

# MAKING WIRELESS COMMUNICATIONS MORE EFFICIENT

by

Jing Guo

A Dissertation

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

Doctor of Philosophy



School of Electrical and Computer Engineering

West Lafayette, Indiana

August 2021

**THE PURDUE UNIVERSITY GRADUATE SCHOOL  
STATEMENT OF COMMITTEE APPROVAL**

**Dr. David J. Love, Chair**

School of Electrical Engineering

**Dr. James V. Krogmeier**

School of Electrical Engineering

**Dr. Michael D. Zoltowski**

School of Electrical Engineering

**Dr. Xiaojun Lin**

School of Electrical Engineering

**Approved by:**

Dr. Dimitrios Peroulis

For my family

## ACKNOWLEDGMENTS

I was changing my interested of field from image processing to wireless communication and physical layer algorithm development when I first started my PhD. I appreciate my advisor, Prof. David Love, for being my advisor during the long journey of my PhD. I learned a lot from him on many different levels. His guidance and mentorship were really invaluable. He provided me many great opportunities to work with many great research laboratories to strengthen my research.

I appreciate many other professors who helped me during my PhD. Prof. James Krogmeier, Prof. Michael Zoltowski, and Prof. Xiaojun Lin were the members in my PhD advising committee. Their suggestions and feedbacks helped me make my work better. I also appreciate Mr. Navid Yazdani and Dr. Raghu Raj for providing me opportunities to intern at MIT Lincoln lab and NRL, respectively.

I am grateful to have all my colleagues at school. The constructive discussions always enlighten me throughout my study. I thank Matthew Booth for helping me get through one of my difficult time. I thank Dr. Feng Zhang's helps and guidance started when I was a undergraduate. Special shoutout to all my friends, Alan Carreno, Tanner Hamann, Irving Garcia, Marcus Currie, Damien Nagle, and Bryan Ehlers.

I would like to thank my family. I appreciate my father who always believed in me, my mother who guides and inspires me, aunts whose love and support surrounded me, and grandparents' wise guidance. Words cannot describe my gratitude to Hui. I would have never been here today without you.

# TABLE OF CONTENTS

LIST OF FIGURES . . . . .	8
ABBREVIATIONS . . . . .	10
ABSTRACT . . . . .	11
1 INTRODUCTION . . . . .	12
1.1 Implementation of rate-adaptive integer forcing compression in distributed wireless relay networking . . . . .	12
1.2 Nonparametric decentralized detection and sparse sensor selection via online kernel scalar quantization . . . . .	14
1.3 A folding approach for multiple antenna arrays using low-resolution ADCs .	16
2 RATE ADAPTIVE INTEGER FORCING SOURCE CODING . . . . .	20
2.1 System Model . . . . .	20
2.2 Rate-Adaptive Integer Forcing Compression . . . . .	23
2.3 Simulation Results . . . . .	26
3 NONPARAMETRIC DECENTRALIZED DETECTION AND SPARSE SENSOR SELECTION VIA ONLINE KERNEL SCALAR QUANTIZATION . . . . .	30
3.1 System Model and Preliminaries . . . . .	30
3.1.1 Wireless Sensor Network System Model . . . . .	30
3.1.2 Kernel and Decision Function . . . . .	31
Kernel . . . . .	31
Decision function . . . . .	33

3.2	Estimation Model . . . . .	34
3.2.1	Prior Work: Batch Learning Problem Formulation . . . . .	34
3.2.2	Online Learning Problem Formulation . . . . .	36
3.3	Online Learning Methodology . . . . .	37
3.3.1	Obtaining the weight of decision function $\mathbf{f}_n$ . . . . .	38
3.3.2	Obtaining $P_{n+1}(\mathbf{q}_n \mathbf{x}_n)$ for all possible $\mathbf{q}_n$ and $\mathbf{x}_n$ . . . . .	39
3.3.3	Obtaining weighted parameter vector $\mathbf{w}_{n+1}$ . . . . .	42
3.4	Performance Analysis . . . . .	44
3.4.1	Assumptions . . . . .	45
3.4.2	Connection Between OL and BL Approach . . . . .	46
3.4.3	Quantization Rules and Weight Parameter Vector Analysis . . . . .	48
	Quantization Rules Analysis . . . . .	48
	Weighted Parameter Vector Analysis . . . . .	50
3.4.4	Convergence Analysis . . . . .	51
3.5	Numerical Performance Analysis . . . . .	53
3.5.1	Randomly Generated Data for Binary Classification Problem . . . . .	54
3.5.2	UCI Iris Data for Multi-Class Problem . . . . .	57
4	A FOLDING APPROACH FOR MULTIPLE ANTENNA ARRAYS USING LOW- RESOLUTION ADCS . . . . .	63
4.1	System Model and Low-Resolution ADC Array Architecture . . . . .	63

4.1.1	System Model . . . . .	63
4.1.2	Low-Resolution ADC Array Construction by Deconstructing a High-Resolution ADC . . . . .	64
4.2	Folded Low-Resolution ADC Array Design and Implementation for Multiple Antenna Communication . . . . .	67
4.2.1	Basics of Folding . . . . .	67
4.2.2	Design of Folded Low-Resolution ADC Array . . . . .	69
4.2.3	Statistical Analysis of Folded Low-Resolution ADC Array in Real Space . . . . .	71
4.3	Folded Low-Resolution ADC Array Enhancements . . . . .	73
4.3.1	Folded Low-Resolution ADC Array using Gray Labeling . . . . .	73
4.3.2	Sign LMB Folded Low-Resolution ADC Array . . . . .	75
4.3.3	Antenna Ordering Strategies . . . . .	76
	Antenna ordering in the low SNR regime . . . . .	76
	Antenna ordering in the moderate-to-high SNR regime . . . . .	76
4.4	System Throughput Analysis . . . . .	77
4.4.1	Mutual Information Preliminary . . . . .	77
4.4.2	System Throughput Analysis in High SNR Regime . . . . .	78
4.5	Numerical Analysis . . . . .	82
5	SUMMARY . . . . .	90
	REFERENCES . . . . .	92

## LIST OF FIGURES

2.1	End-to-end system model . . . . .	20
2.2	Ergodic capacity performance for different IFSC-based compression strategies with four users and Rayleigh fading uplink channels and four relay nodes. . .	27
2.3	Ergodic capacity performance of RAIF compression and approximate Gaussian vector quantization for four users with Rayleigh fading uplink channels.	28
2.4	Ergodic capacity performance of RAIF compression and approximate Gaussian vector quantization for a relay network of ten relay nodes - the number of users is four with a strong uplink channel between the intended user and the first relay node. . . . .	29
3.1	Illusion of the generic decentralized detection system model . . . . .	33
3.2	Construction of $\mathcal{X}_m$ based on (3.71) with $c_1 = 0.1$ and $c_1 = 1$ . . . . .	57
3.3	Demonstrating examples that sensor $m$ observes data/time drawn from the non-separable ( $c = 0.1$ ) and separable dataset ( $c = 1$ ). . . . .	57
3.4	The ACR performance for Algorithm 2 between MSOKSQ without updating $\mathbf{w}_n$ and MSOKSQ without updating $\mathbf{w}_n$ and $P_n(\cdot \cdot)$ assuming a 3-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 1-bit deterministic quantizer given $M = 10$ .	58
3.5	The ACR performance for Algorithm 2 between MSOKSQ with optimal/random sensor selection when $M = 1, 5$ , and 10 sensors are enabled among $M = 11$ and 3-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 2-bit deterministic quantizer given $M = 11$ . . . . .	58
3.6	Demonstrating the value of $P_{m,n_{i_x}}(q_{m,n_{i_x}} = 0.5 x_{m,n_{i_x}} = 0.4329)$ for $i_x = 1, 2, \dots, 30$ , where $i_x$ is the $i_x$ th times that sensor $m$ observes 0.4329 when $\partial_{P_{m,n_{i_x}}(q_{m,n_{i_x}} x_{m,n_{i_x}}=0.4329)}R(s_{n_{i_x}}) \neq 0$ . . . . .	59
3.7	The ACR performance for Algorithm 2 between MSOKSQ with optimal and random sensor selection. $M = 1, 5$ , and 10 sensors are assumed to be enabled and a 2-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 1-bit deterministic quantizer given $M = 11$ . . . . .	61
3.8	The ACR performance for Algorithm 2 between MSOKSQ without updating $\mathbf{w}_n$ and MSOKSQ without updating $\mathbf{w}_n$ and $P_n(\cdot \cdot)$ . $M = 1, 5$ , and 10 sensors are assumed to be enabled and a 2-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 2-bit deterministic quantizer when $M = 11$ . . . . .	61



3.9	The ACR performance for Algorithm 2 between MSOKSQ with optimal and random sensor selection. $M = 1, 5$ , and 10 sensors are assumed to be enabled and a 2-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 3-bit deterministic quantizer when $M = 11$ . . . . .	62
4.1	Implementation of low-resolution ADC array at the receiver. . . . .	64
4.2	Showing the red regions must be integrated over to compute the received bit probabilities, e.g., $\Pr(b_1 = 1 s, h_1)$ . . . . .	72
4.3	Achievable rates for sign LMB folded low-resolution ADC arrays using both labelings and sign-based low-resolution ADC array with single-path UPA structure mmWave channel when $M = 2, 3$ and $N = 16$ . . . . .	85
4.4	Achievable rates for sign LMB folded low-resolution ADC array using both labelings and sign-based low-resolution ADC array with single-path UPA structure mmWave channel when $M = 1, 2, 3, 4$ , $N = 16$ . . . . .	86
4.5	Ratio of achievable rate over approximate achievable rate with single-path UPA mmWave channel when $M = 2, 3$ and $N = 16$ . . . . .	86
4.6	Achievable rates for designs of sign LMB folded low-resolution ADC arrays v.s. folded low-resolution ADC arrays when $M = 2$ and $N = 16$ . (a) and (b) use single-path UPA mmWave channel and Rayleigh fading channel, respectively. . . . .	87
4.7	Achievable rates for LMB folded low-resolution ADC arrays with antenna ordering for moderate-to-high SNR, antenna ordering for low SNR, and no antenna ordering with Rayleigh fading channel when $M = 3$ and $N = 16$ . . . . .	88
4.8	UPA structure mmWave fading channel model achievable rates of sign LMB folded low-resolution ADC arrays using Gray labeling comparing $\mathbf{f}_u$ v.s. $\mathbf{f}_{svd}$ when $M = 1, 2, 3$ and $N = 16$ . . . . .	88
4.9	UPA mmWave channel model achievable rates of sign LMB folded low-resolution ADC arrays using Gray labeling for $N = 16$ and $N = 64$ given $M = 1, 2$ , and 3 and beamforming vector $\mathbf{f}_{svd}$ . . . . .	89

## ABBREVIATIONS

ADC	Analog to digital converter
AWGN	additive white Gaussian noise
BL	Batch learning
CRAN	Cloud radio access network
CSI	Channel state information
CSIR	Channel state information at receiver
FADC	Folded ADC
IFSC	Integer forcing source coding
i.i.d	Independent and identically distributed
LMB	Left-most bit
MISO	Multiple-input single-output
MIMO	Multiple-input multiple-output
mmWave	Millimeter Wave
MSE	Mean square error
MSOKSQ	Multi-sensor online kernel scalar quantization
MRT	Maximum ratio transmission
NORMA	Naive online risk minimization algorithm
OL	Online learning
PKS	Phase shift keying
QAM	Quadrature amplitude modulation
RAIF	Rate adaptive integer forcing
SIMO	Single-input multiple-output
SNR	Signal to noise ratio
UPA	Uniform planar array

## ABSTRACT

Given the increasing importance of mobile data access, extending broadband wireless access have become a global grand challenge. Wireless sensor networks (WSNs) and millimeter wave (mmWave) systems have been introduced to resolve these issues which motivate us to have further investigation. In this paper, the first two work assuming a quantized-and-forward WSN. We first develop a rate adaptive integer forcing source coding (RAIF) scheme to enhance the system throughput by assigning optimal quantization rate to each sensor optimally. Then, we are interested in developing an supervised online technique for solving classification problems. In order to enhance the classification performance, we developed this technique by jointly training the decision function that determines/estimates class label, quantizers across all sensors, and reliability of sensors such that  $M$  most reliable sensors are enabled. Finally, we develop an idea to provide a folded low-resolution ADC array architecture that can utilize any of the widely published centralized folded ADC (FADC) implementation by placing the centralized FADC branches at different antenna elements in a millimeter wave (mmWave) system. With adding a simple analog shift and modulo operations prior to the sign quantizer, we show that the multiple low-resolution ADCs across the array elements can be properly designed such that they can be combined into an effective high-resolution ADC with excellent performance characteristics.

