \mathbb{E}\left[X_i\right]$, $\sigma^2 = \mathrm{Var}\left[\mathrm{X_i}\right]$, $t = \mathbb{E}\left[\left|X_i - \mu_i\right|^3\right]$. Then for any λ, it gives*

$$\left|\mathbb{P}\left[\sum_{i=1}^{n} X_i - n\mu \leq \lambda n\sigma\right] - \Phi\left(\lambda\right)\right| \leq \frac{t}{2\sigma^3\sqrt{n}}, \tag{115}$$

where $\Phi\left(\lambda\right)$ is the cumulative distribution function (CDF) of the normal distribution.

Define new random variables

$$\Lambda_j = i\left(Z_j; X_j\right) - \frac{\log M_1}{n}, j = 1, 2, \cdots, n. \tag{116}$$

Then, it gives $\mathbb{E}\left[\Lambda_j\right] = I_Z - \frac{\log M_1}{n}$, $\mathrm{Var}\left[\Lambda_j\right] = V_Z$ and $\mathbb{E}\left[\left|\frac{\Lambda_j - E[\Lambda_j]}{\sqrt{\mathrm{Var}[\Lambda_j]}}\right|^3\right] = T_Z$. Apply Λ_j with the Berry-Esseen Theorem, it gives

$$\left|\mathbb{P}\left[\sum_{j=1}^{n} \Lambda_j - \left(nI_Z - \log M_1\right) \leq \lambda\sqrt{nV_Z}\right] - \Phi\left(\lambda\right)\right| \leq \frac{T_Z}{2\sqrt{n}}, \tag{117}$$

$$\left|\mathbb{P}\left[I\left(\mathbf{Z}; \mathbf{X}\right) - \log M_1 \leq \lambda\sqrt{nV_Z} + nI_Z - \log M_1\right] - \Phi\left(\lambda\right)\right| \leq \frac{T_Z}{2\sqrt{n}}. \tag{118}$$

Let $\lambda = \frac{\log M_1 - nI_Z}{\sqrt{nV_Z}}$, it gives

$$\left|\mathbb{P}\left[I\left(\mathbf{Z}; \mathbf{X}\right) - \log M_1 \leq 0\right] - \Phi\left(\frac{\log M_1 - nI_Z}{\sqrt{nV_Z}}\right)\right| \leq \frac{T_Z}{2\sqrt{n}}, \tag{119}$$

which is

$$\left|\mathbb{P}\left[I\left(\mathbf{Z}; U\right) > 0\right] - Q\left(\frac{\log M_1 - nI_Z}{\sqrt{nV_Z}}\right)\right| \leq \frac{T_Z}{2\sqrt{n}}. \tag{120}$$

□

Theorem 4. *The blocklength-n average leakage L_A is approximated by*

$$L_A \approx \sqrt{\frac{nV_Z}{2\pi}} e^{-\frac{\left(\log M_1 - nI_Z\right)^2}{2nV_Z}} - \left(\log M_1 - nI_Z\right) Q\left(\frac{\log M_1 - nI_Z}{\sqrt{nV_Z}}\right). \tag{121}$$

Proof.

$$\mathbb{E}\left[\left(\mathbf{Z}; U\right)\right] = \mathbb{E}\left[i\left(\mathbf{Z}; \mathbf{X}\right)\left|i\left(\mathbf{Z}; \mathbf{X}\right) \geq \log M_1\right.\right] - \log M_1 \mathbb{P}\left[i\left(\mathbf{Z}; \mathbf{X}\right) \geq \log M_1\right]. \tag{122}$$

Because

$$i\left(\mathbf{z}; \mathbf{x}\right) = \sum_{i=1}^{n} \log \frac{P_{Z_i|X_i}\left(z_i|x_i\right)}{P_{Z_i}\left(z_i\right)}, \tag{123}$$

according to the central limit theorem, the distribution of $i\left(\mathbf{Z};\mathbf{X}\right)$ can be approximated by the normal distribution with $\mu{=}nI_Z$ and $\sigma^2 = nV_Z$. Then, (122) can be approximated by

$$\int_{\log M_1}^{\infty} i\left(\mathbf{Z};\mathbf{X}\right)\frac{1}{\sqrt{2nV_Z}}e^{-\frac{\left(i(\mathbf{Z};\mathbf{X})-nI_Z\right)^2}{2nV_Z}}di\left(\mathbf{Z};\mathbf{X}\right)-\log M_1 Q\left(\frac{\log M_1-nI_Z}{\sqrt{nV_Z}}\right). \tag{124}$$

Therefore, it gives

$$L_A \approx \sqrt{\frac{nV_Z}{2\pi}}e^{-\frac{\left(\log M_1-nI_Z\right)^2}{2nV_Z}}-\left(\log M_1-nI_Z\right)Q\left(\frac{\log M_1-nI_Z}{\sqrt{nV_Z}}\right). \tag{125}$$

□

Theorem 5. *Take the entropy of the auxiliary message equal to the capacity of the wiretap channel (i.e.,* $\log M_1 = nI_Z$*), the asymptotic secrecy performance is*

$$\lim_{n\to\infty} L_P = \frac{1}{2}; \tag{126}$$

$$\lim_{n\to\infty} L_A = \infty. \tag{127}$$

Theorem 5 just shows that capacity-achieving codes are not good enough for the strong secrecy.

Theorem 6. *i) For a fixed entropy* M_0*, define the effective auxiliary ratio as*

$$\eta_{En} = \frac{\log M_1}{nH\left(X\right)-\log M_0}. \tag{128}$$

When $\eta_{En} < \frac{I_Z}{H(X)}$, $\lim\limits_{n\to\infty} L_p = 1$ *and* $\lim\limits_{n\to\infty} L_A = \log M_0$*;*

When $\eta_{En} > \frac{I_Z}{H(X)}$, $\lim\limits_{n\to\infty} L_p = 0$ *and* $\lim\limits_{n\to\infty} L_A = 0$.

ii) For a fixed rate $R_s = \log M_0/\left(nH\left(X\right)\right)$*, define the effective auxiliary ratio as*

$$\eta_{R_s} = \frac{\log M_1}{nH\left(X\right)\left(1-R_s\right)}. \tag{129}$$

When $\eta_{R_s} < \frac{I_Z}{H(X)(1-R_s)}$, $\lim\limits_{n\to\infty} L_p = 1$ *and* $\lim\limits_{n\to\infty} L_A = \infty$*;*

When $\eta_{R_s} > \frac{I_Z}{H(X)(1-R_s)}$, $\lim\limits_{n\to\infty} L_p = 0$ *and* $\lim\limits_{n\to\infty} L_A = 0$.

Remark: Theorem 6 can be directly proved from Theorem 2 and Theorem 4. For part i), when M_0 is fixed and n goes to infinity, the coding rate goes to zero. However, even the coding rate goes to zero asymptotically, Eve is not able to recover any secret message if η_{En} is above the threshold. For part ii), compared with Theorem 5, it shows that a code achieving the wiretap channel capacity will leak the whole secret message to Eve. However, if the coding rate is greater than the wiretap channel capacity, the strong secrecy can be achieved asymptotically.

Example 1. *Binary Erasure Channel (BEC)*

Use the BEC as an example to explain Theorem 2-6 (the results can be easily extended to other channel models). Assume the input X is uniformly distributed (i.e., $P_X(1) = P_X(0) = \frac{1}{2}$*), the variance is*

$$V_Z = 2\left(\frac{1-\delta}{2}(1-(1-\delta))^2 + \frac{\delta}{2}(0-(1-\delta))^2\right) = (1-\delta)\,\delta, \tag{130}$$

and the skewness is

$$T_Z = (1-\delta)^{\frac{3}{2}}\delta^{\frac{3}{2}}\left((1-\delta)\,\delta^3 + \delta(1-\delta)^3\right). \tag{131}$$

Let the erasure probability $\delta = 0.6$ *and the coding rate* $\eta_{R_s} = 0.6$*, it gives* $I_Z = 0.4$*,* $V_Z = 0.24$ *and* $T_Z = 0.0147$*. The secrecy threshold in Theorem 6 is calculated as* $\eta^*_{R_s} = 0.5$.

Figure 2 shows L_P *as a function of blocklength* n*. For* $\eta^*_{R_s} = 0.5$*,* L_P *converges to* 0.5*, which validates Theorem 5. It is interesting to see that, when* $\eta_{R_s} = 0.51$ *(above the*

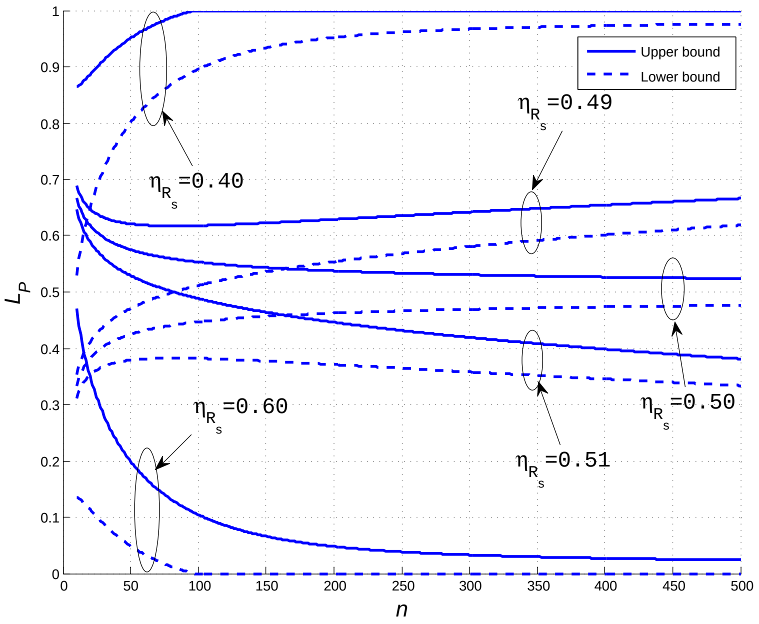


Figure 35: Leak probability L_P vs blocklength n, BEC.

threshold 0.5*),* L_P *will slowly converge to 0, indicating that it requires a large blocklength for an acceptable secrecy performance. When* $\eta_{R_s} = 0.6$*,* L_P *converges to* 0 *more quickly such that a blocklength of 300 is long enough for a small* L_P*. Figure 3 shows* L_A *as a function of blocklength* n*: (1)* L_A *is infinity when* $\eta_{R_s} = 0.5$*. (2)* L_A *curve is not monotone with* n*. When* $\eta_{R_s} = 0.51$*,* L_A *increases with* n *when* $n < 1000$*, which indicates that a small blocklength is beneficial even when* η_{R_s} *is greater than the threshold. Nevertheless, when* η_{R_s} *is large enough, the* L_P *curve is almost monotone with* n*.*

2 Trade-off between the reliability and the secrecy

The performance analysis in Section III only discussed the wiretap channel, while the main channel is an ordinary communication channel with the transmitted message $U \times A$, which can be analyzed using the results in [83]. Combining these results, the wiretap system can be fully evaluated and optimized.

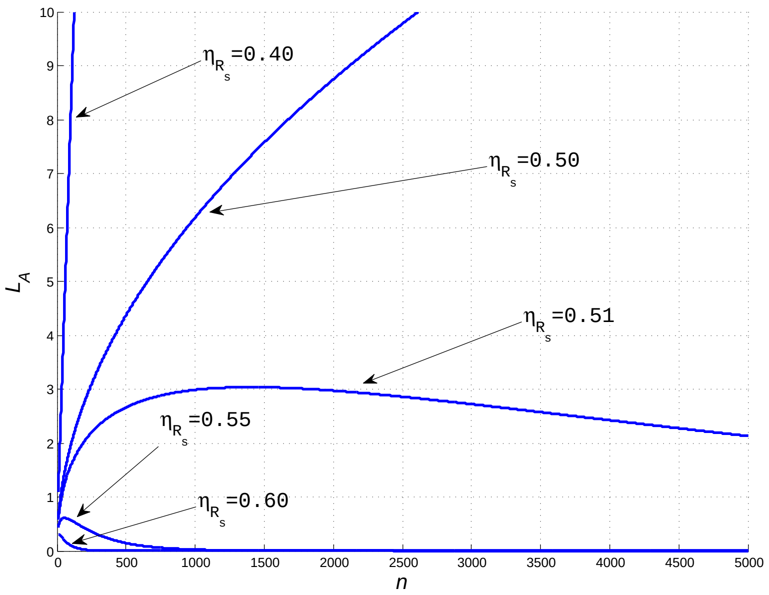


Figure 36: Average leakage L_A vs blocklength n, BEC.

From [83], the average error probability of the main channel is bounded by

$$\epsilon \leq \mathbb{P}\left[i\left(\mathbf{Y};\mathbf{X}\right) \leq \log\frac{M_0M_1-1}{2}\right] + \frac{M_0M_1-1}{2}\mathbb{P}\left[i\left(\bar{\mathbf{Y}};\mathbf{X}\right) > \log\frac{M_0M_1-1}{2}\right], \tag{132}$$

where $\bar{\mathbf{Y}}$ has the same distribution as $\mathbf{Y}$ but is independent of $\mathbf{X}$. Similarly, use the following notations of the main channel:

$$I_Y = \frac{1}{n}\mathbb{E}\left[i\left(\mathbf{Y};\mathbf{X}\right)\right] = \sum_{x\in X}\sum_{y\in Y} P_X\left(x\right)P_{Y|X}\left(y|x\right)\log\frac{P_{Y|X}\left(y|x\right)}{P_Y\left(y\right)}; \tag{133}$$

$$V_Y = \frac{1}{n}\mathbb{E}\left[\left(i\left(\mathbf{Y};\mathbf{X}\right)-nI_Y\right)^2\right] = \sum_{x\in X}\sum_{y\in Y} P_X\left(x\right)P_{Y|X}\left(y|x\right)\log^2\frac{P_{Y|X}\left(y|x\right)}{P_Y\left(y\right)} - I_Y^2. \tag{134}$$

From [83], the average error probability of the main channel can be approximated by

$$\epsilon \approx Q\left(\frac{nI_Y - \log M_0M_1}{\sqrt{nV_Y}}\right). \tag{135}$$

Using Theorem 2, Theorem 4 and (135), it can obtain the secrecy performance conditioning on the desired main channel average error probability.

Theorem 7. *Given the blocklength n, the secret information M_0 and the main channel maximum average error probability ϵ, the secrecy performance is*

$$L_P^*\left(n,\epsilon,M_0\right) \leq Q\left(\frac{n\left(I_Y - I_Z\right) - \sqrt{nV_Y}Q^{-1}\left(\epsilon\right) - \log M_0}{\sqrt{nV_Z}}\right) + \frac{T_Z}{2\sqrt{n}}; \tag{136}$$

$$L_A^*\left(n,\epsilon,M_0\right) = \sqrt{\frac{nV_Z}{2\pi}}e^{-\frac{\left(n\left(I_Y-I_Z\right)-\sqrt{nV_Y}Q^{-1}\left(\epsilon\right)-\log M_0\right)^2}{2nV_Z}} - \left(n\left(I_Y - I_Z\right) - \sqrt{nV_Y}Q^{-1}\left(\epsilon\right) - \log M_0\right) Q\left(\frac{n\left(I_Y - I_Z\right) - \sqrt{nV_Y}Q^{-1}\left(\epsilon\right) - \log M_0}{\sqrt{nV_Z}}\right). \tag{137}$$

Example 2. *(Binary Erasure Channel for both main channel and wiretap channel)*

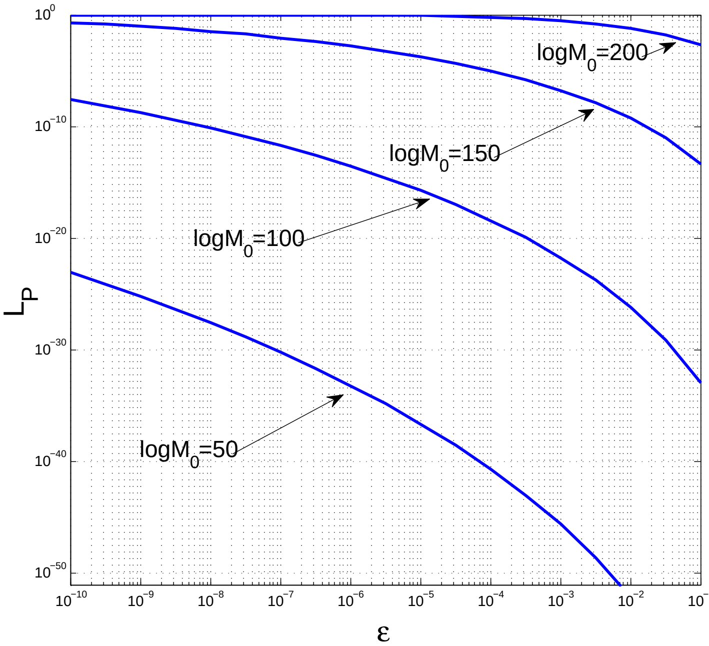


Figure 37: Leak probability L_P vs main channel error probability ϵ, BEC.

In this example, the main channel and the wiretap channel have the erasure probabilities $\delta_Y = 0.1$ and $\delta_Z = 0.6$, respectively. The input X is uniformly distributed with $P_X(1) = P_X(0) = \frac{1}{2}$ and the blocklength is $n = 500$. Figure 37 (resp. Figure 38) shows the trade-off between L_P (resp. L_A) and ϵ. It can seen that, only when $\log M_0$ is

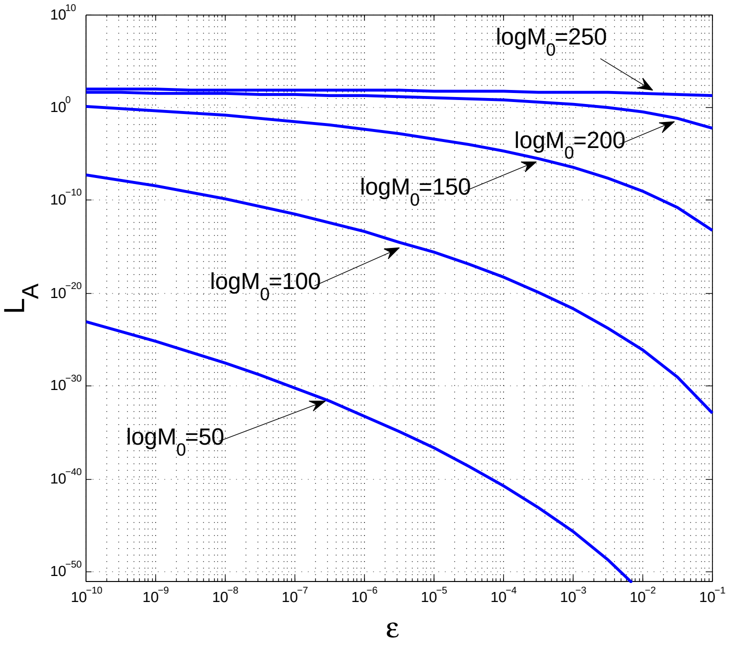


Figure 38: Average leakage L_A vs main channel error probability ϵ, BEC.

smaller than 150, both the error probability and the secrecy performance are able to achieve an acceptable value. Thus, the secrecy rate is below 0.3, which is far below the asymptotic secrecy rate $I_Y - I_Z = 0.9 - 0.4 = 0.5$*. This rate loss is caused by the nature of the finite blocklength.*

B Secure Nested Lattice Coding

Nested lattice codes have already been proved to be able to achieve the strong secrecy over the Gaussian wiretap channel [55, 56]. In the following sections, secure lattice codes are investigated and the blocklength-n secrecy performance is evaluated.

1 Preliminaries

A lattice Λ is a discrete subset of the Euclidean space $\mathbb{R}^n$ and consists of all integer linear combinations of the basis. Thus, an $n \times n$ real-valued matrix G defines a lattice Λ by

$$\Lambda = \left\{ \lambda = G\mathbf{x} : \mathbf{x} \in \mathbb{Z}^n \right\}. \tag{138}$$

The fundamental Voronoi region of Λ denoted by $\mathcal{V}(\Lambda)$ is a set of minimum Euclidean norm coset representatives. The nearest neighbour quantizer is

$$Q_{\mathcal{V}(\Lambda)}\left(\mathbf{x}\right) = \arg\min_{\lambda \in \Lambda} \left\| \mathbf{x} - \lambda \right\|. \tag{139}$$

The modulo-Λ operation corresponding to $\mathcal{V}(\Lambda)$ is defined as

$$\mathbf{x} \bmod \Lambda = \mathbf{x} - Q_{\mathcal{V}(\Lambda)}\left(\mathbf{x}\right). \tag{140}$$

Geometrically, $\mathbb{R}^n$ is the disjoint union of the Voronoi regions,

$$\mathbb{R}^n = \underset{\lambda \in \Lambda}{\cup} \left(\mathcal{V}\left(\Lambda\right) + \lambda \right). \tag{141}$$

The volume $V(\lambda)$ of $\mathcal{V}(\Lambda)$ is therefore the volume of $\mathbb{R}^n$ associated with each point of Λ. The second moment of Λ is the second moment per dimension of the random vector $\mathbf{D}$ uniformly distributed over $\mathcal{V}(\Lambda)$,

$$\sigma^2\left(\mathcal{V}\left(\Lambda\right)\right) = \frac{1}{n}\mathbb{E}\left[\left\|\mathbf{D}\right\|^2\right] = \frac{1}{nV\left(\Lambda\right)}\int_{\mathbf{x}\in\mathcal{V}(\Lambda)} \left\|\mathbf{x}\right\|^2 d\mathbf{x}. \tag{142}$$

The normalized second moment of λ is given by

$$G\left(\Lambda\right) = \frac{\sigma^2\left(\mathcal{V}\left(\Lambda\right)\right)}{V\left(\Lambda\right)^{2/n}} = \frac{1}{n}\frac{\int_{\mathbf{x}\in\mathcal{V}(\Lambda)} \left\|\mathbf{x}\right\|^2 d\mathbf{x}}{V\left(\Lambda\right)^{1+2/n}}. \tag{143}$$

A lattice Λ is said to be *good for quantization* if $G\left(\Lambda\right)$ is close to $\frac{1}{2\pi e}$. Furthermore, [84] has proved that Voronoi region $\mathcal{V}$ approaches a sphere in the sense that $G(\Lambda)$ gets close to

$\frac{1}{2\pi e}$, which means the lattices taking Voronio region as their fundamental region are good-quantization lattices. In this work, only consider the lattices with $\mathcal{V}$ as their fundamental region.

2 Nested Lattice coding

Consider a nested lattice $\Lambda_1 \subset \Lambda_2 \subset \Lambda_3$, where Λ_1 is the shaping lattice for AWGN channel. For the convenience of illustration, use $\mathcal{V}_i$ to represent $\mathcal{V}\left(\Lambda_i\right)$. The nested lattice provides two sets of codebooks:

$$\mathcal{C}_A = \left\{\Lambda_2 \cap \mathcal{V}_1\right\}, \;\; \mathcal{C}_U = \left\{\Lambda_3 \cap \mathcal{V}_2\right\}, \tag{144}$$

where the secret message U and the auxiliary message A are associated with $\mathcal{C}_U$ and $\mathcal{C}_A$, respectively, and the lattice is constructed with

$$\log\left|\mathcal{C}_U\right| = \log M_0, \;\; \log\left|\mathcal{C}_A\right| = \log M_1. \tag{145}$$

For $\lambda_U \in \mathcal{C}_U$, $\lambda_U + \Lambda_2$ is a coset of Λ_2 and

$$\Lambda_3 = \bigcup_{\lambda_U \in \mathcal{C}_U} \lambda_U {+} \Lambda_2. \tag{146}$$

The shaping lattice is chosen according to the power of the transmit symbol S_X. Alice first uniformly selects a $\lambda_A \in \mathcal{C}_A$ as a random message A and selects $\lambda_U \in \left|\mathcal{C}_U\right|$ according to the secret message U. Therefore, the transmitted codewords set at Alice is the coset of $\mathcal{C}_U$,

$$\mathcal{C} = \bigcup_{\lambda_A \in \mathcal{C}_A} \lambda_A {+} \mathcal{C}_U = \left\{\Lambda_3 \cap V_1\right\}. \tag{147}$$