# 1. INTRODUCTION

More than half of the world's population do not have access to the internet due to a lack of communication infrastructure [1]. Two most popular solutions to resolve this probability is wireless sensor network and distributed mmWave MIMO wireless communication network. Although these two areas are extensively studied for the past few decades, we introduce two quantization schemes, which are Rate adaptive integer forcing source coding and folded analog to digital converter, are compatibility with these two networks to further enhance the system throughput.

## 1.1 Implementation of rate-adaptive integer forcing compression in distributed wireless relay networking

The relay networking provide the reliable transmission and broad coverage. It provides [2] make them good candidates to extend coverage to underdeveloped areas or to regions that suffered from natural disasters. One example is the deployment of Google balloons in Peru and Puerto Rico to provide wireless coverage when infrastructure was destroyed because of natural disasters.

For the past two decades, there has been much work focusing on different forms of relay networking. This work can be mostly divided into four categories; amplify-and-forward [2]–[7], decode-and-forward [8]–[12], compress-and-forward [13]–[17], and compute-and-forward relaying [18]–[23]. Amplify-and-forward relaying is the simplest among the four because the relay network linearly processes its received signals and retransmits the signals to the receivers. On the contrary, decode-and-forward relaying decodes or partially decodes the received signal before forwarding the decoded symbols to the receiver while compress-and-forward relaying quantizes the observations at the relay network and forwards the quantized observations to the receiver side to decode. Recently, compute-and-forward has been studied to harness the interference by decoding linear equations of the transmitted messages using the noisy linear combinations at the output of the uplink channel. The relay network then forwards many of these linear combinations to the receiver to decode the desired message.

We consider a compress-and-forward relaying strategy utilizing a cooperative distributed relay network consisting of many relay nodes. The relay nodes cooperate together to transmit the data of an intended transmit user to a receiving user. The relay nodes perform a quantization strategy based on integer forcing that is proposed in [24]. Integer forcing source coding (IFSC) is a recently proposed quantization scheme that is based on lattice vector quantization [25]. It is proposed as a distributed lossy compression strategy that requires symmetric quantization rates and equal desired distortions. IFSC encoders are implemented inside the relay network to perform the dithered quantization modulo operation on observed signals. The processed signal is then forwarded from the relay network to the receiving user. Decoders at the receiving user recover the integer linear combinations of the quantized signals subject to desired distortion levels.

An extension of IFSC has been studied in [26] which we denote by modified IFSC in this thesis. The modified IFSC is proposed to be applied in cloud radio access networks (CRAN) [26]. In contrast to IFSC, different coarse and fine lattices are considered to achieve different distortion levels while maintaining symmetric quantization rates. Although the work in [26] enhances the throughput of the system, it increases the computational complexity because of the assumption of different coarse and fine lattices.

Different from [24] and [26], we propose a novel IFSC-based quantization strategy that can support different quantization rates with lower computational complexity. We denote the proposed strategy by rate-adaptive integer forcing (RAIF) compression. We deploy RAIF compression at the relay network in our cooperative distributed relaying system. Each node quantizes its observation using a predetermined quantization rate. The quantization rates are optimally allocated to minimize the difference between the mean square error (MSE) when the relay nodes does not perform any quantization and the MSE when the relay nodes quantize their observations. The proposed rate-allocation algorithm is inspired by the work in [27]. The relay network then broadcasts its signals to all users. When the receiving user is decoding the transmission of a certain user, the transmissions of the rest of the users are treated as interference. In contrast to IFSC in [24] and modified IFSC in [26], our compression strategy uses asymmetric quantization rates to achieve different desired distortion levels. In

addition, we consider a lattice design that uses the same coarse lattice and different fine lattices which reduces the computational complexity as explained in following chapter

## 1.2 Nonparametric decentralized detection and sparse sensor selection via on-line kernel scalar quantization

Wireless sensor networks (WSNs), which leverage intermediate sensors to sample and process the data from an environment, have been extensively studied in the past few decades [1], [28]–[42]. A variety of WSN applications in, e.g., wireless communications [1], [28]–[36] and radar [37] have already or will soon impact our day-to-day lives significantly. This widespread use of WSNs is driving research in a number of communication and signal processing directions. Of particular continuing interest is leveraging WSNs to provide detector/estimator of one or more underlying phenomena, e.g., for signal classification tasks [38]–[42].

In this paper, we focus on a decentralized detection problem [38]–[48] exploiting a compress-and-forward sensor network [28]–[36], [38]. In these settings, each sensor forwards the processed information based on the local decision rule to a fusion center for further processing and decision-making. Decentralized detection problems can be divided into parametric [44]–[46] and non-parametric categories [38]–[42], [47], [48]. A parametric decentralized detection problem assumes that a fusion center knows or partially knows the joint distribution of  $\mathbf{X}$  (random vector) and  $Y$  (random variable), denoted as  $P(\mathbf{X}, Y)$ , where  $\mathbf{X}$  is the observation vector made by the sensors and  $Y$  is the class label of  $\mathbf{X}$ . However, this assumption could lead to performance limitations and an increase in system complexity [39]. Thus, non-parametric decentralized detection can be considered.

We are interested specifically in non-parametric decentralized detection for signal classification tasks. In the past, authors have studied this topic by leveraging kernel-based learning techniques, e.g., [38]–[42]. Kernel methods are commonly employed in learning problems to deal with non-linearly separable statistical models [38]–[42]. The scenarios studied in [38]–[40] most closely align with our focus in this paper. In these works, each sensor forwards the quantized observation based on a quantization rule (refers to the local decision) to a fusion center. Then, the fusion center employs a decision function chosen from a high-dimensional function space that takes the mapped quantized observations by the kernel to the set of

reals. The output of the decision function is used to infer the corresponding signal classification class label. To maximize classification performance, a joint optimization problem with respect to the quantizers (i.e., quantization rules across the sensors) and the decision function is considered [38]–[40].

The methods in [39], [40] are based on batch learning (BL), where model training is conducted after all data is collected. For contemporary WSN applications which require learning over large volumes of samples, BL approaches potentially consume high computational power and require a significant amount of memory. The detector/estimator learnt by the BL approach is highly dependent on the input training sample distribution, i.e.,  $P(\mathbf{X}, Y)$ . If the testing dataset distribution is mismatched to the current WSN inputs, which will happen when the environment evolves over time, the performance (e.g., classification) can be degraded significantly. To address these challenges, we consider an online learning (OL) approach to signal classification that processes the data over time dynamically. Note that the OL problem has a variety of applications [41], [42], [49]–[53], but, generally, the role of quantization has not been considered.

In this paper, our *primary goal* is to develop an OL methodology for WSNs that jointly finds quantizers and the decision function dynamically. Our theoretical analysis will also establish relationships between the OL and the BL cases studied in [39]. In doing so, we will see how the proposed algorithm affects the quantization rules based on some computation friendly marginalized kernels.

In a WSN, it can be expected that quantization outputs will furnish varying contributions to the signal classifier at the fusion center, e.g., some sensors are naturally less reliable when they are deployed over a large area. Hence, to promote the network power efficiency, we are interested in the sensor selection problem [40], [54], [55], where a certain number of lowest contributing sensors are disabled. We can accomplish this within our framework by introducing a marginalized weighted kernel [40]. Compared to our prior work [38], we consider multiple sensors, introduce new techniques for obtaining the quantization rules across the sensors and the decision function, and provide sensor selection strategy.

Notations for Chapter 2: Scalars and column vectors are written using lower case letters, e.g.,  $x$ , and a boldface lower case letters, e.g.,  $\mathbf{x}$ , respectively. Random variables and vec-

tors are written using upper case letter, e.g.,  $X$ , and a boldface upper case letter, e.g.,  $\mathbf{X}$ , respectively.  $\text{sgn}(\cdot)$  is the signum function.  $\|\cdot\|_1$  denotes the  $L_1$  norm.  $\|\cdot\|_0$  is the count of the number of non-zero elements.  $\|\cdot\|$  denotes the  $L_2$  norm.  $|\cdot|$  means cardinality of the set.  $O(\cdot)$  is big  $O$  notation.  $\mathbb{1}(\cdot, \cdot)$  is an indicator function.

### 1.3 A folding approach for multiple antenna arrays using low-resolution ADCs

Multiple antenna wireless systems are now considered necessary for any high-rate wireless standard, and the number of antennas continues to increase with each new broadband wireless standard release. It is expected that beyond-5G systems operating in sub-6 GHz and millimeter wave (mmWave) bands will utilize large-scale multiple antenna arrays, with much focus on massive multiple-input multiple-output (MIMO) and mmWave beamforming/pre-coding systems [56]–[66]. Massive MIMO techniques utilize large numbers of antennas at the base station to increase the number of users that can be supported across the time-frequency resources [62]. The mmWave bands offer wide bandwidths and enhance system throughput for future wireless deployments [56]. However, mmWave comes with propagation challenges, which necessitate the use of a large antenna array and multiple antenna signal processing at the transmitter and receiver.

Past MIMO work has used “small” numbers of antennas at the receiver and assumed a high-speed high-resolution (e.g., six bits or more) analog-to-digital converter (ADC) for each of the in-phase and quadrature streams at each element [67], [68]. When the number of antennas grows large, the economics and power requirements of the array hardware change, causing high-resolution ADCs to become costly in dollars and power. To overcome this, there has been much work on signal processing techniques that utilize a low-resolution ADC at each element [63], [64], [67]–[86], with the most popular choice being to use a sign quantizer for each of the in-phase and quadrature signals at each element [77]–[80]. At a high-level, the multitude of low-resolution quantization outputs coming from the antennas can be fused together by using the diversity advantages offered by the spatially-varying channels corresponding to the different array elements.