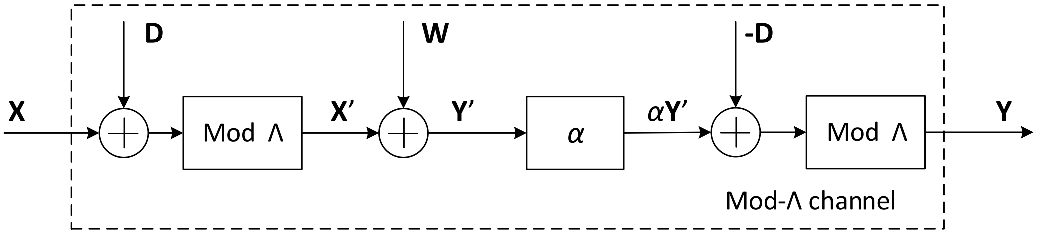


Figure 39: Mod-Λ channel

3 Mod-Λ channel with Euclidean decoder

The mod-Λ_1 channel is demonstrated by Figure 39, where the input $\mathbf{X}$ is first added with a dither $\mathbf{D}$ before the shaping lattice Λ_1. The transmitted signal is

$$\mathbf{X}' = [\mathbf{X} + \mathbf{D}] \bmod \Lambda_1. \tag{148}$$

The received signal is

$$\mathbf{Y}' = \mathbf{X}' + \mathbf{W} = [\mathbf{X} + \mathbf{D}] \bmod \Lambda_1 + \mathbf{W}, \tag{149}$$

where $\mathbf{W}$ is an n-dimension i.i.d Gaussian noise vector. The receiver applies the linear MMSE estimation for $\mathbf{Y}'$ with coefficient α and subtract the dither $\mathbf{D}$, followed by the shaping lattice modulo-Λ operation. This whole process produces an output of the mod-Λ channel as

$$\mathbf{Y} = [\alpha\mathbf{Y}' - \mathbf{D}] \bmod \Lambda_1 \tag{150}$$

$$= [[\mathbf{X} + \mathbf{D}] \bmod \Lambda_1 - \mathbf{D} - (1-\alpha)\,\mathbf{X}' + \alpha\mathbf{W}] \bmod \Lambda_1 \tag{151}$$

$$= [\mathbf{X} + \alpha\mathbf{W} - (1-\alpha)\,\mathbf{D}] \bmod \Lambda_1 \tag{152}$$

In (152), used the fact that the modulo operation is distributive so the dither cancels out, and $\mathbf{X}'$ has the same distribution as $\mathbf{D}$ according to Lemma 1.

Lemma 1. *For an uniform random variable* $\mathbf{D}$ *(over* $\mathcal{V}$*) and an arbitrary random variable* $\mathbf{X}$ *(over the same* $\mathcal{V}$ *but independent of* $\mathbf{D}$*), it gives that* $[\mathbf{X}+\mathbf{D}] \bmod \Lambda$ *is uniformly distributed over* $\mathcal{V}$ *and is independent of* $\mathbf{X}$*.*

The proof of Lemma 1 can be found in [85]. Define the equivalent noise as

$$\mathbf{W}' = \alpha \mathbf{W} - (1-\alpha)\,\mathbf{D}, \tag{153}$$

the transfer function is given by

$$f_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) = \sum_{\lambda_1 \in \Lambda_1} f_{\mathbf{W}'}\left(\mathbf{y} + \lambda_1 - \mathbf{x}\right). \tag{154}$$

To obtain the Euclidean decoder, apply two simplification methods adopted in [85]. First, approximate $\mathbf{W}'$ as a n-dimensional Gaussian random variable using the following Lemma.

Lemma 2. *For a* n*-dimensional Gaussian random variable* $\mathbf{X} \sim N\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$ *and any* n*-dimension random variable* $\mathbf{Y}$*, to minimize the KL divergence between* $\mathbf{Y}$ *and* $\mathbf{X}$

$$\min_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} \mathrm{KL}\left(f_{\mathbf{Y}} \,\|f_{\mathbf{X}}\right), \tag{155}$$

the optimal solution is

$$\boldsymbol{\mu} = \mathbb{E}\left[\mathbf{Y}\right]_{f_{\mathbf{Y}}}, \quad \boldsymbol{\Sigma} = \mathbb{E}\left[\mathbf{Y} \otimes \mathbf{Y}^*\right]_{f_{\mathbf{Y}}}, \tag{156}$$

where $\otimes$ *is Kronecker product and* $\mathbf{Y}^*$ *is the conjugate transpose of* $\mathbf{Y}$*.*

Lemma 2 is proved in Appendix A. Using Lemma 2 to minimize KL divergence of the Gaussian variable $N\left(\mu_{W'}, \sigma_{W'}^2\right)$ and W', it gives

$$\mu_{W'^2} = E\left[W'\right] = \alpha E\left[W\right] - \left(1-\alpha\right) E\left[D\right] = 0; \tag{157}$$

$$\sigma^2_{W'^2} = E\left[W'^2\right] = \alpha^2 E\left[W^2\right] + (1-\alpha)^2 E\left[D^2\right] = \alpha^2\sigma^2_W + (1-\alpha)^2 S_X. \tag{158}$$

Since the elements in both vector $\mathbf{W}$ and $\mathbf{U}$ are i.i.d., the PDF of $\mathbf{W}'$ is approximated as

$$\tilde{f}_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) = \left(2\pi\sigma^2_{W'}\right)^{-\frac{n}{2}} \sum_{\lambda_1\in\Lambda_1} e^{-\frac{\|\mathbf{y}+\lambda_{\mathbf{1}}-\mathbf{x}\|^2}{2\sigma^2_{W'}}}. \tag{159}$$

Then, the largest term is

$$\tilde{\tilde{f}}_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) = \max_{\lambda_1\in\Lambda_1}\left(2\pi\sigma^2_{W'}\right)^{-\frac{n}{2}} e^{-\frac{\|\mathbf{y}+\lambda_{\mathbf{1}}-\mathbf{x}\|^2}{2\sigma^2_{W'}}} = \left(2\pi\sigma^2_{W'}\right)^{-\frac{n}{2}} e^{-\frac{\|[\mathbf{y}-\mathbf{x}] \bmod \Lambda_1\|^2}{2\sigma^2_{W'}}}. \tag{160}$$

Based on (160), the decoder will apply the maximum a posterior (MAP) decoding as

$$\max_{\mathbf{x}\in\mathcal{V}_1} \tilde{\tilde{f}}_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right), \tag{161}$$

which is equivalent to the Euclidean decoding

$$\min_{\mathbf{x}\in\mathcal{V}_1} \|[\mathbf{y}-\mathbf{x}] \bmod \Lambda\|^2 = \min_{\mathbf{x}\in\mathcal{V}_1} \|\mathbf{w}\left(\mathbf{y},\mathbf{x}\right)\|^2, \mathbf{w}\in\mathcal{V}_{\mathbf{1}}. \tag{162}$$

Thus, the Mod-Λ Euclidean decoder has the following equivalent transfer probability function as

$$f^{\text{mod}}_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) = \frac{\tilde{\tilde{f}}_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right)}{\int\limits_{\mathbf{y}\in\mathcal{V}_1} \tilde{\tilde{f}}_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) d\mathbf{y}} = \frac{\left(2\pi\sigma^2_{W'}\right)^{-\frac{n}{2}} e^{-\frac{\|\mathbf{w}(\mathbf{y},\mathbf{x})\|^2}{2\sigma^2_{W'}}}}{\int\limits_{\mathbf{w}\in\mathcal{V}_1} \left(2\pi\sigma^2_{W'}\right)^{-\frac{n}{2}} e^{-\frac{\|\mathbf{w}\|^2}{2\sigma^2_{W'}}} d\mathbf{w}}, \mathbf{w}\in\mathcal{V}_1. \tag{163}$$

C Secrecy performance of nested lattice codes

To facilitate the discussion, first introduce function Q_n.

Definition 2. *The n-dimension Q_n function is defined as*

$$Q_n\left(\rho^2\right) = (2\pi)^{-\frac{n}{2}} \int_{\mathbf{w}\in\Theta(\rho^2)} \exp\left(-\frac{\|\mathbf{w}\|^2}{2}\right) d\mathbf{w}, \tag{164}$$

where

$$\Theta\left(\rho^2\right) = \left\{\mathbf{w}\left|\|\mathbf{w}\|^2 \geq n\rho^2\right.\right\}. \tag{165}$$

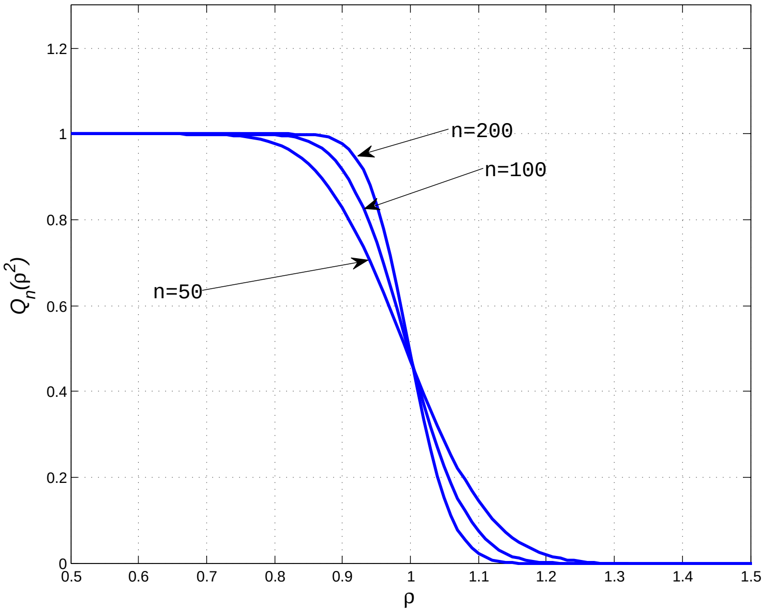


Figure 40: Q_n function

Plot Q_n in Figure 40.

Lemma 3. *Q_n can be obtained numerically by*

$$Q_n\left(\rho^2\right) = \left(\frac{n}{\pi}\right)^{\frac{1}{2}} \left(\frac{e}{n}\right)^{\frac{n}{2}} \int_{\sqrt{n}\rho}^{\infty} t^{n-1} \exp\left(-\frac{t^2}{2}\right) dt. \tag{166}$$

Lemma 3 is proved in Appendix B.

Lemma 4. *The asymptotic characteristic of Q_n is*

$$\lim_{n\to\infty} Q_n\left(\rho^2\right) = 1 \ \text{if}\rho < 1 \ \text{ and } \ \lim_{\mathrm{n}\to\infty} \mathrm{Q}_\mathrm{n}\left(\rho^2\right) = 0 \ \text{if}\rho > 1. \tag{167}$$

Lemma 4 is proved in Appendix C.

Proposition 1.

$$1 - Q_n\left(\rho^2\right) \in o\left(n^{-1}\right), \tag{168}$$

which gives

$$\lim_{n\to\infty} n\left(1 - Q_n\left(\rho^2\right)\right) = 0 \quad \text{when } \rho < 1. \tag{169}$$

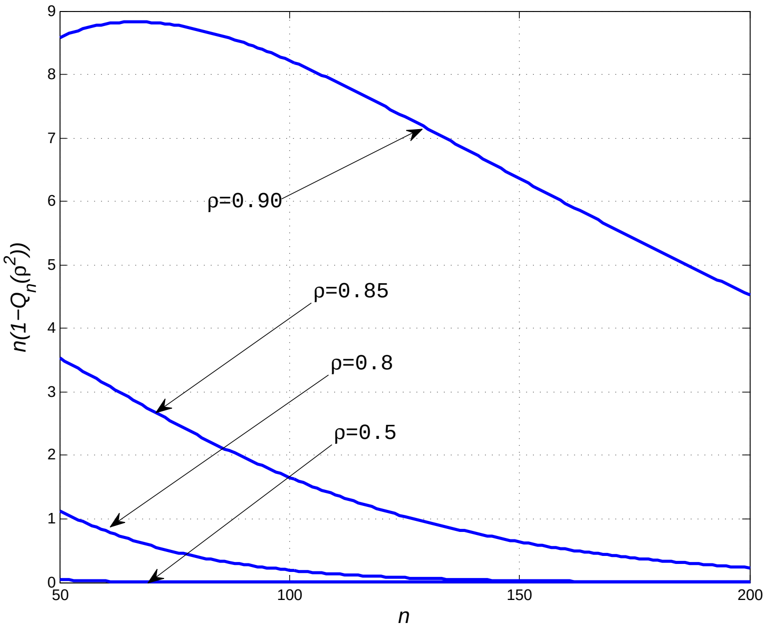


Figure 41: $n\left(1-Q_n\left(\rho^2\right)\right)$ vs n

Proposition 1 is illustrated by Figure 41.

Lemma 5. *For a random variable* $\mathbf{X}$*, if* $\mathbf{X}$ *is uniformly distributed in a* n*-dimension sphere with power* S_X *per dimension, the radius* ρ *of this sphere is given by*

$$\rho^2=(n+2)\,S_X. \tag{170}$$

Lemma 5 is proved in Appendix D. When n is sufficiently large, the Voroni region can be expressed as $\mathcal{V}=\left\{\mathbf{x}\left|\|\mathbf{x}\|^2\leq(n+2)\,S_X\right.\right\}$, and $\rho^2\approx nS_X$. Using Definition 2 and Lemma 5, (163) is re-written as

$$f_{\mathbf{Y}|\mathbf{X}}^{\text{mod}}\left(\mathbf{y}|\mathbf{x}\right)=\frac{\left(2\pi\sigma_{W'}^2\right)^{-\frac{n}{2}}e^{-\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2\sigma_{W'}^2}}}{1-Q_n\left(\frac{S_X}{\sigma_{W'}^2}\right)}\mathbf{y}\in V,\mathbf{x}\in V. \tag{171}$$

Lemma 6. *The capacity of the Mod-*Λ *channel is achieved when* $\mathbf{X}$ *is uniformly distributed in* $\mathcal{V}_1$ *as*

$$\arg\max_{f_{\mathbf{X}}(\mathbf{x})} I\left(\mathbf{Y};\mathbf{X}\right)=\frac{1}{V_1},\;\;\mathbf{x}\in\mathcal{V}_1. \tag{172}$$

Lemma 6 is proved in Appendix E. As Bob adopts the Mod-Λ Euclidean decoder, assume that Alice always intends to maximize the mutual information to Bob. Thus, X is assumed to be uniformly distributed over $\mathcal{V}_1$ according to Lemma 6. Then the mutual information density of the Mod-Λ Euclidean decoder is obtained as

$$\begin{aligned} i^{\text{ mod }}(\mathbf{z};\mathbf{x}) &= \log \frac{f_{\mathbf{Z}|\mathbf{X}}^{\text{mod}}\left(\mathbf{z}|\mathbf{x}\right)}{f_{\mathbf{Z}}\left(\mathbf{z}\right)} \qquad (173)\\ &= \log\left(V_1\left(\left(2\pi\sigma^2_{W_z{}'}\right)^{-\frac{n}{2}}\exp\left(-\frac{\left\|\mathbf{z}-\mathbf{x}\right\|^2}{2\sigma^2_{W_z{}'}}\right)\right)\Big/\left(1-Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right)\right) \qquad (174)\\ &= \frac{n}{2}\log\left(\frac{eS_X}{\sigma^2_{W_z{}'}}\right)-\frac{\left\|\mathbf{w}\right\|^2}{2\sigma^2_{W_z{}'}}\log e-\frac{1}{2}\log\pi n-\log\left(1-Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right). \qquad (175)\end{aligned}$$

Theorem 8. *There exists nested lattice codes with the Mod-Λ Euclidean decoder that can achieve the Gaussian channel capacity.*

Proof. Nested lattice codes with the the Mod-Λ Euclidean decoder have already been proved to be capacity-achieving by [85]. Alternatively, this can also be proved using the mutual information density.

Lemma 7. *The MMSE coefficient of* $\mathbb{E}\left[\left\|\alpha\mathbf{y}-\mathbf{x}\right\|^2\right]$ *is* $\alpha^* = \frac{S_X}{S_X+\sigma_w^2}$.

Lemma 7 is proved in Appendix F. Using the MMSE coefficient stated by Lemma 7,

$$\begin{aligned} &\lim_{n\to\infty}\frac{1}{n}i^{\text{ mod }}(\mathbf{z};\mathbf{x}) \qquad (176)\\ &= \frac{1}{2}\log\left(\frac{eS_X}{\sigma^2_{W_z{}'}}\right)-\lim_{n\to\infty}\frac{\left\|\mathbf{w}'\right\|^2}{2n\sigma^2_{W_z{}'}}\log e-\frac{1}{2}\lim_{n\to\infty}\frac{\log\pi n}{n}-\lim_{n\to\infty}\frac{1}{n}\log Q_n\left(\frac{S_X}{\sigma^2_{W'}}\right) \qquad (177)\\ &= \frac{1}{2}\log\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)=\frac{1}{2}\log\left(1+\frac{S_X}{\sigma^2_{W}}\right). \qquad (178)\end{aligned}$$

□

1 Blocklength-n secrecy performance

Now, let's analyze blocklength-n secrecy performance. First define the threshold value

$$\rho^2_{\text{mod}}\left(n, M_1\right) = \ln\left(\frac{eS_X}{\sigma^2_{W_y}}\right) - \frac{2}{n}\ln M_1 - \frac{1}{n}\ln \pi n - \frac{2}{n}\log\left(1 - Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right). \quad (179)$$

Then, the set of $\mathbf{W}$ satisfying $i_{\mathbf{Z}|\mathbf{X}}\left(\mathbf{z}|\mathbf{x}\right) > \log M_1$ can be rewritten as

$$\Omega_{\text{ mod}}\left(\rho^2_{\text{mod}}\left(n, M_1\right)\right) = \left\{\mathbf{w}\left|\frac{\|\mathbf{w}\|^2}{\sigma^2_{W_z{}'}} < n\rho^2_{\text{mod}}\left(n, M_1\right)\right.\right\}. \quad (180)$$

Therefore, the leak probability is

$$L_P = P\left[i^{\text{ mod}}\left(\mathbf{z};\mathbf{x}\right) > \log M_1\right] = \int_{\mathbf{w}'\in\Omega_{\text{ mod}}\cap\mathcal{V}_1}\int_{\mathbf{x}\in\mathcal{V}_1} f_{\mathbf{X}}\left(\mathbf{x}\right) f_{\mathbf{W}_z{}'}\left(\mathbf{w}'\right) d\mathbf{x}d\mathbf{w}'. \quad (181)$$

Theorem 9. *Given the MMSE coefficient* α *and* $\forall\ n > 1$*, it gives*

$$\Omega_{\text{ mod}} \cap \mathcal{V}_1 = \Omega_{\text{ mod}} \quad (182)$$

Proof. Note that

$$\Omega_{\text{ mod}} = \left\{\mathbf{w}'\left|\frac{\|\mathbf{w}'\|^2}{\sigma^2_{W_z{}'}} < n\log\left(\frac{eS_X}{\sigma^2_{W_z{}'}}\right) - \ln\pi n - \ln\left(1 - Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right)\right.\right\}. \quad (183)$$

Taking the MMSE coefficient, it gives

$$Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right) = Q_n\left(1 + \frac{S_X}{\sigma^2_{W_z}}\right) < Q_2\left(1 + \frac{S_X}{\sigma^2_{W_z}}\right) < Q_2\left(1\right) \quad (184)$$

Now, check Q_2 as

$$Q_2\left(1\right) = \frac{1}{2\pi}\int_{w_1^2+w_2^2\geq 2} e^{-\frac{w_1^2+w_2^2}{2}} dw_1 dw_2 = e^{-1}. \quad (185)$$

Then

$$\ln\left(1-Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right) > \ln\left(1-e^{-1}\right), \tag{186}$$

and

$$-\ln \pi n-\ln\left(1-Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right) < -\ln\left(\pi n\left(1-e^{-1}\right)\right) < 0. \tag{187}$$

Therefore,

$$\Omega_{\text{ mod }} \subset \left\{\mathbf{w}' \left| \frac{\|\mathbf{w}'\|^2}{n\sigma^2_{W_z{}'}} < \ln\left(\frac{eS_X}{\sigma^2_{W_z{}'}}\right)\right.\right\} = \left\{\mathbf{w}' \left| \frac{\|\mathbf{w}'\|^2}{n\sigma^2_{W_z{}'}} < 1+\ln\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right.\right\}. \tag{188}$$

Lemma 8. *For any $S_X > 0$ and the given coefficient α, it gives*

$$\frac{S_X}{\sigma^2_{W_z{}'}} > 1+\ln\frac{S_X}{\sigma^2_{W_z{}'}}. \tag{189}$$

Lemma 8 is proved in Appendix G. Using Lemma 8, it gives

$$\Omega_{\text{ mod }} \subset \left\{\mathbf{w}' \left| \frac{\|\mathbf{w}'\|^2}{n\sigma^2_{W_z{}'}} < 1+\ln\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right.\right\} \subset \left\{\mathbf{w}' \left| \frac{\|\mathbf{w}'\|^2}{n\sigma^2_{W_z{}'}} < \frac{S_X}{\sigma^2_{W_z{}'}}\right.\right\} = \left\{\mathbf{w}' \left| \|\mathbf{w}'\|^2 < nS_X\right.\right\}. \tag{190}$$

According to Lemma 5, (190) becomes

$$\Omega_{\text{ mod }} \subset \left\{\mathbf{w}' \left| \|\mathbf{w}'\|^2 < nS_X\right.\right\} \subset \left\{\mathbf{w}' \left| \|\mathbf{w}'\|^2 < (n+2)\, S_X\right.\right\} = \mathcal{V}_1. \tag{191}$$

□

Using Theorem 9, the leak probability is simplified as

$$L_P = \int_{\mathbf{w}\in\Omega_{\text{ mod }}} \int_{\mathbf{x}\in\mathcal{V}_1} f_{\mathbf{X}}(\mathbf{x})\, f_{\mathbf{W}_z{}'}(\mathbf{w}')\, d\mathbf{x}d\mathbf{w}' = 1-Q_n\left(\rho^2_{\text{mod }}(n, M_1)\right). \tag{192}$$

Similarly, the average leakage is

$$L_A = \mathbb{E}\left[\left[i\left(\mathbf{z};\mathbf{x}|u\right) - \log M_1\right]^+\right] \tag{193}$$

$$= \left(2\pi\sigma^2_{W_z'}\right)^{-\frac{n}{2}} \int_{\mathbf{w}'\in\Omega_{\text{ mod }}\cap\mathcal{V}_1} \left(\frac{n}{2}\rho^2_{Eu}\left(n, M_1\right)\log e - \frac{\left\|\mathbf{w}'\right\|^2}{2\sigma^2_{W_z'}}\log e\right) e^{-\frac{\left\|\mathbf{w}\right\|^2}{2\sigma^2_{W_z'}}} d\mathbf{w}' \tag{194}$$

$$= \left(2\pi\sigma^2_{W_z'}\right)^{-\frac{n}{2}} \int_{\mathbf{w}'\in\Omega_{\text{ mod }}} \left(\frac{n}{2}\rho^2_{Eu}\left(n, M_1\right)\log e - \frac{\left\|\mathbf{w}'\right\|^2}{2\sigma^2_{W_z'}}\log e\right) e^{-\frac{\left\|\mathbf{w}\right\|^2}{2\sigma^2_{W_z'}}} d\mathbf{w}' \tag{195}$$

$$= \left(2\pi\sigma^2_{W_z'}\right)^{-\frac{n}{2}} \log e \int_{\mathbf{w}'\in\Omega_{\text{ mod }}} -\frac{\left\|\mathbf{w}'\right\|^2}{2\sigma^2_{W_z'}} e^{-\frac{\left\|\mathbf{w}'\right\|^2}{2\sigma^2_{W_z'}}} d\mathbf{w}' + \tag{196}$$

$$\frac{\log e}{2}\rho^2_{\text{mod }}\left(n, M_1\right) n\left(1 - Q_n\left(\rho^2_{\text{mod }}\left(n, M_1\right)\right)\right). \tag{197}$$