Because of the channel dependence, the foremost problem of the sign-based low-resolution ADC array is that the sign bits from the different antennas are not necessarily complementary. A simple example is to study the case when the channel realization between the transmitter and every receiver is one. In this scenario, all output bits from the array are noisy versions of the same bit, and the throughput is limited to at most one bit per channel use per dimension. For more general channels, the achievable rate performance of a system with sign quantization has an upper-bound that grows *logarithmically* with the number of array elements in the high signal-to-noise ratio (SNR) regime [79]. This is problematic because ideally the rate of a properly designed quantizer should scale *linearly* with the number of antennas. The other problem is amplitude uncertainty. Unless highly accurate statistical channel models are available [73], the sign quantizer is unable to distinguish between two transmit symbols that are scaled versions of each other (i.e., two symbols  $s_1$  and  $s_1 = \alpha s_2$  with  $\alpha \neq 1$ ). In most practical scenarios, this limits the transmit symbol to phase shift keying (PSK) constellations. In summary, the sign-based low-resolution ADC arrays generally offer far from optimal performance, especially when the channels are not sufficiently spatially diverse.

In this thesis, we show how to overcome these challenges by designing simple analog processing that varies across the array elements in a carefully chosen way. This diversity of analog processing acts to supplement the channels' spatial diversity to facilitate detection that can recover high-rate constellations. This per-element analog processing is inspired from techniques previously used in the high-resolution ADC literature, building upon past work in [83]. For example, we show how to design an  $M$ -antenna array that uses two single-bit per sample ADCs at each element, and we show to use signal processing theory developed for a single high-resolution ADC that produces  $2M$  bits of information for each sample.

There are a variety of techniques developed for use in high-resolution ADCs. In initial work, researchers, including us, focused on the structure of the  $\Delta\Sigma$  (or  $\Sigma\Delta$ ) ADC [63], [64], [83]–[86]. However,  $\Delta\Sigma$  ADCs require complex signal processing in the analog domain to perform noise shaping in order to overcome the quantization loss. In this thesis, we focus on other popular high-resolution ADCs called folded ADCs (FADCs) (or folding-related ADCs) e.g., [40], [67], [68], [87]–[98]. Note that folded ADCs are widely studied by researchers in

circuits and used commercially [67], [68]. The word *folded* comes from the analog modulo operation which can be implemented based on a simple low-power consumption analog signal processing circuits, e.g., [87].

In our framework, each of the in-phase and quadrature signals at each element utilizes simple shift and modulo processing followed by a sign quantizer. To select the shift and modulo processing across the multiple array elements, we make use of techniques from the multi-branch FADC. Our framework is not circuit specific and can be implemented with any of the widely published per-branch circuits employed for centralized high-sampling rate FADC implementations, e.g., [67], [68]. Compared to prior work on the low-resolution ADC array based on  $\Delta\Sigma$  ADC theory, the analog signal processing of the proposed design at each antenna in the analog domain is simpler. The proposed array also outperforms the sign-based low-resolution ADC array that use only a sign quantizer at each array element with our technique easily supporting multiple amplitude constellations such as quadrature amplitude modulation (QAM) even when the channel lacks spatial variation. Most importantly, we show that our proposed low-resolution array that uses folding at each element can provide an achievable rate that increases linearly with the number of elements in the high SNR regime, a strong improvement over the logarithmic scaling of the sign-based low-resolution ADC array. In addition, our folded low-resolution ADC array offers deep connections to other quantization and source coding schemes. In particular, the modulo quantization structure of the proposed design has connections to lattice codes using integer forcing source coding, but the constructions of the quantizers the assumptions used for the system performance analysis are completely different [24], [26], [99], [100]. The unique reconstruction of the high-rate constellations such as QAM also relates to labeling systems such as those used in the residue number system [101].

Notations for Chapter 3: Scalars and column vectors are written using lower case letters, and a boldface letter denotes a vector. All logarithms are assumed to be base two. The function  $\text{sgn}(\cdot)$  is the signum function. We denote  $\text{Re}(x)$  and  $\text{Im}(x)$  as the real and imaginary part of  $x$ , respectively.  $\log[\cdot]$  is denoted as log base two. The probability density function and the entropy are denoted by  $p(\cdot)$  and  $\mathcal{H}(\cdot)$ , respectively. The absolute value of  $x$  is defined as  $|x|$  and the modulo of  $a$  of size  $b$  is written as  $[a] \bmod b$ . The transpose and the conjugate

transpose are denoted by  $(\cdot)^T$  and  $(\cdot)^*$ , respectively. The Kronecker product is written as  $\otimes$ .  $\text{Card}(\mathcal{S})$  means cardinality of the set  $\mathcal{S}$ .

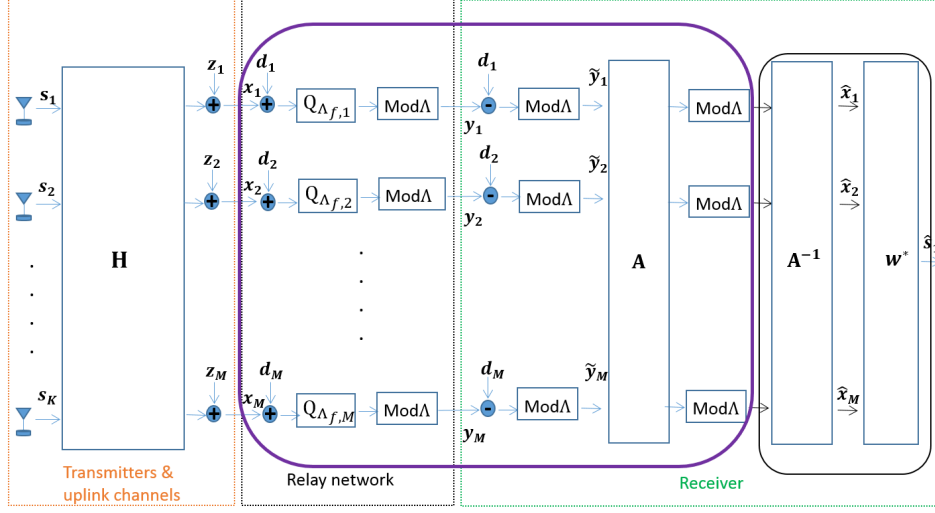


Figure 2.1. End-to-end system model

## 2. RATE ADAPTIVE INTEGER FORCING SOURCE CODING

### 2.1 System Model

We consider a multi-user distributed relay network where communication between any pair of  $K$  users is facilitated by a relay network that consists of  $M$  relay nodes as shown in Fig. 1. In particular, the receiver not only captures the signal that is formed by the intended transmit user, but also receives the transmitted signals from the interfering users (i.e., the other  $K - 1$  users). The relay nodes and the users are assumed to be equipped with a signal antenna.

In the uplink (transmission from the transmit user to the relay network), all of the users transmit their signals simultaneously through fading channels to the relay network. The observed signal at the  $m^{\text{th}}$  relay node at the  $\ell^{\text{th}}$  channel use can be written as

$$\begin{aligned}
 x_m[\ell] &= \sum_{k=1}^K h_{m,k}[\ell] s_k[\ell] + z_m[\ell] \\
 &= h_{m,1} s_1[\ell] + \sum_{k=2}^K h_{m,k} s_k[\ell] + z_m[\ell]
 \end{aligned} \tag{2.1}$$

where  $s_k[\ell] \in \mathbb{R}$  is the signal transmitted by the  $k^{\text{th}}$  user at the  $\ell^{\text{th}}$  channel use satisfying a power constraint that

$$\mathbb{E}[s_k[\ell]^2] = P_k, \quad (2.2)$$

$h_{m,k}[\ell] \in \mathbb{R}$  is the fading uplink channel between the  $k^{\text{th}}$  user and the  $m^{\text{th}}$  relay node which is distributed as  $\mathcal{N}(0, 1)$ , and  $z_m[\ell] \in \mathbb{R}$  is the additive white Gaussian noise (AWGN) at the  $m^{\text{th}}$  relay node which is spatially and temporally independent and identically distributed (i.i.d)  $\mathcal{N}(0, \sigma^2)$ . We assume that the uplink channels are static over the length of the quantization block of size  $L$  (i.e.,  $h_{m,k}[\ell] = h_{m,k}$  for  $\ell = 0, 1, \dots, L-1$ ). Without loss of generality, the first user (i.e.,  $k = 1$ ) is considered to be the intended user.

In this thesis, we assume that any lattice in high dimensional nested lattice chains is a discrete subgroup of  $\mathbb{R}^L$  which is denoted as  $\Lambda \subseteq \Lambda_{f,1} \subseteq \dots \subseteq \Lambda_{f,m}$ . All of the fine lattices  $\Lambda_{f,m}$  for  $m = 1, \dots, M$  are considered to be good for MSE quantization such that the relationship between the distortion levels and quantization rates can be found. The coarse lattice  $\Lambda$  is considered to be good for channel coding such that the overload error probability can be neglected with high probability [24]. This assumption is a relaxed version of “good” nested lattice chains, whose existence has been proposed and rigorously analyzed in [102].

The  $m^{\text{th}}$  relay node collects the observations through  $L$  channel uses constructing the combined row vector  $\mathbf{x}_m = [x[0], x[1], \dots, x_M[L-1]] \in \mathbb{R}^{1 \times L}$ . The  $m^{\text{th}}$  relay node then adds the dither vector  $\mathbf{d}_m$ . The reason we add the dither vector is to make the quantization error independent of the observed signal [24]. The vector  $\mathbf{d}_m$  is independent of  $\mathbf{x}_m$  and is uniformly distributed in  $\mathcal{V}_{f,m}$  which is the fundamental Voronoi region of the fine lattice at the  $m^{\text{th}}$  relay node. The dither vector is assumed to be known to both the encoder and the decoder. The sum (i.e.,  $\mathbf{x}_m + \mathbf{d}_m$ ) is quantized onto the fine lattice  $\Lambda_{f,m}$  using  $Q_{\Lambda_{f,m}}$  and then mapped onto the fundamental Voronoi region of the coarse lattice using  $\text{mod } \Lambda$ . The quantization function  $Q_{\Lambda_{f,m}}$  is defined to be [18], [24]

$$Q_{\Lambda_{f,m}}(\mathbf{x} + \mathbf{d}) = \underset{\bar{\mathbf{\lambda}} \in \Lambda_{f,m}}{\text{argmin}} \|\mathbf{x} + \mathbf{d} - \bar{\mathbf{\lambda}}\|. \quad (2.3)$$

The quantized signals can be written as

$$\mathbf{y}_m = \left[ Q_{\Lambda_{f,m}}(\mathbf{x}_m + \mathbf{d}_m) \right] \bmod \Lambda. \quad (2.4)$$

We assume the downlink channel (between the relay network and the receiver) is robust. This assumption follows from the bandwidth constraint enforced on the total number of bits that can be broadcast from the relay network. The  $m^{\text{th}}$  relay node forwards  $\mathbf{y}_m$  defined in (2.4) to the receiver. At the receiver, the decoder subtracts the dither vectors and reduces the result using  $\bmod \Lambda$ . The reconstructed vector at the receiving user is

$$\tilde{\mathbf{y}}_m = \left[ [Q_{\Lambda_{f,m}}(\mathbf{x}_m + \mathbf{d}_m)] \bmod \Lambda - \mathbf{d}_m \right] \bmod \Lambda. \quad (2.5)$$