Note that (196) can be further calculated as

$$\left(2\pi\sigma^2_{W_z'}\right)^{-\frac{n}{2}} \log e \int_{\mathbf{w}'\in\Omega_{\text{ mod }}} -\frac{\left\|\mathbf{w}\right\|^2}{2\sigma^2_{W_z'}} e^{-\frac{\left\|\mathbf{w}'\right\|^2}{2\sigma^2_{W_z'}}} d\mathbf{w}' \tag{198}$$

$$= -\left(2\pi\sigma^2_{W_z'}\right)^{-\frac{n}{2}} \log e \int_{\rho<\rho_{\text{ mod }}} \left(\frac{n}{\pi}\right)^{\frac{1}{2}} \left(\frac{2\pi e}{n}\right)^{\frac{n}{2}} \rho^{n-1} \frac{\rho^2}{2\sigma^2_{W_z'}} e^{-\frac{\rho^2}{2\sigma^2_{W_z'}}} d\rho \tag{199}$$

$$= -\frac{\log e}{2e}\left(n+2\right)\left(\frac{n}{n+2}\right)^{\frac{1}{2}}\left(\frac{n+2}{n}\right)^{\frac{n}{2}}\left(2\pi\sigma^2_{W_z'}\right)^{-\frac{n+2}{2}} \times \tag{200}$$

$$\int_{\rho<\rho_{\text{ mod }}} \left(\frac{n+2}{\pi}\right)^{\frac{1}{2}}\left(\frac{2\pi e}{n+2}\right)^{\frac{n+2}{2}} \rho^{n+1} e^{-\frac{\rho^2}{2\sigma^2_{W_z'}}} d\rho \tag{201}$$

$$= -\frac{\log e}{2e}\left(n+2\right)\left(\frac{n}{n+2}\right)^{\frac{1}{2}}\left(\frac{n+2}{n}\right)^{\frac{n}{2}}\left(1 - Q_{n+2}\left(\frac{n}{n+2}\rho^2_{\text{mod }}\left(n, M_1\right)\right)\right) \tag{202}$$

$$= -\frac{\log e}{2}\left(n+2\right)\left(1 - Q_{n+2}\left(\frac{n}{n+2}\rho^2_{\text{mod }}\left(n, M_1\right)\right)\right), \tag{203}$$

where used the fact $\left(\frac{n}{n+2}\right)^{\frac{1}{2}} \approx 1$ and $\left(\frac{n+2}{n}\right)^{\frac{n}{2}} \approx e$ for a large n. Finally, L_A is calculated as

$$\begin{aligned} L_A =& \frac{\log e}{2}\rho^2_{\text{mod }}\left(n, M_1\right) n\left(1 - Q_n\left(\rho^2_{\text{mod }}\left(n, M_1\right)\right)\right) \\ &- \frac{\log e}{2}\left(n+2\right)\left(1 - Q_{n+2}\left(\frac{n}{n+2}\rho^2_{\text{mod }}\left(n, M_1\right)\right)\right). \end{aligned} \tag{204}$$

Example 3. *(The Mod-Λ Euclidean decoder at Eve)*

Assume Eve adopts the Mod-Λ Euclidean decoder with a received SNR of 10dB, evaluate L_P and compared it with the result of the blocklength-n Gaussian channel in Section III. Figure 42 shows L_P versus n, when $\log M_1 = 0.93I_Z$, $0.98I_Z$, I_Z, $1.02I_Z$ *and* $1.07I_Z$. *From Figure 42, it gives the following observation. For the secure lattice codes with finite blocklengths, L_P is below the theoretical value when* $\log M_1$ *is greater than the channel capacity, and L_P will be greater than its theoretical value when* $\log M_1$ *becomes smaller than the channel capacity. Nevertheless, the observation for L_P doesn't hold for L_A. Figure 43 shows L_A versus n, clearly, L_A of the Mod-Λ Euclidean decoder is always smaller than the theoretical L_A of the n-dimensional Gaussian channel.*

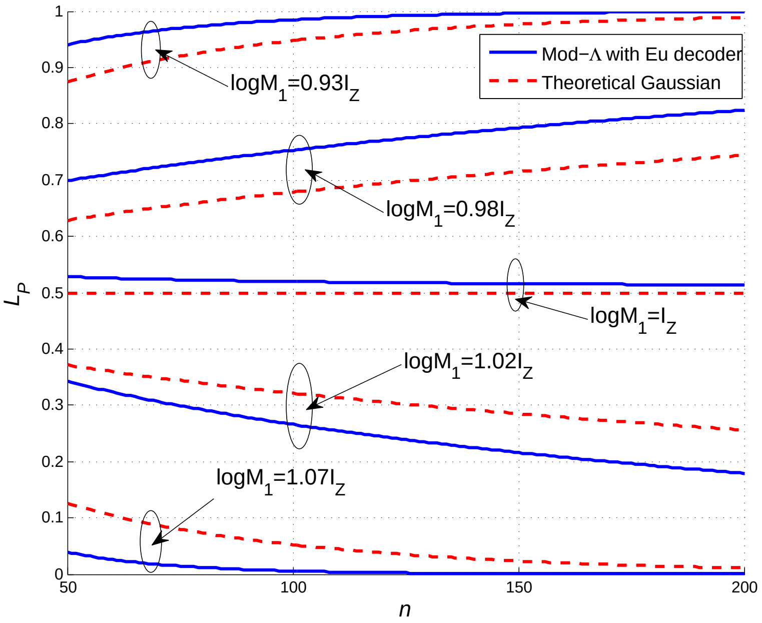


Figure 42: Leak probability L_P vs blocklength n, Comparison of the Mod-Λ Euclidean decoder and the blocklength-n Gaussian channel.

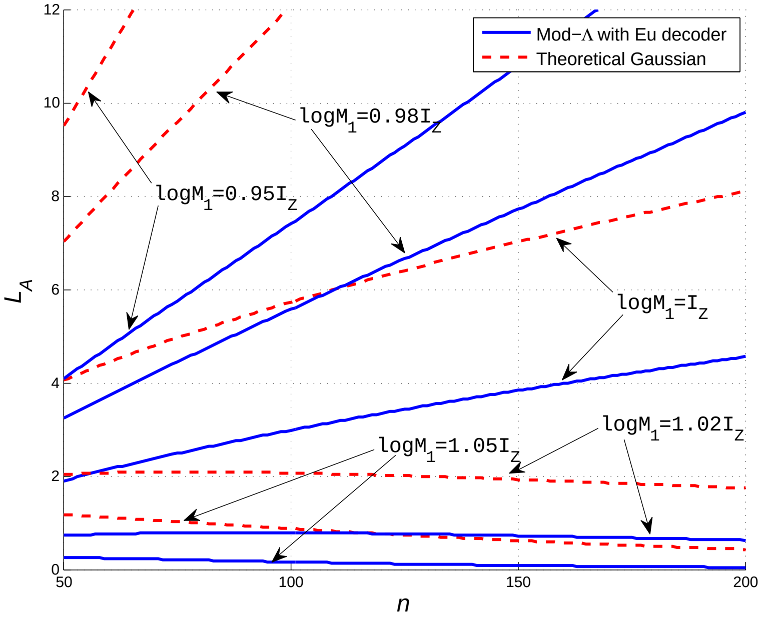


Figure 43: Average leakage L_A vs blocklength n, Comparison of the Mod-Λ Euclidean decoder and the blocklength-n Gaussian channel.

2 Asymptotic analysis

In this section, the results is re-examined when n goes to infinity. From Lemma 4, when $\rho^2_{\text{mod}}\,(n, M_1) < 1$, the asymptotic L_P is given by

$$\lim_{n\to\infty} L_P = 1 - \lim_{n\to\infty} Q_n\left(\rho^2_{\text{mod}}\,(n, M_1)\right) = 0. \tag{205}$$

Similarly, from Proposition 1, the asymptotic L_A is given by

$$\begin{aligned}\lim_{n\to\infty} L_A = \frac{\log e}{2} \times \Big(&\lim_{n\to\infty} \rho^2_{\text{mod}}\,(n, M_1)\, n \left(1 - Q_n\left(\rho^2_{\text{mod}}\,(n, M_1)\right)\right) - \\ &\lim_{n\to\infty} (n+2)\left(1 - Q_{n+2}\left(\frac{n}{n+2}\rho^2_{\text{mod}}\,(n, M_1)\right)\right)\Big) \qquad (206)\\ = 0 - 0 = 0.&\end{aligned}$$

When $\rho^2_{\text{mod}}\,(n, M_1) > 1$, the asymptotic L_P becomes

$$\lim_{n\to\infty} L_P = 1 - \lim_{n\to\infty} Q_n\left(\rho^2_{\text{mod}}\,(n, M_1)\right) = 1, \tag{207}$$

and the asymptotic L_A becomes

$$\lim_{n\to\infty} L_A = \frac{\log e}{2}\lim_{n\to\infty}\rho^2_{\text{mod}}\left(n, M_1\right) n - \frac{\log e}{2}\lim_{n\to\infty}\left(n+2\right) \tag{208}$$

$$= \frac{\log e}{2}\left(\rho^2_{\text{mod}}\left(n, M_1\right) - 1\right)\lim_{n\to\infty} n - O\left(1\right) = \infty, \tag{209}$$

where Lemma 4 and the fact that $\lim\limits_{n\to\infty} \frac{n}{n+2}\rho^2_{Eu}\left(n, M_1\right) > 1$ when $\rho^2_{\text{mod}}\left(n, M_1\right) > 1$ are used.

From (205)-(209), the asymptotic performance threshold is $\rho^2_{\text{mod}} = 1$. Compared with Theorem 6, it gives

$$\eta_{R_s} = \frac{\log M_1}{nH\left(X\right)\left(1-R_s\right)} = \frac{I_Z + \left(1-\rho^2_{\text{mod}}\right)\log e}{H\left(X\right)\left(1-R_s\right)}. \tag{210}$$

Note that $\rho^2_{\text{mod}} = 1$ is equivalent to $\eta_{R_s} = \frac{I_Z}{H(X)(1-\eta)}$, which is exactly the same as the strong secrecy threshold obtained. This means, when $H\left(A\right) > I_Z$, the nested lattice codes with the Mod-Λ Euclidean decoder can achieve the strong secrecy. This conclusion can also be verified by $\rho^2_{\text{mod}}\left(n, M_1\right)$. Using the threshold $\rho^2_{\text{mod}}\left(n, M_1\right) = 1$ and the fact $\log\left(1 - Q_n\left(\frac{S_X}{\sigma^2_{W_z{}'}}\right)\right) \sim O\left(1\right)$ for a sufficiently large n, it gives

$$\frac{1}{n}\log M_1 = \frac{1}{2}\log\left(1 + \frac{S_X}{\sigma^2_{W_z}}\right) - \frac{1}{n}O\left(\log n\right), \tag{211}$$

which means that the auxiliary message A achieves the capacity of wiretap channel as $n \to \infty$.

3 Trade-off between the reliability and the secrecy

In [85], Erez proved there exists lattice codes whose decoding error probability with the Mod-Λ Euclidean decoder satisfies

$$P_e \leq e^{-n\left(E_P\left(e^{2\left(I_Y - I\right)}\right) - O(1)\right)}, \tag{212}$$

where E_p is the Poltyrev exponent

$$E_P(x) = \begin{cases} \frac{1}{2}\left((x-1)-\ln x\right) & 1 < x \le 2 \\ \frac{1}{2}\ln\frac{ex}{4} & 2 \le x \le 4 \\ \frac{x}{8} & x \ge 4 \end{cases} .$$

Therefore, for a required M_0 and ϵ (the upper bound of P_e), the minimal M_1 is given by

$$\frac{1}{n}\log M_1 = I_Y - \frac{1}{n}\log M_0 - \frac{1}{2}\ln\left(E_p^{-1}\left(-\frac{\ln\epsilon}{n}\right)\right). \tag{213}$$

Plug (213) into (179), ρ_{mod} can be rewritten as a function of M_0 and ϵ.

$$\rho^2_{\text{mod}}\left(n, M_0, \epsilon\right) = 1 - \frac{\ln 4}{n}\left(I_Y - I_Z - \log M_0\right) + \frac{\ln 2}{n}\ln\left(E_P^{-1}\left(-\frac{\ln\epsilon}{n}\right)\right) - \frac{\ln \pi n}{n}. \tag{214}$$

Then, the trade-off between ϵ and the secrecy performance can be expressed as

$$L_P^*\left(M_0, \epsilon, n\right) = 1 - Q_n\left(\rho^2_{\text{Un}}\left(M_0, \epsilon, n\right)\right); \tag{215}$$

$$\begin{aligned} L_A^*\left(M_0, \epsilon, n\right) =& \frac{\log e}{2}\rho^2_{\text{Un}}\left(M_0, \epsilon, n\right) n\left(1 - Q_n\left(\rho^2_{\text{Un}}\left(M_0, \epsilon, n\right)\right)\right) \\ &- \frac{\log e}{2}(n+2)\left(1 - Q_{n+2}\left(\frac{n}{n+2}\rho^2_{\text{Un}}\left(M_0, \epsilon, n\right)\right)\right). \end{aligned} \tag{216}$$