Adding and subtracting  $\mathbf{x}_m$  in the right hand side of (2.5) gives

$$\begin{aligned} \tilde{\mathbf{y}}_m &= \left( \mathbf{x}_m - \underbrace{[(\mathbf{x}_m + \mathbf{d}_m) - Q_{\Lambda_{f,m}}(\mathbf{x}_m + \mathbf{d}_m)]}_{\text{Quantization Error}} \right) \bmod \Lambda \\ &= (\mathbf{x}_m + \boldsymbol{\delta}_{q,m}) \bmod \Lambda. \end{aligned} \quad (2.6)$$

The vector  $\boldsymbol{\delta}_{q,m}$  is the quantization error at the  $m^{\text{th}}$  relay node which is statistically independent of  $\mathbf{x}_m$  for  $m = 1, \dots, M$  and uniformly distributed over  $\mathcal{V}_{f,m}$  by Crypto Lemma which is stated in [103, Lemma 1]. As the number of channel uses increases without bound (i.e.,  $L \rightarrow \infty$ ), the cumulative density function of  $\boldsymbol{\delta}_{q,m}$  is upper bounded by the cumulative density function of an i.i.d zero-mean Gaussian vector whose covariance matrix is diagonal in [18, Lemma 8]. Thus, in this thesis, we use this upper bound for the quantization error  $\boldsymbol{\delta}_{q,m}$  which is assumed to be an i.i.d zero-mean Gaussian vector whose covariance matrix is  $\mathbf{U}_q = \text{diag} \{ \mu_1, \mu_2, \dots, \mu_M \}$  and  $\mu_m$  is the distortion level of the fine lattice at the  $m^{\text{th}}$  relay node. We can write (2.6) in a matrix form such that

$$\tilde{\mathbf{Y}} = (\mathbf{X} + \boldsymbol{\Delta}_q) \bmod \Lambda, \quad (2.7)$$

where

$$\begin{aligned}\mathbf{X} &= [\mathbf{x}_1^T, \dots, \mathbf{x}_M^T]^T \in \mathbb{R}^{M \times L}, \\ \Delta_q &= [\delta_1^T, \delta_2^T, \dots, \delta_M^T]^T \in \mathbb{R}^{M \times L}, \\ \tilde{\mathbf{Y}} &= [\tilde{\mathbf{y}}_1^T, \dots, \tilde{\mathbf{y}}_M^T]^T \in \mathbb{R}^{M \times L}.\end{aligned}$$

The observations of the relay nodes can then be estimated at the receiver side to be

$$\begin{aligned}\hat{\mathbf{X}} &= \mathbf{A}^{-1}([\mathbf{A}\tilde{\mathbf{Y}}] \bmod \Lambda) \\ &\stackrel{(a)}{=} (\mathbf{X} + \Delta_q),\end{aligned}\tag{2.8}$$

where  $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1^T, \dots, \hat{\mathbf{x}}_M^T]^T \in \mathbb{R}^{M \times L}$  is the estimate of the matrix  $\mathbf{X}$ ,  $\mathbf{A} \in \mathbb{Z}^{M \times M}$  is the IFSC integer coefficients matrix as defined in [24], and the equation (a) only holds with high probability as it follows from the properties of the modulo operation and the assumption that the coarse lattice is good for channel coding.

The receiver estimates the intended signals that were transmitted over  $L$  channel uses (i.e., the vector  $\mathbf{s}_1 = [s_1[0], s_1[1], \dots, s_1[L-1]]$ ) by combining the output of the  $M$  relay nodes to find  $\hat{\mathbf{s}}_1$  as in

$$\hat{\mathbf{s}}_1 = \mathbf{w}^T \hat{\mathbf{X}},\tag{2.9}$$

where  $\mathbf{w} \in \mathbb{R}^M$  is the receive combiner at the receiving user side.

## 2.2 Rate-Adaptive Integer Forcing Compression

In this section, we propose a novel integer forcing compression strategy that utilizes lattice vector quantization. We dub the proposed strategy RAIF compression. The construction of the lattice vector quantizer depends on the second moment of the fine and the coarse lattices. There is a correlation between the quantization rates and the second moments of the fine and coarse lattices. In particular, the quantization rate of the  $m^{\text{th}}$  relay node is

$$r_m = \min_{\substack{\mathbf{A} \in \mathbb{Z}^{L \times L} \\ \det(\mathbf{A}) \neq 0}} \max_{i=1, \dots, M} \frac{1}{2} \log \frac{\sigma_{\Lambda}^2(\mathbf{a}_i)}{\sigma_{\Lambda_{f,m}}^2}, \quad m = 1, \dots, M,\tag{2.10}$$

where  $\sigma_{\Lambda_{f,m}}^2$  is the second moment (distortion level) of the fine lattice at the  $m^{\text{th}}$  relay node because of the assumption of the fine lattice is good for quantization [24],  $\sigma_{\Lambda}^2(\mathbf{a}_i)$  is the second moment of the coarse lattice corresponding to  $\mathbf{a}_i$ , and  $\mathbf{a}_i$  is the  $i^{\text{th}}$  row of the matrix  $\mathbf{A}$ . Note that the integer coefficient matrix  $\mathbf{A}$  in the equation (2.10) is the same for all  $1 \leq m \leq M$ . Therefore, we can write

$$\sigma_{\Lambda_{f,m}}^2 = \mu_m. \quad (2.11)$$

The second moment of the coarse lattice with respect the  $i^{\text{th}}$  row of  $\mathbf{A}$  is [24]

$$\sigma_{\Lambda}^2(\mathbf{a}_i) = \mathbf{a}_i^T (\mathbf{Q}_{\mathbf{X}\mathbf{X}} + \mathbf{U}_q) \mathbf{a}_i, \quad (2.12)$$

where  $\mathbf{Q}_{\mathbf{X}\mathbf{X}} = \mathbb{E}[\mathbf{X}\mathbf{X}^T]$  is the correlation matrix of  $\mathbf{X}$ . Hence, the optimal coarse lattice is found by its second moment which depends on designing the optimal integer coefficient matrix  $\mathbf{A}$ . The matrix  $\mathbf{A}$  can be obtained by solving the optimization problem

$$r_m = \min_{\substack{\mathbf{A} \in \mathbb{Z}^{L \times L} \\ \det(\mathbf{A}) \neq 0}} \max_{i=1, \dots, M} \frac{1}{2} \log \frac{\mathbf{a}_i^T (\mathbf{Q}_{\mathbf{X}\mathbf{X}} + \mathbf{U}_q) \mathbf{a}_i}{\mu_m}. \quad (2.13)$$

Different from the case of symmetric quantization rate, we assume that each relay node is assigned a different quantization rate. We denote the allocated quantization rate at the  $m^{\text{th}}$  relay node to be  $\tilde{r}_m$ . Allocation of quantization rates is done such that the MSE caused by the quantization performed at the relay nodes is minimized. The second moments of the fine and the coarse lattice are all related to the distortion levels which, in turn, depend on the allocated quantization rates. Therefore, before being able to solve (2.13), we derive the optimal quantization rates that are allocated to each of the relay nodes.

Similar to [27], we introduce a rate-allocation optimization problem based on the minimization of the difference between the MSE of the system for the case when the receiver has access to the quantized observations of the relay nodes and the case when the receiver



has access to the unquantized observations of the relay nodes. The optimization problem is written as

$$\begin{aligned} & \underset{\tilde{r}_m}{\operatorname{argmin}} && \tilde{\mathbf{h}}_1^T \mathbf{U}_q \tilde{\mathbf{h}}_1, \\ & \text{subject to} && \sum_{m=1}^M r_m = r_{\text{tot}}, \end{aligned} \quad (2.14)$$

where  $\mathbf{U}_q$  is a function of the quantization rate ( $\tilde{r}_m$ ) as will be shown later in (2.16),  $\tilde{\mathbf{h}}_1 = \mathbf{C}^{-1} \mathbf{h}_1$ , and  $\mathbf{C}$  is the covariance matrix of the noise and interference during the fading block of length  $L$  use which is defined as

$$\mathbf{C} = \mathbb{E} \left[ \left( \sum_{k=2}^K \mathbf{h}_k \mathbf{s}_k + \mathbf{z} \right) \left( \sum_{k=2}^K \mathbf{h}_k \mathbf{s}_k + \mathbf{z} \right)^T \right].$$

The rate-allocation is described in the following corollary.

**Corollary 1.** *Given the coarse lattice is good for channel coding and the fine lattices are good for the quantization, where goodness is defined as in [24] and for any second moment of the coarse lattice  $\sigma_\Lambda^2(\mathbf{a}_i)$ , the optimal quantization rate is*

$$\tilde{r}_m = \frac{r_{\text{tot}}}{M} + \frac{1}{2M} \sum_{j=1}^M \log \left( \frac{\tilde{h}_{m,1}^2}{\tilde{h}_{j,1}^2} \right), \quad (2.15)$$

where  $\tilde{h}_{j,1}$  is the  $j^{\text{th}}$  entry of  $\tilde{\mathbf{h}}_1$ . In the case of having the negative quantization rates, we remove the corresponding relay nodes and redo the optimization problem on the rest of the relay nodes.

Given the optimal quantization rate of each relay node, we can find the optimal integer coefficient matrix  $\mathbf{A}$  that solves the optimization problem in (2.13). Since the optimization problem in (2.13) is NP-hard, a common way for finding  $\mathbf{A}$  is to apply the LLL algorithm [104] as shown in Algorithm 1. In this algorithm, we use the fact that we have the same coarse lattice to map the distortions of all the relay nodes to the distortion of the  $m^{\text{th}}$  relay node to be

$$\mu_j = \mu_m 4^{\tilde{r}_m - \tilde{r}_j} \quad \text{for } j = 1, \dots, M. \quad (2.16)$$

Therefore, instead of having the distortions of all the relay nodes affecting the optimization problem in (2.13), the algorithm maps the distortions of all relay nodes to the distortion

---

**Algorithm 1** Rate-Adaptive Integer Forcing Compression

---

```

procedure RAIF( $\mathbf{Q}_{\mathbf{X}\mathbf{X}}, \tilde{\mathbf{r}}$ )  $\triangleright \tilde{\mathbf{r}} = [\tilde{r}_1, \dots, \tilde{r}_M^T]$ 
  Initialization: Fix  $\mu_m = 10^{-3}$ ,  $\mu_m = \mu_j 4^{\tilde{r}_j - \tilde{r}_m}$ .
  Calculate:  $\mathbf{A} = \text{LLL}(\mathbf{Q}_{\mathbf{X}\mathbf{X}} + \mathbf{U}_q)$ .

  Calculate:  $r_m = \frac{1}{2} \log \left( \frac{\mathbf{a}_i^T (\mathbf{Q}_{\mathbf{X}\mathbf{X}} + \mathbf{U}_q) \mathbf{a}_i}{\mu_m} \right)$ .

  while  $(|r_m - \tilde{r}_m| > \epsilon \text{ OR } r_m > \tilde{r}_m)$  do
    Increase  $\mu_m$ 
    Calculate:  $\mathbf{A} = \text{LLL}(\mathbf{Q}_{\mathbf{X}\mathbf{X}} + \mathbf{U}_q)$ .

    Calculate:  $r_m = \frac{1}{2} \log \left( \frac{\mathbf{a}_i^T (\mathbf{Q}_{\mathbf{X}\mathbf{X}} + \mathbf{U}_q) \mathbf{a}_i}{\mu_m} \right)$ .

  return  $\mathbf{A}, \mathbf{U}_q$ .
end procedure

```

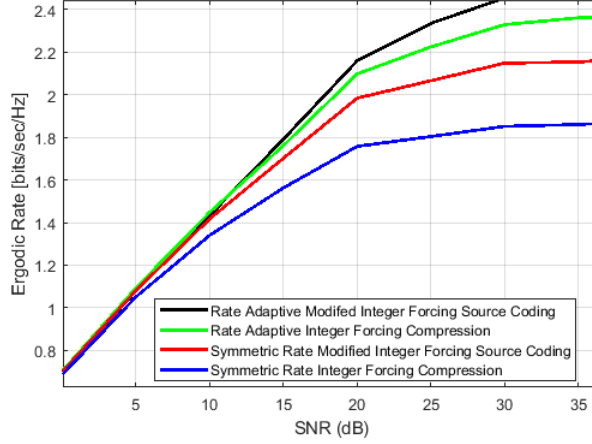
---

of the  $m^{\text{th}}$  relay node using (2.16). Hence, we can write  $\mathbf{U}_q$  in terms of  $\mu_m$  only. As mentioned earlier, we enforce the same coarse lattice for RAIF compression. Therefore, the computational complexity of the algorithm does not increase exponentially with the number of relay nodes as happens if multiple coarse lattices are used.

### 2.3 Simulation Results

In this section, Monte Carlo simulations are used to study the effect of different integer forcing source coding strategies, namely; RAIF compression, Modified IFSC [26] and IFSC [24], on the system throughput. We also compare against the throughput in the case of theoretical Gaussian approximation of the distortion which is shown in [27]. The effect of uplink channel variations on the throughput for each strategy is also studied. The uplink channels are modeled as spatially i.i.d Rayleigh fading channels. Four users are considered and the power of the transmitted signals from all users are equivalent. Our performance metric is the overall system throughput which is defined as

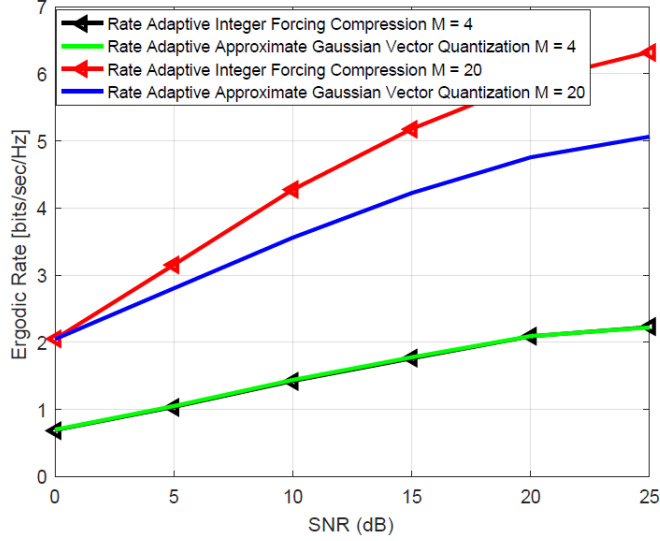
$$\mathcal{R} = \mathbb{E}_{\mathbf{h}_1} \left[ \frac{1}{2} \log \left( 1 + \text{SNR} \mathbf{h}_1^T (\mathbf{C} + \mathbf{U}_q)^{-1} \mathbf{h}_1 \right) \right].$$



**Figure 2.2.** Ergodic capacity performance for different IFSC-based compression strategies with four users and Rayleigh fading uplink channels and four relay nodes.

In Fig. 2, we assume that the relay network consists of four relay nodes and that the total number of quantization bits is 12 bits/channel use. In this figure, we show the throughput when the relay nodes use RAIF compression (defined in Algorithm 1), IFSC [24] (which can be derived by assuming equal rates in Algorithm 1), modified IFSC [26], and modified IFSC with optimized rates. It should be noted that even though the algorithm in [26] can be modified to run a different quantization rate for each relay node, it does not have a way to find these assigned rates. Therefore, we use the optimized quantization rates that we derived in (2.15).

As shown in Fig. 2, the throughput is higher when the optimal quantization rates are applied than when quantization rates are distributed equally among the relay nodes. When the values of SNR are higher than 30 dB, the throughput saturates for all the strategies and RAIF compression is 0.4 bits/sec/Hz higher than IFSC. Although the throughput of the rate-adaptive modified IFSC is slightly higher than our proposed algorithm, this comes at a cost of exponential increase in the computational complexity with the number of relay nodes. The exponential increase in computational complexity comes from the fact that rate-adaptive modified IFSC algorithm optimizes the distortion level of each relay node separately to achieve the optimal quantization rate assigned to this relay node.

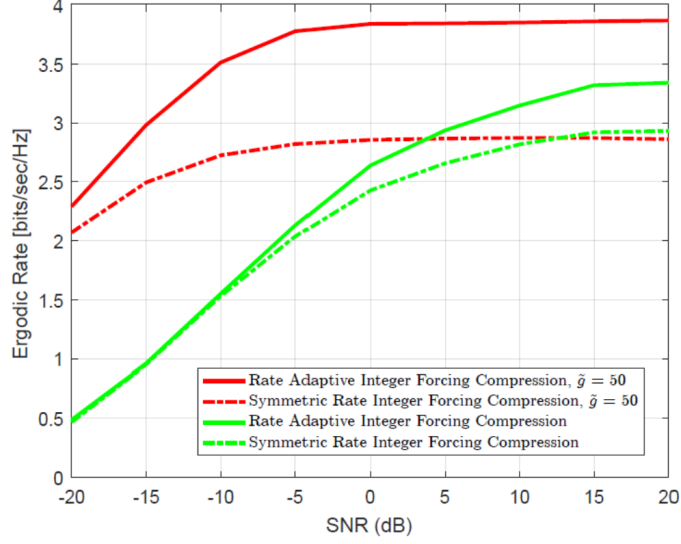


**Figure 2.3.** Ergodic capacity performance of RAIF compression and approximate Gaussian vector quantization for four users with Rayleigh fading uplink channels.

A comparison of the throughput of RAIF compression versus the theoretical throughput derived in [27] is shown in Fig. 3. In this figure, we consider two cases where the number of relay nodes are 4 and 20 while the total number of quantization bits at the relay network are 12 and 60 bits/channel use, respectively. As the number of the relay nodes increases, the throughput of our proposed algorithms becomes better than the throughput in [27]. The reason for this is that the quantization in [27] is done at each relay node separately while the relay nodes in our proposed algorithm are cooperating together.

In Fig. 4, we show how the quantization rate optimization in (2.15) makes the system sensitive to severe channel variations. The relay network contains ten relay nodes and the total number of quantization bits is 30 bits/channel use. The uplink channels are distributed as  $\mathcal{N}(0, 1)$ . Two cases are considered. In the first case, we assume that the gain of the uplink channel between the intended transmit user and one of the relay nodes which is denoted as  $\tilde{g}$  is higher than the average gain of the other uplink channels. In the second case, we consider all uplink channels are generated without severe channel variations.

The throughput of RAIF compression and IFSC are saturated at 0 dB in the first case because  $\tilde{g} = 50$  is applied into the system. The throughput of RAIF compression is almost



**Figure 2.4.** Ergodic capacity performance of RAIF compression and approximate Gaussian vector quantization for a relay network of ten relay nodes - the number of users is four with a strong uplink channel between the intended user and the first relay node.

one bit/sec/Hz better than IFSC that is using symmetric quantization rates. At 15 dB, the throughputs of both cases are saturated. Moreover, the gap of the throughput between RAIF compression and IFSC in the first case is twice as bit as the gap of the throughput between RAIF compression and IFSC in the second case. This shows that our proposed algorithm is adaptive to changes in the uplink channel conditions.

### 3. NONPARAMETRIC DECENTRALIZED DETECTION AND SPARSE SENSOR SELECTION VIA ONLINE KERNEL SCALAR QUANTIZATION

#### 3.1 System Model and Preliminaries

In this section, we first describe the system model (Section 3.1.1). Then, we provide basic background on kernels and decision functions which will be later employed (Section 3.1.2).

##### 3.1.1 Wireless Sensor Network System Model

We consider a decentralized problem consisting of  $M$  sensors connected to a common fusion center as shown in Fig. 4.1. The fusion center must utilize the sensors to train a decision function to determine the true class, which is represented by  $y \in \mathcal{Y}$ , representing a statistical model for the target (dependent) variable of interest.  $\mathcal{Y}$  is considered as a finite discrete set as we will focus on the cases of classification. Although the proposed system model and methods are applicable for any multi-class problem, our *prime focus* is the binary classification case, i.e.,  $\mathcal{Y} \in \{-1, 1\}$ , in this paper.

We assume perfect synchronization among sensors. The fusion center *primarily* i) trains a decision function chosen from a function space  $\mathcal{H}$  to determine the class and ii) determines a quantization rule to be employed by each sensor's quantizer in order to maximize classification performance. Followed by this, we are interested in making the fusion center iii) utilize  $M$  of  $M$  quantized observations which can promote the network power efficiency, for  $1 \leq m \leq M$ .

Ideally, the fusion center and quantizers would have access to  $P(\mathbf{X}, Y)$  of the sensor observations and the class, where  $Y \in \mathcal{Y}$  and  $\mathbf{X} = [X_1, X_2, \dots, X_M]^T$ . However, we consider a non-parametric decentralized detection approach where  $P(\mathbf{X}, Y)$  is unknown. To overcome this lack of knowledge, we solve this problem based on a sequence of training samples  $\mathcal{S} = (s_n)_{n=1}^N$  such that  $s_n = (\mathbf{x}_n, y_n)$ , where  $x_{m,n} \in \mathcal{X}_m$  is element  $m$  of  $\mathbf{x}_n \in \mathcal{X}$ ,  $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_M$ , and  $y_n \in \mathcal{Y}$  for all  $n$ . We assume  $\mathcal{X}_m$  and  $\mathcal{X}$  are finite sets for all  $m$ . Note that we use a

subscript  $n$  to denote the time index. This motivate us to leverage empirically-driven learning techniques, e.g., OL and BL approaches.