Example 4. *(Mod-Λ channel with the Euclidean decoder at both Bob and Eve)*

Assume Bob and Eve both adopt the module operation and the Euclidean decoder with received SNRs being 20dB and 10dB respectively, and the blocklength n is 200. From the asymptotic result, the maximal secrecy entropy is $n\left(I_Y - I_Z\right) = 330.9$. However, the secrecy entropy for the finite blocklength is much smaller than 330.9. Figure 44 and Figure 45 show L_p and L_A vs ϵ. First of all, only when $\log M_0$ is smaller than 150 can L_P, L_A and ϵ simultaneously achieve acceptable values, which clearly shows that the finite blocklength penalty is quiet significant and cannot be ignored. Another observation from

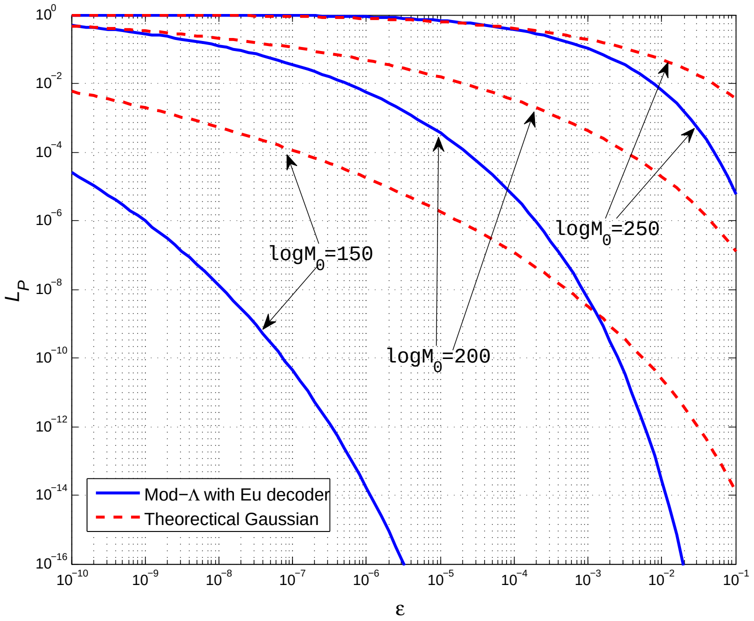


Figure 44: Leak probability L_P vs main channel error probability ϵ. Comparison of the Mod-Λ Euclidean decoder and the blocklength-n Gaussian channel.

Figure 44 and Figure 45 is that the performance of the Mod-λ Euclidean decoder is close to the blocklength-n Gaussian channel, especially when a smaller ϵ is required at Bob, which indicates that the Mod-Λ Euclidean decoder is acceptable for Eve if Alice emphasizes on the reliability more than the secrecy.

D Conclusion

In this Chapter, a practical approach to analyze secure lattice codes with finite blocklengths was provided. The idea to find computationally trackable analysis is to utilize the second-order. The idea and technique presented can be applied to various scenarios for the analysis of the finite-blocklength secrecy. For a specific code design, it is necessary to analyze its trade-off between the reliability and the secrecy.

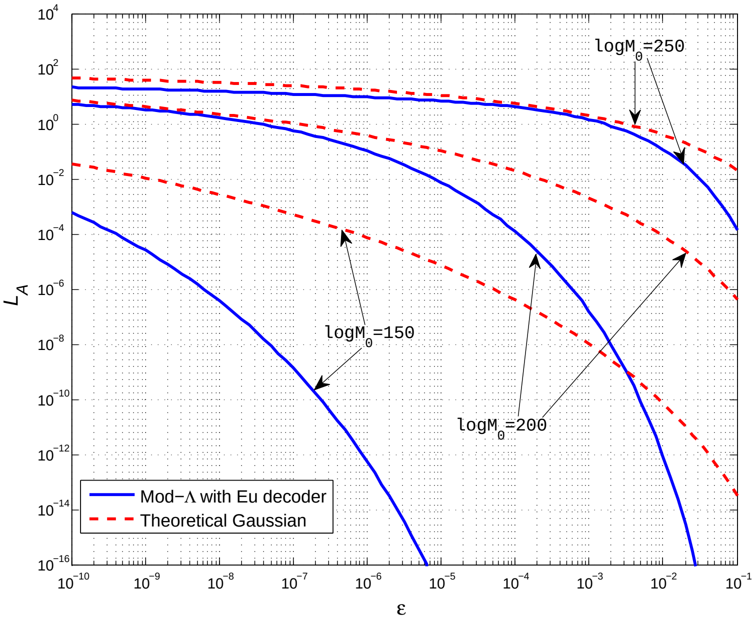


Figure 45: Average leakage L_A vs main channel error probability ϵ. Comparison of the Mod-Λ Euclidean decoder and the blocklength-n Gaussian channel.

REFERENCES

[1] R. H. Walden, "Analog-to-digital converter survey and analysis," *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 4, pp. 539–550, 1999.

[2] B. Murmann, "Adc performance survey 1997-2014," *[Online].Available:http://www.stanford.edu/bmurmann/adcsurvey.html*, 2014.

[3] T. Sundström, B. Murmann, and C. Svensson, "Power dissipation bounds for high-speed nyquist analog-to-digital converters," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 56, no. 3, pp. 509–518, 2009.

[4] J. Singh, S. Ponnuru, and U. Madhow, "Multi-gigabit communication: the adc bottleneck," in *Ultra-Wideband, 2009. ICUWB 2009. IEEE International Conference on*, pp. 22–27, IEEE, 2009.

[5] J. Mo and R. W. Heath, "High snr capacity of millimeter wave mimo systems with one-bit quantization," in *Information Theory and Applications Workshop (ITA), 2014*, pp. 1–5, IEEE, 2014.

[6] J. Mo and R. W. Heath Jr, "Capacity analysis of one-bit quantized mimo systems with transmitter channel state information," *arXiv preprint arXiv:1410.7353*, 2014.

[7] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Transactions on Communications*, vol. 57, no. 12, pp. 3629–3639, 2009.

[8] A. Mezghani and J. A. Nossek, "On ultra-wideband mimo systems with 1-bit quantized outputs: Performance analysis and input optimization," in *2007 IEEE International Symposium on Information Theory*, pp. 1286–1289, IEEE, 2007.

[9] A. Mezghani and J. A. Nossek, "Analysis of rayleigh-fading channels with 1-bit quantized output," in *2008 IEEE International Symposium on Information Theory*, pp. 260–264, IEEE, 2008.

[10] A. Mezghani and J. A. Nossek, "Analysis of 1-bit output noncoherent fading channels in the low snr regime," in *2009 IEEE International Symposium on Information Theory*, pp. 1080–1084, IEEE, 2009.

[11] W. R. Bennett, "Spectra of quantized signals," *Bell System Technical Journal*, vol. 27, no. 3, pp. 446–472, 1948.

[12] C. Risi, D. Persson, and E. G. Larsson, "Massive mimo with 1-bit adc," *arXiv preprint arXiv:1404.7736*, 2014.

[13] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, “One-bit massive mimo: Channel estimation and high-order modulations,” *arXiv preprint arXiv:1504.04540*, 2015.

[14] 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 Transactions on Communications*, vol. 64, no. 5, pp. 2005–2018, 2016.

[15] 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,” 2015.

[16] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath Jr, “Channel estimation in millimeter wave mimo systems with one-bit quantization,” in *Proc. Asilomar Conf. on Signals, Systems and Computers*, 2014.

[17] C. Cao, H. Li, Z. Hu, and H. Zeng, “1-bit relay cluster for long-distance transmission,” *IEEE Communications Letters*, 2016.

[18] C. Studer and G. Durisi, “Quantized massive mu-mimo-ofdm uplink,” *IEEE Transactions on Communications*, vol. 64, pp. 2387–2399, June 2016.

[19] N. Liang and W. Zhang, “Mixed-adc massive mimo uplink in frequency-selective channels,” *arXiv preprint arXiv:1601.02082*, 2016.

[20] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, “Serial concatenated trellis coded modulation with iterative decoding: Design and performance,” 1997.

[21] S. Benedetto and G. Montorsi, “Unveiling turbo codes: Some results on parallel concatenated coding schemes,” *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 409–428, 1996.

[22] X. Wang and H. V. Poor, “Iterative (turbo) soft interference cancellation and decoding for coded cdma,” *IEEE Transactions on communications*, vol. 47, no. 7, pp. 1046–1061, 1999.

[23] C. Douillard, M. Jézéquel, C. Berrou, D. Electronique, A. Picart, P. Didier, and A. Glavieux, “Iterative correction of intersymbol interference: Turbo-equalization,” *European transactions on telecommunications*, vol. 6, no. 5, pp. 507–511, 1995.

[24] G. Bauch and V. Franz, “A comparison of soft-in/soft-out algorithms for turbo detection,” in *Proc. Int. Conf. Telecomm*, pp. 259–263, Citeseer, 1998.

[25] A. Anastasopoulos and K. M. Chugg, “Iterative equalization/decoding of tcm for frequency-selective fading channels,” in *Signals, Systems &amp; Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*, vol. 1, pp. 177–181, IEEE, 1997.

[26] M. Tuchler, R. Koetter, and A. C. Singer, “Turbo equalization: principles and new results,” *IEEE transactions on communications*, vol. 50, no. 5, pp. 754–767, 2002.

[27] M. Tuchler, A. C. Singer, and R. Koetter, “Minimum mean squared error equalization using a priori information,” *IEEE Transactions on Signal processing*, vol. 50, no. 3, pp. 673–683, 2002.

[28] A. Glavieux, C. Laot, and J. Labat, “Turbo equalization over a frequency selective channel,” in *Proc. Int. Symp. Turbo Codes*, vol. 962102, 1997.

[29] R. El Chall, F. Nouvel, M. Hélard, and M. Liu, “Iterative receivers combining mimo detection with turbo decoding: performance-complexity trade-offs,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–19, 2015.

[30] A. Tomasoni, M. Ferrari, D. Gatti, F. Osnato, and S. Bellini, “A low complexity turbo mmse receiver for w-lan mimo systems,” in *2006 IEEE International Conference on Communications*, vol. 9, pp. 4119–4124, IEEE, 2006.

[31] P. Shang, S. Kim, and K. Choi, “Soft mmse receiver for turbo coded mimo system,” in *2011 IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 471–475, IEEE, 2011.

[32] S. Salari, M. Ardebilipour, M. Ahmadian, J.-P. Cances, and V. Meghdadi, “Turbo receiver design with carrier-frequency offset estimation for ldpc-coded mimo ofdm systems,” in *The 9th International Conference on Advanced Communication Technology*, vol. 3, pp. 1911–1915, IEEE, 2007.

[33] J. Guo, Y. Shang, S. Ren, and H. Xiang, “A turbo receiver combined with cfo compensation in frequency-domain for ofdma uplink,” *Signal Processing*, vol. 90, no. 7, pp. 2253–2264, 2010.

[34] T. Kang and R. A. Iltis, “Iterative carrier frequency offset and channel estimation for underwater acoustic ofdm systems,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 9, pp. 1650–1661, 2008.

[35] J. Boutros and G. Caire, “Iterative multiuser joint decoding: Unified framework and asymptotic analysis,” *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1772–1793, 2002.

[36] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, “The factor graph approach to model-based signal processing,” *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, 2007.

[37] G. Colavolpe and G. Germi, “On the application of factor graphs and the sum-product algorithm to isi channels,” *IEEE Transactions on Communications*, vol. 53, no. 5, pp. 818–825, 2005.

[38] C. Novak, G. Matz, and F. Hlawatsch, “A factor graph approach to joint iterative data detection and channel estimation in pilot-assisted idma transmissions,” in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2697–

2700, IEEE, 2008.

[39] A. P. Worthen and W. E. Stark, “Unified design of iterative receivers using factor graphs,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 843–849, 2001.

[40] D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18914–18919, 2009.

[41] D. L. Donoho, A. Maleki, and A. Montanari, “Message passing algorithms for compressed sensing: I. motivation and construction,” in *Information Theory Workshop (ITW), 2010 IEEE*, pp. 1–5, IEEE, 2010.