Formally, we collect the sensor observations at time  $n$  into a vector  $\mathbf{x}_n$  and the quantized outputs into a vector  $\mathbf{q}_n = [q_{1,n} \cdots q_{M,n}]^T \in \mathcal{Q}$  such that  $\mathcal{Q} = \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_M$ . We assume  $\mathcal{Q}_m$  and  $\mathcal{Q}$  are assume to be finite sets such that  $|\mathcal{Q}_m| \leq |\mathcal{Q}|$ ,  $|\mathcal{Q}_m| \leq |\mathcal{X}_m|$ , and  $|\mathcal{Q}| \leq |\mathcal{X}|$ , for all  $m$ . Each element of  $\mathbf{q}_n$ , e.g.,  $q_{m,n}$ , is produced based on the quantization rule  $P_{m,n}(q_{m,n}|x_{m,n})$  such that  $P_{m,n}(q_{m,n}|x_{m,n}) \in [0, 1]$  and  $\sum_{q_{m,n} \in \mathcal{Q}_m} P_{m,n}(q_{m,n}|x_{m,n}) = 1$  follows axioms of probability [105], for all  $q_{m,n} \in \mathcal{Q}_m$  and  $x_{m,n} \in \mathcal{X}_m$ . Moreover, the quantization rules across the sensors inspired by [38]–[40], are assumed to be

$$P_n(\mathbf{q}_n|\mathbf{x}_n) = \prod_{m=1}^M P_{m,n}(q_{m,n}|x_{m,n}), \quad (3.1)$$

for all possible  $\mathbf{q}_n$  and  $\mathbf{x}_n$ . As a result, we can form a set  $\mathcal{P}$  of the feasible quantization rules across the sensors as

$$\mathcal{P} = \left\{ P(\mathbf{Q}|\mathbf{X}) : \forall \mathbf{q} \in \mathcal{Q}, \mathbf{x} \in \mathcal{X}, P(\mathbf{q}|\mathbf{x}) \geq 0, \right. \\ \left. \sum_{\mathbf{q} \in \mathcal{Q}} P(\mathbf{q}|\mathbf{x}) = 1, \text{ and } P(\mathbf{q}|\mathbf{x}) = \prod_{m=1}^M P_m(q_m|x_m) \right\}. \quad (3.2)$$

After quantization, the WSN forwards  $\mathbf{q}_n$  to the fusion center. Similar to [38]–[40], we assume that the fusion center perfectly observes  $\mathbf{q}_n$ , i.e., the channel does not introduce significant noise to the quantized measurements. The fusion center is assumed to know the behavior of the quantizers perfectly (i.e., quantization rules across the sensors,  $P_n(\mathbf{q}_n|\mathbf{x}_n)$ ) at time  $n$ . Hence, the fusion center jointly trains the decision function, updates quantization rules, and performs sensor subset selection, which will be discussed in Section 3.2.2.

### 3.1.2 Kernel and Decision Function

#### Kernel

We are interested in developing an OL algorithm that can work with both linearly and non-linearly separable statistical models of the quantized data. One possible approach to

dealing with non-linearly separable statistical models is to map a vector to a high-dimensional space  $\mathcal{H}$  using a mapping function  $\phi(\cdot) : \mathcal{Q} \rightarrow \mathcal{H}$  [38]–[42], [49], [106]. Generally, this mapping function is not unique. An algorithm is developed based on the inner product between points, i.e.,  $\langle \phi(\mathbf{q}^\tau), \phi(\mathbf{q}^\tau) \rangle \in \mathbb{R}$ , for all possible quantization points  $\mathbf{q}^\tau, \mathbf{q}^\tau \in \mathcal{Q}$  so that the problem is solved using one or more (approximately) linear decision boundaries [38]–[42], [49], [106]. Although it might be challenging to construct  $\phi(\cdot)$ , it is easy to construct the output of  $\langle \phi(\mathbf{q}^\tau), \phi(\mathbf{q}^\tau) \rangle$  [106]. To accomplish this, we utilize kernel-based techniques to replace  $\langle \phi(\mathbf{q}^\tau), \phi(\mathbf{q}^\tau) \rangle$  using the definition of the kernel [106] below.

**Definition 3.1.1.** *A real-valued symmetric function  $k(\cdot, \cdot) : \mathcal{Q} \times \mathcal{Q} \rightarrow \mathbb{R}$  is called a kernel if the Gram matrix  $\mathbf{K} = [K_{\tau, \tau}]$ ,  $K_{\tau, \tau} = k(\mathbf{q}^\tau, \mathbf{q}^\tau)$  for  $\tau, \tau = 1, \dots, |\mathcal{Q}|$ , is positive semi-definite.*

The mapping function is related to the kernel according to  $\phi(\mathbf{q}^\tau) = k(\mathbf{q}^\tau, \cdot)$  so that  $\langle k(\mathbf{q}^\tau, \cdot), k(\mathbf{q}^\tau, \cdot) \rangle = k(\mathbf{q}^\tau, \mathbf{q}^\tau)$ . In this paper, we are interested in using a special kernel, called weighted kernel [40]. A weighted kernel  $k_{\mathbf{w}}(\mathbf{q}^\tau, \mathbf{q}^\tau)$  is used to convey a level of reliability from the sensors using a weight parameter vector  $\mathbf{w}$  that satisfies

$$\|\mathbf{w}\|_1 = M \text{ and } \mathbf{w} \succeq \mathbf{0}, \quad (3.3)$$

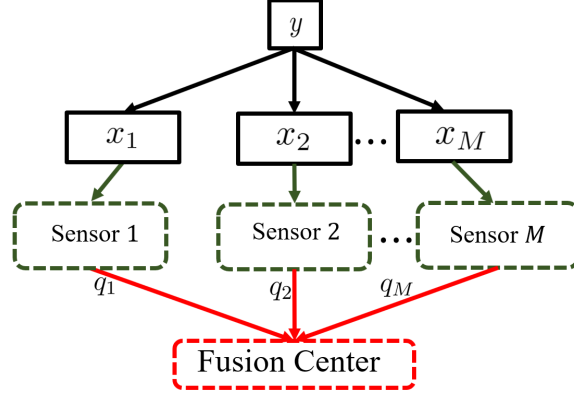
where  $\mathbf{w} \succeq \mathbf{0}$  indicates  $w_m \geq 0$ , for all  $m$ . Thus, a large weight parameter of a sensor (e.g.,  $w_m$ ) indicates that the corresponding sensor is more reliable [40]. We consider sensor  $m$  to be disabled when  $w_m = 0$  for some  $m = 1, 2, \dots, M$ . For this reason, the weighted kernel will be critical to our OL algorithm development. Note that the construction of the weighted kernel is not unique. An example weighted kernel is the weighted count kernel described by [39], [40]<sup>1</sup>

$$k_{\mathbf{w}}(\mathbf{q}^\tau, \mathbf{q}^\tau) = \sum_{m=1}^M w_m^2 \mathbb{1}(\tau = \tau). \quad (3.4)$$

---

<sup>1</sup>Another example inspired by [39] is the weighed linear kernel, e.g.,  $k_{\mathbf{w}}(\mathbf{q}^\tau, \mathbf{q}^\tau) = \sum_{m=1}^M w_m^2 q_m^\tau q_m^\tau$ .





**Figure 3.1.** Illusion of the generic decentralized detection system model

### Decision function

The purpose of the decision function is to estimate/predict the class of the input samples. For any given kernel (including the weighted kernel), the decision function induced by the mapping space  $\mathcal{H}$  can be defined as

$$f(\mathbf{q}^\tau) = \sum_{\tau=1}^{|\mathcal{Q}|} \alpha_\tau k_{\mathbf{w}}(\mathbf{q}^\tau, \mathbf{q}^\tau), \quad (3.5)$$

where  $\alpha_\tau \in \mathbb{R}$ . We assume that  $\mathcal{H}$  is a reproducing kernel Hilbert space (RKHS). A RKHS  $\mathcal{H}$  is one for which there is a kernel  $k_{\mathbf{w}} : \mathcal{Q} \times \mathcal{Q} \rightarrow \mathbb{R}$  and a dot product  $\langle \cdot, \cdot \rangle$  in  $\mathcal{H}$  such that

- $k_{\mathbf{w}}$  has the producing property:

$$f(\mathbf{q}^\tau) = \langle \mathbf{f}, k_{\mathbf{w}}(\mathbf{q}^\tau, \cdot) \rangle, \text{ for all } \mathbf{f} \in \mathcal{H} \text{ and } \tau. \quad (3.6)$$

- $\mathcal{H}$  is the closure of the space of all  $k_{\mathbf{w}}(\mathbf{q}^\tau, \cdot)$ , for all  $\tau$ .

We refer to  $\mathbf{f}$  as the weight of  $f(\cdot)$ , which is also expressed as  $\mathbf{f} = \sum_{\tau=1}^{|\mathcal{Q}|} \alpha_\tau k_{\mathbf{w}}(\cdot, \mathbf{q}^\tau)$  based on (3.5) and (3.6).

## 3.2 Estimation Model

In this section, we introduce the binary classification problem for the BL (Section III-A) and OL (Section III-B) settings.

### 3.2.1 Prior Work: Batch Learning Problem Formulation

We assume that the fusion center does not have access to  $P(\mathbf{X}, Y)$ . Hence, a BL approach [39], [40], [106] can be employed based on the i.i.d. sample sequence  $\mathcal{S}$  which will be useful in our subsequent development of the OL case. The BL approach is one of the popular, empirically-driven learning approaches when the dataset has been collected in advance. To simplify the presentation, we denote a solution set for the BL approach, which is  $\gamma = \{\mathbf{f}, \{P(\mathbf{q}_n|\mathbf{x}_n)\}_{n=1}^N, \mathbf{w}\}$  for all  $\mathbf{q}_n \in \mathcal{Q}$  and  $\mathbf{x}_n \in \mathcal{X}$ . The weight of the decision function  $\mathbf{f}$ , the quantization rules across the sensors  $P(\mathbf{q}_n|\mathbf{x}_n)$ , and the weight parameter vector  $\mathbf{w}$  are determined offline based on  $\mathcal{S}$ . We use **subscript**  $n$  in the BL setting denote the **sample index**.

An empirical risk function is introduced by [39], [40] to obtain optimal  $\gamma$ , which is given by

$$R_{\text{emp}}(\mathcal{S}) = \sum_{n=1}^N \sum_{\mathbf{q}_n \in \mathcal{Q}} \frac{1}{N} \ell(f(\mathbf{q}_n), y_n) P(\mathbf{q}_n|\mathbf{x}_n). \quad (3.7)$$

where  $\ell(\cdot, \cdot)$  is a loss function which can either be a differentiable or non-differentiable convex function. For example, the logistic loss is a differentiable function, which is

$$\ell(f(\mathbf{Q}), Y) := \log \left( 1 + \exp \left\{ -Y f(\mathbf{Q}) \right\} \right). \quad (3.8)$$

On the other hand, the soft margin loss function takes the following non-differentiable form [49], which is

$$\ell(f(\mathbf{Q}), Y) := \max \{0, \rho - Y f(\mathbf{Q})\}, \quad (3.9)$$

where  $\rho > 0$  is the margin parameter. (3.9) denoted as the hinge loss function when  $\rho = 1$ . Eqn. (3.9) is most notably used for the binary classification problem [38]–[40], [106] To

prevent model overfilling and perform sensor selection, the regularized empirical risk function described by [40] is given by

$$\bar{R}_{\text{reg}}(\mathcal{S}) = R_{\text{emp}}(\mathcal{S}) + \frac{\lambda_1 \|\mathbf{f}\|^2}{2} + \lambda_2 \|\mathbf{w}\|_1. \quad (3.10)$$

We treat the regularization parameters  $\lambda_1, \lambda_2 \geq 0$  as being fixed.  $\lambda_2 \geq 0$  controls how many reliable sensors are selected for the classification [40];  $\|\cdot\|_1$  helps enforce a sparse solution while retaining convexity.