[42] M. A. Maleki, *Approximate message passing algorithms for compressed sensing*. Stanford University, 2010.

[43] S. Rangan, “Generalized approximate message passing for estimation with random linear mixing,” in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pp. 2168–2172, IEEE, 2011.

[44] C. Cao, H. Li, and Z. Hu, “An amp based decoder for massive mu-mimo-ofdm with low-resolution adcs,” *ICNC*, 2017.

[45] E. Nash, Y. Toh, and G. Hendrickson, “Single chip realizes direct-conversion rx,” *MICROWAVES & RF*, pp. 55–67, 2002.

[46] J. Mo and R. W. Heath, “Capacity analysis of one-bit quantized mimo systems with transmitter channel state information,” *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5498–5512, 2015.

[47] A. D. Wyner, “The wire-tap channel,” *Bell System Technical Journal, The*, vol. 54, no. 8, pp. 1355–1387, 1975.

[48] I. Csiszár and J. Korner, “Broadcast channels with confidential messages,” *Information Theory, IEEE Transactions on*, vol. 24, no. 3, pp. 339–348, 1978.

[49] U. Maurer and S. Wolf, “Information-theoretic key agreement: From weak to strong secrecy for free,” in *Advances in Cryptology EUROCRYPT 2000*, pp. 351–368, Springer, 2000.

[50] M. Bloch and J. Barros, *Physical-layer security: from information theory to security engineering*. Cambridge University Press, 2011.

[51] I. Csiszár, “Almost independence and secrecy capacity,” *Problemy Peredachi Informatsii*, vol. 32, no. 1, pp. 48–57, 1996.

[52] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, “Applications of ldpc codes to the wiretap channel,” *Information Theory, IEEE Transactions on*, vol. 53, no. 8, pp. 2933–2945, 2007.

[53] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *Information Theory, IEEE Transactions on*, vol. 59, no. 12, pp. 8077–8098, 2013.

[54] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *Information Theory, IEEE Transactions on*, vol. 57, no. 10, pp. 6428–6443, 2011.

[55] L.-C. Choo, C. Ling, and K.-K. Wong, "Achievable rates for lattice coded gaussian wiretap channels," in *2011 IEEE International Conference on Communications Workshops (ICC)*, 2011.

[56] C. Ling, L. Luzzi, J. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the gaussian wiretap channel," 2012.

[57] V. Rathi, M. Andersson, R. Thobaben, J. Kliewer, and M. Skoglund, "Performance analysis and design of two edge-type ldpc codes for the bec wiretap channel," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 1048–1064, 2013.

[58] V. Rathi, R. Urbanke, M. Andersson, and M. Skoglund, "Rate-equivocation optimal spatially coupled ldpc codes for the bec wiretap channel," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pp. 2393–2397, IEEE, 2011.

[59] C. W. Wong, T. F. Wong, and J. M. Shea, "Secret-sharing ldpc codes for the bpsk-constrained gaussian wiretap channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 551–564, 2011.

[60] A. Subramanian, A. Thangaraj, M. Bloch, and S. W. McLaughlin, "Strong secrecy on the binary erasure wiretap channel using large-girth ldpc codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 585–594, 2011.

[61] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[62] E. Sasoglu and A. Vardy, "A new polar coding scheme for strong security on wiretap channels," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pp. 1117–1121, IEEE, 2013.

[63] M. Hayashi and R. Matsumoto, "Construction of wiretap codes from ordinary channel codes," *arXiv preprint arXiv:1001.1197*, 2010.

[64] M. Cheraghchi, F. Didier, and A. Shokrollahi, "Invertible extractors and wiretap protocols," *Information Theory, IEEE Transactions on*, vol. 58, no. 2, pp. 1254–1274, 2012.

[65] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Advances in Cryptology–CRYPTO 2012*, pp. 294–311, Springer, 2012.

[66] H. Tyagi and A. Vardy, "Explicit capacity-achieving coding scheme for the gaussian wiretap channel," in *2014 IEEE International Symposium on Information Theory*, pp. 956–960, IEEE, 2014.

[67] F. Oggier and B. Hassibi, "The secrecy capacity of the mimo wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4961–4972, 2011.

[68] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas i: The misome wiretap channel," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3088–3104, 2010.

[69] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas ii: The mimome wiretap channel," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5515–5532, 2010.

[70] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, 2009.

[71] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the gaussian mimo wiretap channel," in *2007 IEEE International Symposium on Information Theory*, pp. 2471–2475, IEEE, 2007.

[72] E. Ekrem and S. Ulukus, "The secrecy capacity region of the gaussian mimo multi-receiver wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2083–2114, 2011.

[73] E. Ekrem and S. Ulukus, "Capacity-equivocation region of the gaussian mimo wiretap channel," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5699–5710, 2012.

[74] J. Li and A. P. Petropulu, "On ergodic secrecy rate for gaussian miso wiretap channels," *IEEE Transactions on Wireless communications*, vol. 10, no. 4, pp. 1176–1187, 2011.

[75] L. Lai, H. El Gamal, and H. V. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5059–5067, 2008.

[76] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y.-H. Kim, "Wiretap channel with secure rate-limited feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5353–5361, 2009.

[77] Z. Rezki, A. Khisti, and M.-S. Alouini, "Ergodic secret message capacity of the wiretap channel with finite-rate feedback," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3364–3379, 2014.

[78] Y. Chen *et al.*, "Wiretap channel with side information," in *2006 IEEE International*

Symposium on Information Theory, pp. 2607–2611, IEEE, 2006.

[79] C. Mitrpant, A. H. Vinck, and Y. Luo, "An achievable region for the gaussian wiretap channel with side information," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2181–2190, 2006.

[80] E. Ekrem and S. Ulukus, "On the secrecy of multiple access wiretap channel," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, pp. 1014–1021, IEEE, 2008.

[81] W. Yang, R. F. Schaefer, and H. V. Poor, "Finite-blocklength bounds for wiretap channels," in *Information Theory (ISIT), 2016 IEEE International Symposium on*, pp. 3087–3091, IEEE, 2016.

[82] W. K. Harrison, D. Sarmento, J. P. Vilela, and M. Gomes, "Analysis of short blocklength codes for secrecy," *arXiv preprint arXiv:1509.07092*, 2015.

[83] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *Information Theory, IEEE Transactions on*, vol. 56, no. 5, pp. 2307–2359, 2010.

[84] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3401–3416, 2005.

[85] U. Erez and R. Zamir, "Achieving 1/2 log (1+ snr) on the awgn channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, 2004.

APPENDIX

E Proof of Equation (42)(43)

First it gives

$$f\left(x\right)=e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}=e^{-\frac{(x-\hat{\mu})^2}{2\hat{\sigma}^2}}e^{-\frac{(\mu_1-\mu_2)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}}, \tag{217}$$

where

$$\hat{\sigma}=\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}, \quad \hat{\mu}=\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mu_1+\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\mu_2. \tag{218}$$

Apply Equation(36) in,

$$\mu_1=\left(y-\tau_{y_{smt}\backslash b_{smnk}}\right)e^{-j2\pi(t-1)(k-1)/K}, \mu_2=0, \tag{219}$$

$$\sigma_1^2=\nu_{y_{smt}\backslash b_{smnk}}+\sigma_w^2, \sigma_2^2=\sigma_b^2. \tag{220}$$

Then it can derive the nominator of Equation(38) as

$$\int\limits_{L(y_{smt})}^{U(y_{smt})}\int\limits_{x} xf\left(x\right)dxdy=\int\limits_{L(y_{smt})}^{U(y_{smt})}\sqrt{2\pi\hat{\sigma}^2}\hat{\mu}e^{-\frac{(\mu_1-\mu_2)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}}dy \tag{221}$$

$$=\sqrt{2\pi\hat{\sigma}^2}e^{-j2\pi(t-1)(k-1)/K}\sigma_2^2\int\limits_{L(y_{smt})}^{U(y_{smt})}\frac{\left(y-\tau_{y_{smt}\backslash b_{smnk}}\right)}{\sigma_1^2+\sigma_2^2}e^{-\frac{\left(y-\tau_{y_{smt}\backslash b_{smnk}}\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}}dy \tag{222}$$

$$=\sqrt{2\pi\hat{\sigma}^2}e^{-j2\pi(t-1)(k-1)/K}\sigma_2^2\Gamma \tag{223}$$

$$\Gamma = \int\limits_{L(y_{smt})}^{U(y_{smt})} \frac{\left(y - \tau_{y_{smt} \setminus b_{smnk}}\right)}{\sigma_1^2 + \sigma_2^2} e^{-\frac{\left(y-\tau_{y_{smt} \setminus b_{smnk}}\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} dy \tag{224}$$

$$= \int\limits_{L(\mathcal{R}(y_{smt}))}^{U(\mathcal{R}(y_{smt}))} \int\limits_{L(\mathcal{I}(y_{smt}))}^{U(\mathcal{I}(y_{smt}))} \frac{\left(\mathcal{R}\left(y\right) - \mathcal{R}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)}{\sigma_1^2 + \sigma_2^2} \tag{225}$$

$$\times \exp\left(-\frac{\left(\mathcal{R}\left(y\right) - \mathcal{R}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2 + \left(\mathcal{I}\left(y\right) - \mathcal{I}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2 + \sigma_2^2\right)}\right) d\mathcal{R}\left(y\right) d\mathcal{I}\left(y\right) \tag{226}$$

$$+ i \int\limits_{L(\mathcal{R}(y_{smt}))}^{U(\mathcal{R}(y_{smt}))} \int\limits_{L(\mathcal{I}(y_{smt}))}^{U(\mathcal{I}(y_{smt}))} \frac{\left(\mathcal{I}\left(y\right) - \mathcal{I}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)}{\sigma_1^2 + \sigma_2^2} \tag{227}$$

$$e^{-\frac{\left(R(y)-\mathcal{R}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2+\left(\mathcal{I}(y)-\mathcal{I}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} d\mathcal{R}\left(y\right) d\mathcal{I}\left(y\right) \tag{228}$$

$$= e^{-\frac{\left(\mathcal{R}(y)-\mathcal{R}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} \Bigg|_{\mathcal{R}(y)=U(\mathcal{R}(y_{smt}))}^{\mathcal{R}(y)=L(\mathcal{R}(y_{smt}))} \times \int\limits_{L(\mathcal{I}(y_{smt}))}^{U(\mathcal{I}(y_{smt}))} e^{-\frac{\left(\mathcal{I}(y)-\mathcal{I}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} d\mathcal{I}\left(y\right) \tag{229}$$

$$+ ie^{-\frac{\left(\mathcal{I}(y)-\mathcal{I}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} \Bigg|_{\mathcal{I}(y)=U(\mathcal{I}(y_{smt}))}^{\mathcal{I}(y)=L(\mathcal{I}(y_{smt}))} \times \int\limits_{L(\mathcal{R}(y_{smt}))}^{U(\mathcal{R}(y_{smt}))} e^{-\frac{\left(\mathcal{R}(y)-\mathcal{R}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} d\mathcal{R}\left(y\right) \tag{230}$$

$$= \sqrt{2\pi\left(\sigma_1^2 + \sigma_2^2\right)} e^{-\frac{\left(\mathcal{R}(y)-\mathcal{R}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} \Bigg|_{\mathcal{R}(y)=U(\mathcal{R}(y_{smt}))}^{\mathcal{R}(y)=L(\mathcal{R}(y_{smt}))} \tag{231}$$

$$\times\ \mathcal{Q}\left(\frac{\mathcal{I}\left(y\right) - \mathcal{I}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\Bigg|_{\mathcal{I}(y)=U(\mathcal{I}(y_{smt}))}^{\mathcal{I}(y)=L(\mathcal{I}(y_{smt}))} \tag{232}$$

$$+ i\sqrt{2\pi\left(\sigma_1^2 + \sigma_2^2\right)} e^{-\frac{\left(\mathcal{I}(y)-\mathcal{I}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)\right)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}} \Bigg|_{\mathcal{I}(y)=U(\mathcal{I}(y_{smt}))}^{\mathcal{I}(y)=L(\mathcal{I}(y_{smt}))} \tag{233}$$

$$\times\ \mathcal{Q}\left(\frac{\mathcal{R}\left(y\right) - \mathcal{R}\left(\tau_{y_{smt} \setminus b_{smnk}}\right)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\Bigg|_{\mathcal{R}(y)=U(\mathcal{R}(y_{smt}))}^{\mathcal{R}(y)=L(\mathcal{R}(y_{smt}))}. \tag{234}$$

Note that

$$f^{prior}\left(y_{smt}\right)=\mathcal{Q}\left(\frac{\mathcal{I}\left(y\right)-\mathcal{I}\left(\tau_{y_{smt}\setminus b_{smnk}}\right)}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)\Bigg|_{\mathcal{I}(y)=U(\mathcal{I}(y_{smt}))}^{\mathcal{I}(y)=L(\mathcal{I}(y_{smt}))} \tag{235}$$

$$\times\,\mathcal{Q}\left(\frac{\mathcal{R}\left(y\right)-\mathcal{R}\left(\tau_{y_{smt}\setminus b_{smnk}}\right)}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)\Bigg|_{\mathcal{R}(y)=U(\mathcal{R}(y_{smt}))}^{\mathcal{R}(y)=L(\mathcal{R}(y_{smt}))}, \tag{236}$$

It can obtain Equation(42)(43).