It is difficult to learn the desired parameters in (3.10) based on expectation of the loss function over  $\mathbf{q}_n$  for all  $n$  [39], [40]. Moreover, it also requires huge computational power to obtain the desired parameters when we solve the upper bound of (3.11). To resolve this, [39] and [40] apply Jensen's inequality to work with the lower bound of (3.10),

$$R_{\text{reg}}(\mathcal{S}) = \sum_{n=1}^N \frac{\ell(\langle \mathbf{f}, \Phi(\mathbf{x}_n) \rangle, y_n)}{N} + \frac{\lambda_1 \|\mathbf{f}\|^2}{2} + \lambda_2 \|\mathbf{w}\|_1, \quad (3.11)$$

where  $\langle \mathbf{f}, \Phi(\mathbf{x}_n) \rangle$  is a *modified* version of the decision function. We define  $\Phi(\cdot) \in \mathcal{H}$  as a marginalized mapping function based on [39], [40]

$$\Phi(\mathbf{x}_n) = \sum_{\mathbf{q}_n \in \mathcal{Q}} P(\mathbf{q}_n | \mathbf{x}_n) k_{\mathbf{w}}(\mathbf{q}_n, \cdot). \quad (3.12)$$

Note that (3.10) and (3.11) are equal when the quantizers are *deterministic* [39], [40], i.e.,

$$P_m(q_{m,n} | x_{m,n}) = \begin{cases} 1, & \text{if } q_{m,n} = q_m, \\ 0, & \text{otherwise,} \end{cases} \quad \text{for some } q_m \in \mathcal{Q}_m, \quad (3.13)$$

for all  $m$  and sample  $n$ . Although the optimal solutions are generally derived based on the upper bound, the optimal  $\gamma$  can be derived based on (3.13), has been studied in [39], [40]. With this in hand, we move on formulating a risk function for the OL setting, which will build upon (3.11) for the BL case.

### 3.2.2 Online Learning Problem Formulation

Instead of solving for the parameters offline (i.e., a BL approach), one of the main contributions is to solve the problem in the online setting. Specifically, we need to update the weight of the decision function  $\mathbf{f}_n \in \mathcal{H}$ , the quantization rules across the sensors  $P_n(\mathbf{q}_n|\mathbf{x}_n) \in \mathcal{P}$ , and the weighted parameter vector  $\mathbf{w}_n$  satisfying (3.3) at each time  $n$ . To simplify the presentation, we will use  $(\gamma_n = \{\mathbf{f}_n, P_n(\mathbf{q}_n|\mathbf{x}_n), \mathbf{w}_n\})_{n=1}^N$  as a solution set sequence across the sensors per time  $n$ .

The risk function must accommodate new data points in each time. Thus, we formulate the following instantaneous regularized risk function, inspired by [40], [49],

$$R(s_n) = \ell(\langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle, y_n) + \frac{\lambda_1 \|\mathbf{f}_n\|^2}{2} + \lambda_{2,n} \|\mathbf{w}_n\|_1, \quad (3.14)$$

which is an approximation of  $R_{\text{reg}}(\mathcal{S})$  using the data at time  $n$ . As before,  $\lambda_1 > 0$  and  $\lambda_{2,n} \geq 0$  for all  $n$  are regularization terms. Note that  $\lambda_{2,n}$  controls how many sensors are disabled which evolves over time. The modified decision function at time  $n$  is  $\langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle$ , inspired by (3.11). The marginalized mapping function  $\Phi_n(\mathbf{x}_n)$  at time  $n$  is

$$\Phi_n(\mathbf{x}_n) = \sum_{\mathbf{q}_n \in \mathcal{Q}} P_n(\mathbf{q}_n|\mathbf{x}_n) k_{\mathbf{w}_{\cdot,n}}(\cdot, \mathbf{q}_n). \quad (3.15)$$

We define  $k_{\Phi_{i,n}}(\mathbf{x}_i, \mathbf{x}_n) = \langle \Phi_i(\mathbf{x}_i), \Phi_n(\mathbf{x}_n) \rangle$  the marginalized kernel [38]–[40], which is given by

$$k_{\Phi_{i,n}}(\mathbf{x}_i, \mathbf{x}_n) = \sum_{\mathbf{q}_i, \mathbf{q}_n \in \mathcal{Q}} P_i(\mathbf{q}_i|\mathbf{x}_i) P_n(\mathbf{q}_n|\mathbf{x}_n) k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n), \quad (3.16)$$

where  $k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n)$  is the weighted kernel with respect to time  $i$  and  $n$ .  $k_{\mathbf{w}_{\cdot,n}}(\cdot, \mathbf{q}_n)$  can be considered as the mapping function associated with  $\mathbf{w}_n$  at time  $n$ .

In this paper, we will consider a computationally friendly weighted marginalized kernel inspired by [107]<sup>2</sup>, namely, the weighted count marginalized kernel with time  $i$  and  $n$ , i.e.,

$$\begin{aligned} k_{\Phi_{i,n}}(\mathbf{x}_i, \mathbf{x}_n) &= \sum_{m=1}^M \sum_{q_{m,n}, q_{m,i} \in \mathcal{Q}_m} w_{m,i} w_{m,n} P_{m,i}(q_{m,i} | x_{m,i}) \\ &\quad \times P_{m,n}(q_{m,n} | x_{m,n}) \mathbb{1}(q_{m,i} = q_{m,n}, x_{i,m} = x_{n,m}), \end{aligned} \quad (3.17)$$

where an additional constraint  $x_{i,m} = x_{n,m}$  is added in  $\mathbb{1}(\cdot, \cdot)$  compared to [40], for all  $i = 1, 2, \dots, n$ . This change benefits the classification performance while ensuring the quantizer of each sensor is a(n) (approximate) deterministic quantizer. A detailed discussion is shown in Section 3.4.3. There are various ways to construct  $k_{\Phi_{i,n}}(\mathbf{x}_i, \mathbf{x}_n)$ , as long as the construction of  $k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n)$  in (3.16) follows Definition 3.1.1. In reality, the BL and OL approaches require a feedback system which is not our focus to modify/disable the quantizers across the sensors.

### 3.3 Online Learning Methodology

Inspired by [38]–[40], we develop an MSOKSQ algorithm exploiting (sub)-gradient descent-based methods to solve (3.14) by iteratively updating each element in  $(\gamma_n)_{n=1}^N$ . This is because (3.14) is convex assuming two of three elements in  $\gamma_n$  are given. In this paper, we focus on a non-differential loss function used in (3.14), where the *sub-gradients* of elements in  $\gamma_n$  prior to each update need to be derived. We can extend the results of the algorithm to any differentiable loss function since the sub-gradient of a differentiable loss function is its gradient [106]. Hence, after we initialize  $\mathbf{f}_1$ ,  $\mathbf{w}_1$ , and  $P_1(\mathbf{q}_1 | \mathbf{x}_1)$  for all possible  $\mathbf{q}_1 \in \mathcal{Q}$  and  $\mathbf{x}_1 \in \mathcal{X}$ , we first derive  $\mathbf{f}_n$  given  $P_n(\mathbf{q}_{n-1} | \mathbf{x}_{n-1})$  and  $\mathbf{w}_n$  at time  $n-1$  (Section 3.3.1), where  $P_n(\mathbf{q}_{n-1} | \mathbf{x}_{n-1})$  is the updated version of  $P_{n-1}(\mathbf{q}_{n-1} | \mathbf{x}_{n-1})$ . Then, we derive  $P_{n+1}(\mathbf{q}_n | \mathbf{x}_n)$  given  $\mathbf{f}_n$  and  $\mathbf{w}_n$  at time  $n$  (Section 3.3.2). Finally, we derive  $\mathbf{w}_{n+1}$  given  $\mathbf{f}_n$  and  $P_{n+1}(\mathbf{q}_n | \mathbf{x}_n)$  at time  $n$ , and Algorithm 2 (Section 3.3.3).

<sup>2</sup>↑ Another example is the marginalized weighed linear kernel with time  $i$  and  $n$ , i.e.,  $k_{\Phi_{i,n}}(\mathbf{x}_i, \mathbf{x}_n) = \sum_{m=1}^M w_{m,i} w_{m,n} \mathbb{E}[\phi(Q_m) | x_{m,i}] \mathbb{E}[\phi(Q_m) | x_{m,n}]$ .

### 3.3.1 Obtaining the weight of decision function $\mathbf{f}_n$

Motivated by [49], given  $P_n(\mathbf{q}_{n-1}|\mathbf{x}_{n-1})$  and  $\mathbf{w}_n$ , the update of  $\mathbf{f}_n$  is occurred at time  $n-1$  based on (3.14) using

$$\mathbf{f}_n = \mathbf{f}_{n-1} - \eta_{n-1} \partial_{\mathbf{f}_{n-1}} R(s_{n-1}), \quad (3.18)$$

where  $\eta_{n-1} \geq 0$  is the learning rate.  $\partial_{\mathbf{f}_{n-1}} R(s_{n-1})$  is the sub-gradient of  $R(s_{n-1})$  with respect to  $\mathbf{f}_{n-1}$ , which is

$$\begin{aligned} & \partial_{\mathbf{f}_{n-1}} R(s_{n-1}) \\ &= \partial_{\langle \mathbf{f}_{n-1}, \tilde{\Phi}_{n-1}(\mathbf{x}_{n-1}) \rangle} \ell(\langle \mathbf{f}_{n-1}, \tilde{\Phi}_{n-1}(\mathbf{x}_{n-1}) \rangle, y_{n-1}) \\ & \quad \times \tilde{\Phi}_{n-1}(\mathbf{x}_{n-1}) + \lambda_1 \mathbf{f}_{n-1} \end{aligned} \quad (3.19)$$

using the chain rule. We define  $\tilde{\Phi}_{n-1}(\cdot)$  as the updated marginalized mapping function at time  $n-1$  such that

$$\tilde{\Phi}_{n-1}(\mathbf{x}_{n-1}) = \sum_{\mathbf{q}_{n-1} \in \mathcal{Q}} P_n(\mathbf{q}_{n-1}|\mathbf{x}_{n-1}) k_{\mathbf{w}, n-1}(\cdot, \mathbf{q}_{n-1}). \quad (3.20)$$

Initializing  $\mathbf{f}_1 = \mathbf{0}$ , we update  $\mathbf{f}_n$  as

$$\mathbf{f}_n = \sum_{i=1}^{n-1} \alpha_{i,n} \tilde{\Phi}_i(\mathbf{x}_i), \quad (3.21)$$

where  $\alpha_{i,n}$  is the coefficient at time  $n$  of  $\tilde{\Phi}_i(\mathbf{x}_i)$ , which is,

$$\alpha_{i,n} = \begin{cases} -\eta_i \partial_{\langle \mathbf{f}_i, \tilde{\Phi}_i(\mathbf{x}_i) \rangle} \ell(\langle \mathbf{f}_i, \tilde{\Phi}_i(\mathbf{x}_i) \rangle, y_i), & \text{if } i = n-1, \\ (1 - \eta_{n-1} \lambda_1) \alpha_{i,n-1}, & \text{if } i < n-1. \end{cases} \quad (3.22)$$

To form  $\mathbf{f}_n$ , the new coefficient  $\alpha_{n-1,n}$  is generated with  $\tilde{\Phi}_{n-1}(\mathbf{x}_{n-1})$  based on the sub-gradient of  $\ell(\cdot, \cdot)$ , and  $\alpha_{i,n}$  on the previous  $\tilde{\Phi}_i(\mathbf{x}_i)$  are multiplied by  $(1 - \eta_n \lambda_1)$  for  $i = 1, 2, \dots, n-1$ . Followed by this, we can write the modified decision function, which is given by

$$\langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle = \sum_{i=1}^{n-1} \alpha_{i,n} k_{\Phi_{i,n}}(\mathbf{x}_i, \mathbf{x}_n). \quad (3.23)$$