F Proof of Lemma 3

Proof. Following the definition of Q_n,

$$Q_n\left(\rho^2\right)=(2\pi)^{-\frac{n}{2}}\int_{\mathbf{w}\in\Theta(\rho^2)}\exp\left(-\frac{\|\mathbf{w}\|^2}{2}\right)d\mathbf{w}, \tag{237}$$

and $\Theta\left(\rho^2\right)=\left\{\mathbf{w}\,\middle|\,\|\mathbf{w}\|^2\geq n\rho^2\right\}$. Apparently, $\Theta\left(\rho^2\right)$ is a sphere with radius $\sqrt{n}\rho$. Thus, the integral along $\mathbf{w}$ can be transformed into integral along radius as

$$Q_n\left(\rho^2\right)=(2\pi)^{-\frac{n}{2}}\int_{\sqrt{n}\rho}^{\infty}\frac{n\pi^{\frac{n}{2}}}{(n/2)!}r^{n-1}e^{-\frac{r^2}{2}}dr. \tag{238}$$

Using the Stirling's approximation, $(n/2)!\approx\sqrt{\pi n}(n/2e)^{\frac{n}{2}}$,

$$\begin{aligned}Q_n\left(\rho^2\right)&=(2\pi)^{-\frac{n}{2}}\int_{\sqrt{n}\rho}^{\infty}\frac{n\pi^{\frac{n}{2}}}{\sqrt{\pi n}(n/2e)^{\frac{n}{2}}}r^{n-1}e^{-\frac{r^2}{2}}dr\\&=\left(\frac{n}{\pi}\right)^{\frac{1}{2}}\left(\frac{e}{n}\right)^{\frac{n}{2}}\int_{\sqrt{n}\rho}^{\infty}r^{n-1}e^{-\frac{r^2}{2}}dr.\end{aligned} \tag{239}$$

□

G Proof of Lemma 4

Proof. Following the definition of Q_n, $\Theta\left(\rho^2\right)=\left\{\mathbf{w}\,\middle|\,\|\mathbf{w}\|^2\geq n\rho^2\right\}$. According to the law of large number, for the random variable $\mathbf{w}$ with variance n,

$$\lim_{n\to\infty}P\left\{\left|\frac{\|\mathbf{w}\|^2}{n}-1\right|>\tau\right\}=0. \tag{240}$$

Thus,

$$\lim_{n\to\infty} P\left\{\mathbf{w}\in\Theta\left(\rho^2\right)\right\} = 1 \text{ if } \rho < 1; \tag{241}$$

$$\lim_{n\to\infty} P\left\{\mathbf{w}\in\Theta\left(\rho^2\right)\right\} = 0 \text{ if } \rho > 1. \tag{242}$$

Because Q_n is the integral along $\forall\mathbf{w}\in\Theta\left(\rho^2\right)$, (241)(242) give the result

$$\lim_{n\to\infty} Q_n\left(\rho^2\right) = 1 \text{ if } \rho < 1; \tag{243}$$

$$\lim_{n\to\infty} Q_n\left(\rho^2\right) = 0 \text{ if } \rho > 1. \tag{244}$$

□

H Proof of Lemma 5

Proof. Denote the n-dimension sphere as $\mathcal{S}_n$. For the random variable $\mathbf{X}$ uniformly distributed in this sphere, the power of $\mathbf{X}$ is

$$S_{\mathbf{X}} = \frac{1}{V\left(\mathcal{S}_n\right)}\int_{\mathbf{x}\in\mathcal{S}_n}\|\mathbf{x}\|^2 d\mathbf{x}. \tag{245}$$

When n is even, the volume of an n-dimension sphere is $\frac{\left(\pi\rho^2\right)^{\frac{n}{2}}}{(n/2)!}$, and the surface is $\frac{n\pi^{\frac{n}{2}}}{(n/2)!}\rho^{n-1}$. (245) is obtained as

$$\begin{aligned} S_{\mathbf{X}} &= \frac{(n/2)!}{\left(\pi\rho^2\right)^{\frac{n}{2}}}\int_0^\rho \frac{n\pi^{\frac{n}{2}}}{(n/2)!} r^{n-1} r^2 dr \\ &= \rho^{-n}\int_0^\rho n r^{n+1} dr = \frac{n}{n+2}\rho^2. \end{aligned} \tag{246}$$

For the given power per dimension S_X,

$$S_{\mathbf{X}} = nS_X = \frac{n}{n+2}\rho^2. \tag{247}$$

Therefore, $\rho^2 = (n+2)\,S_X$. □

I Proof of Lemma 6

Proof. From (154) it gives

$$f_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) = \sum_{\lambda\in\Lambda} f_{\mathbf{W}'}\left(\mathbf{y}+\lambda-\mathbf{x}\right).$$

Take $\tilde{\mathbf{w}} = (\mathbf{y}-\mathbf{x}) \bmod \Lambda_1$, it gives

$$H\left(\mathbf{Y}|\mathbf{X}\right) \tag{248}$$

$$= -\int_{\mathbf{x}\in V_1}\int_{\mathbf{y}\in V_1} f_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) f_{\mathbf{X}}\left(\mathbf{x}\right)\log f_{\mathbf{Y}|\mathbf{X}}\left(\mathbf{y}|\mathbf{x}\right) d\mathbf{y} d\mathbf{x} \tag{249}$$

$$= -\int_{\mathbf{x}\in V_1}\int_{\tilde{\mathbf{w}}\in V_1} f_{\mathbf{Y}|\mathbf{X}}\left(\tilde{\mathbf{w}}\right) f_{\mathbf{X}}\left(\mathbf{x}\right)\log f_{\mathbf{Y}|\mathbf{X}}\left(\tilde{\mathbf{w}}\right) d\tilde{\mathbf{w}} d\mathbf{x} \tag{250}$$

$$= -\int_{\tilde{\mathbf{w}}\in V_1} f_{\mathbf{Y}|\mathbf{X}}\left(\tilde{\mathbf{w}}\right)\log f_{\mathbf{Y}|\mathbf{X}}\left(\tilde{\mathbf{w}}\right) d\mathbf{w}'. \tag{251}$$

which is independent of $\mathbf{Y}$. Therefore, to maximize $I\left(\mathbf{X};\mathbf{Y}\right)$ is only to maximize $H\left(\mathbf{Y}\right)$. Therefore, the optimal PDF for Y is

$$f_{\mathbf{Y}}\left(\mathbf{y}\right) = \frac{1}{V_1}, \mathbf{y} \in \mathcal{V}_{\mathbf{1}}. \tag{252}$$

According to Lemma 1, if the input $\mathbf{X}$ is uniformly distributed in $\mathcal{V}_1$, the output $\mathbf{Y}$ will also be uniformly distributed in $\mathcal{V}_1$ for the Mod-Λ channel. Therefore, the optimal PDF of $\mathbf{X}$ to maximize $I\left(\mathbf{X};\mathbf{Y}\right)$ is

$$f_{\mathbf{X}}\left(\mathbf{x}\right) = \frac{1}{V_1}, \mathbf{x} \in \mathcal{V}_{\mathbf{1}}. \tag{253}$$

□

J Proof of Lemma 7

Proof. The mean square error between the received signal $\mathbf{y}$ and the original signal $\mathbf{x}$ is $\mathbb{E}\left[\|\alpha\mathbf{y}-\mathbf{x}\|^2\right]$ and the optimal coefficient is

$$\alpha^* = \arg\min_{\alpha} \mathbb{E}\left[\|\alpha\mathbf{y}-\mathbf{x}\|^2\right]. \tag{254}$$

The first derivation is

$$\begin{aligned} \frac{d\mathbb{E}\left[\|\alpha\mathbf{y}-\mathbf{x}\|^2\right]}{d\alpha} &= \alpha\mathbb{E}\left[\|\mathbf{y}\|^2\right] - \mathbb{E}\left[\mathbf{y}\mathbf{x}^T\right] \\ &= \alpha\mathbb{E}\left[\|\mathbf{y}\|^2\right] - \mathbb{E}\left[(\mathbf{x}+\mathbf{w})\,\mathbf{x}^T\right]. \end{aligned} \tag{255}$$

As $\mathbf{x}$ is independent of $\mathbf{w}$,

$$\frac{d\mathbb{E}\left[\|\alpha\mathbf{y}-\mathbf{x}\|^2\right]}{d\alpha} = \alpha\mathbb{E}\left[\|\mathbf{y}\|^2\right] - \mathbb{E}\left[\|\mathbf{x}\|^2\right]. \tag{256}$$

And the second derivative is

$$\frac{d^2\mathbb{E}\left[\|\alpha\mathbf{y}-\mathbf{x}\|^2\right]}{d\alpha^2} = \mathbb{E}\left[\|\mathbf{y}\|^2\right] > 0. \tag{257}$$

(256)(257) give

$$\alpha^* = \frac{\mathbb{E}\left[\|\mathbf{x}\|^2\right]}{\mathbb{E}\left[\|\mathbf{y}\|^2\right]} = \frac{S_X}{S_X+\sigma_w^2}. \tag{258}$$

□

K Proof of Lemma 8

Proof. Define the function $f(x) = x-1-\ln x$, then $f'(x) = 1-x^{-1}$ and $f''(x) = x^{-2} > 0$. It gives $\min f(x) = f(1) = 0$. with the MMSE coefficient $\frac{S_X}{\sigma^2_{W_z{}'}} = 1+\frac{S_X}{\sigma^2_{W_z}} > 1$, it gives

$$\frac{S_X}{\sigma^2_{W_z{}'}} > 1 + \ln\frac{S_X}{\sigma^2_{W_z{}'}}. \tag{259}$$

□

CURRICULUM VITAE

NAME: Chen Cao

ADDRESS: W.S. Speed Hall
University of Louisville
Louisville, KY 40208

EDUCATION: Ph.D. ECE, University of Louisville, 2017

M.S. ECE, Xi'an Jiaotong University, 2013

B.S. ECE, Xi'an Jiaotong University, 2010

RESEARCH: Digital Signal Processing, Wireless Communication

PUBLICATION:

[1] **Chen Cao**, Hongxiang Li, " Turbo Receiver for Uplink MU-MIMO-OFDM Systems with Low-Resolution ADCs", submitted on *IEEE Transactions on Wireless Communication*.

[2] **Chen Cao**, Hongxiang Li, "A Practical Analysis of the Finite-Blocklength Secrecy Performance of Lattice Codes," submitted on *IEEE Transactions on Information Forensics and Security*.

[3] **Chen Cao**, Hongxiang Li, Zixia Hu, Huacheng Zeng, "1-Bit Transceiver Cluster for Relay Transmission," *IEEE Communications Letters*, 2016.

[4] **Chen Cao**, Hongxiang Li, Zixia Hu, "A Novel Cross-Level Decoding Method for LT Codes," *IEEE Communications Letters*, 2014.

[5] **Chen Cao**, Hongxiang Li, Zixia Hu, "Physical-layer secrecy performance in finite blocklength case," *GLOBELCOM 2015*.

[6] **Chen Cao**, Hongxiang Li, Zixia Hu, “An AMP Based Decoder for Massive MU-MIMO-OFDM with Low-Resolution ADCs,” *ICNC 2017*.

[7] Huacheng Zeng, **Chen Cao**, Hongxiang Li, Qiben Yan “HyRx: Enabling Harmonious Coexistence of WiFi and LTE,” submitted on *MobiCom ’17* .