---

**Algorithm 2** MSOKSQ algorithm for signal classification.

---

**procedure** MSOKSQ( $\mathcal{S}$ )

Set system parameters:  $N, M > 0$

Initialize hyper-parameters:  $\eta_1, \eta_1^P, \eta_1^W, \lambda_1, \lambda_{2,1} > 0$

Initialize hyper-parameters:  $\mathbf{f}_1 = \mathbf{0}$ ,  $\mathbf{w}_1 = \mathbf{1}$ , and  $P(\mathbf{q}|\mathbf{x}) \in \mathcal{P}$  for all possible  $\mathbf{q} \in \mathcal{Q}$  and  $\mathbf{x} \in \mathcal{X}$

**for**  $n = 1 : 1 : N - 1$  **do**

Calculate  $P_{m,n+1}(q_{m,n}|x_{m,n})$  using Proposition 3.3.1, (3.29), (3.30), and (3.31), given  $\mathbf{w}_n$  and  $\mathbf{f}_n$

Update  $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n)$  using (3.1)

Calculate  $w_{m,n+1}$  using Proposition 3.3.2, (3.38), and (3.40), given  $\mathbf{f}_n$  and  $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n)$

Update  $\mathbf{w}_{n+1}$  based on  $w_{m,n}$

Update  $\mathbf{f}_{n+1}$  using (3.21), given  $\mathbf{w}_{n+1}$  and  $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n)$

**return**  $(\gamma_n)_{n=1}^N$   
**end procedure**

---

### 3.3.2 Obtaining $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n)$ for all possible $\mathbf{q}_n$ and $\mathbf{x}_n$

We derive  $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n) \in \mathcal{P}$  given  $\mathbf{f}_n$  and  $\mathbf{w}_n$  for all possible  $\mathbf{q}_n \in \mathcal{Q}$  and some  $\mathbf{x}_n \in \mathcal{X}$ . This is done by first deriving  $P_{m,n+1}(q_{m,n}|x_{m,n})$  using a conditional (coordinate) sub-gradient method described by [38]–[40] for all  $q_{m,n} \in \mathcal{Q}_m$ , some  $x_{m,n} \in \mathcal{X}_m$ , and  $m = 1, 2, \dots, M$ . Then, we use the relationship described in (3.1) to obtain the resulting  $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n)$ . Note that  $P_{m,n+1}(q_{m,n}|x_{m,n})$  is the updated  $P_{m,n}(q_{m,n}|x_{m,n})$ . The following Theorem gives the sub-gradient of  $R(s_n)$  with respect to  $P_{m,n}(q_{m,n}|x_{m,n})$  for any weighted marginalized kernel at time  $n$ , for all  $m$ .

**Theorem 3.3.1.** *Given  $\mathbf{f}_n$ ,  $\mathbf{w}_n$ , and an i.i.d sequence  $\mathcal{S}$ , for any marginalized weighted kernel defined in (3.16),  $\partial_{P_{m,n}(q_{m,n}|x_{m,n})}R(s_n)$  is*

$$\begin{aligned} & \partial_{P_{m,n}(q_{m,n}|x_{m,n})}R(s_n) \\ &= \sum_{i=1}^{n-1} \sum_{\mathbf{q}_i, \mathbf{q}_n \in \mathcal{Q}} \mu_{1,n} \alpha_{i,n} P_i(\mathbf{q}_i|\mathbf{x}_i) \frac{P_n(\mathbf{q}_n|\mathbf{x}_n)}{P_{m,n}(q_{m,n}|x_{m,n})} \\ & \quad \times k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n), \end{aligned} \tag{3.24}$$

where  $\mathbf{q}_i$  and  $\mathbf{q}_n$  are possible quantization outputs across the sensors at time  $i$  and  $n$ , respectively, and

$$\mu_{1,n} = \partial_{\langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle} \ell(\langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle, y_n). \quad (3.25)$$

*Proof.* Using the chain rule, we can write

$$\begin{aligned} & \partial_{P_{m,n}(q_{m,n}|x_{m,n})} R(s_n) \\ &= \partial_{\langle \mathbf{f}_n, \Phi_n(s_n) \rangle} \ell(\langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle, y_n) \\ & \quad \times \partial_{P_{m,n}(q_{m,n}|x_{m,n})} \langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle. \end{aligned} \quad (3.26)$$

The first term on the right hand side (RHS) of (3.26) can be represented by (3.25). The second term on the RHS of (3.26) is

$$\begin{aligned} & \partial_{P_{m,n}(q_{m,n}|x_{m,n})} \langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle \\ & \stackrel{(a)}{=} \partial_{P_{m,n}(q_{m,n}|x_{m,n})} \sum_{i=1}^{n-1} \sum_{\mathbf{q}_i, \mathbf{q}_n \in \mathcal{Q}} \alpha_{i,n} P_i(\mathbf{q}_i | \mathbf{x}_i) \\ & \quad \times P_n(\mathbf{q}_n | \mathbf{x}_n) k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n), \end{aligned} \quad (3.27)$$

where (3.27a) holds based on (3.16) and (3.23). Thus, (3.24) is obtained by applying (3.1) and (3.27a), where

$$\partial_{P_{m,n}(q_{m,n}|x_{m,n})} P_n(\mathbf{q}_n | \mathbf{x}_n) = \frac{P_n(\mathbf{q}_n | x_n)}{P_{m,n}(q_{m,n}|x_{m,n})}. \quad \square$$

Based on Theorem 3.3.1, the computational power of computing the sub-gradient increases significantly when  $|\mathcal{Q}|$  increases. To resolve this, the following proposition provides a special form of  $\partial_{P_{m,n}(q_{m,n}|x_{m,n})} R(s_n)$  based on (3.17).

**Proposition 3.3.1.** *Given  $\mathbf{f}_n$ ,  $\mathbf{w}_n$ , and an i.i.d sequence  $\mathcal{S}$ ,  $\partial_{P_{m,n}(q_{m,n}|x_{m,n})} R(s_n)$  can be shown based on (3.17) as*

$$\begin{aligned} & \partial_{P_{m,n}(q_{m,n}|x_{m,n})} R(s_n) \\ &= \sum_{i=1}^{n-1} \mu_{1,n} \alpha_{i,n} w_{m,i} w_{m,n} P_{m,i}(q_{m,i} | x_{m,i}) \\ & \quad \times \mathbb{1}(q_{m,i} = q_{m,n}, x_{m,i} = x_{m,n}), \end{aligned} \quad (3.28)$$

where  $\mu_{1,n}$  is defined in (3.25).



*Proof.* As we use (3.26) in Theorem 3.3.1, we can further expand (3.27) using (3.17) to get (3.28).  $\square$

Since  $P_{m,n}(q_{m,n}|x_{m,n})$  must satisfy the summation constraint described in (3.2) for all  $m$  and  $n$ , we can update  $P_{m,n}(q_{m,n}|x_{m,n})$  to  $P_{m,n+1}(q_{m,n}|x_{m,n})$  through the conditional sub-gradient method with a simplex problem [39], [40], [108]. Note that a simplex constraint refers to the summation in (3.2). We focus on this particular method because the quantizers will become to deterministic quantizers after some time instance which will be discussed in Section (3.4.3) in detail. The OL setting considered in this paper has the added complexity of dealing with data per time. We must update  $P_{m,n}(q_{m,n}|x_{m,n})$  at time  $n$  for some  $x_{m,n} \in \mathcal{X}_m$  that is observed, as opposed to all possible  $x_{m,n} \in \mathcal{X}_m$  comparing to the BL setting [39], [40], simultaneously. Hence,  $P_{m,n+1}(q_{m,n+1}|x_{m,n+1})$  does not necessarily correspond to the updated version of  $P_{m,n}(q_{m,n}|x_{m,n})$ , e.g.,  $P_{m,n+1}(q_{m,n}|x_{m,n})$ , because  $x_{m,n}$  may not equal  $x_{m,n+1}$ , for all possible  $q_{m,n}, q_{m,n+1} \in \mathcal{Q}_m$  and some  $n = 1, 2, \dots, N - 1$ .

For this reason, we introduce a time index  $n_{i_x} \in \{1, 2, \dots, N\}$  with a positive integer  $i_x$ , where  $i_x = 1, 2, \dots, N_x$ , (e.g.,  $N_x \leq N$ ) represents the number of times that sensor  $m$  observes  $x$ , i.e.,  $x_{n_{i_x}} = x$  for some  $x \in \mathcal{X}_m$ , when  $\mu_{1,n_{i_x}} \neq 0$  and  $\alpha_{i,n_{i_x}} \neq 0$  considering (3.24) for all  $i = 1, 2, \dots, n_{i_x}$ . One benefit of introducing this time index is to make sure that the quantization rules strictly follows the summation described in (3.2). If the updated quantization rules do strictly follow the summation, then it contradicts **axioms of probability**. Hence, to strictly follow the summation constraint, we show the updated as follows. If  $\partial_{P_{m,n_{i_x}}(q_{m,n_{i_x}}|x_{m,n_{i_x}})} R(s_{n_{i_x}}) < 0$ , then the update of  $P_{m,n_{i_x}}(q_{m,n_{i_x}}|x_{m,n_{i_x}})$  inspired by [108], is given by

$$\begin{aligned} & P_{m,n_{i_x}+1}(q_{m,n_{i_x}}|x_{m,n_{i_x}}) \\ &= (1 - \eta_{n_{i_x}}^P) P_{m,n_{i_x}}(q_{m,n_{i_x}}|x_{m,n_{i_x}}) - \mathbb{1}(q_{m,n_{i_x}} = q_m) \\ & \quad \times \eta_{n_{i_x}}^P \operatorname{sgn}\left(\partial_{P_{m,n_{i_x}}(q_{m,n_{i_x}}|x_{m,n_{i_x}})} R(s_{n_{i_x}})\right). \end{aligned} \tag{3.29}$$

However, if  $\partial_{P_{m,n_{i_x}}(q_{m,n_{i_x}}|x_{m,n_{i_x}})} R(s_{n_{i_x}}) \geq 0$ , then the updates are not performed, e.g.,

$$P_{m,n_{i_x}+1}(q_{m,n_{i_x}}|x_{m,n_{i_x}}) = P_{m,n_{i_x}}(q_{m,n_{i_x}}|x_{m,n_{i_x}}) \tag{3.30}$$

for all  $q_{m,n} \in \mathcal{Q}_m$ .  $\eta_{n_{i_x}}^P > 0$  is the learning rate of the quantization rule.  $q_m$  is picked as

$$q_m = \operatorname{argmax}_{q_m \in \mathcal{Q}_m} \left\| \partial_{P_{m,n_{i_x}}(q_{m,n_{i_x}}=q_m|x_{m,n_{i_x}})} R(s_{n_{i_x}}) \right\|. \quad (3.31)$$

Followed by this, the quantization rule when sensor  $m$  observes  $x$  for the  $i_x + 1$  times at time  $n_{i_x} + 1$  is given by

$$P_{m,n_{i_x}+1}(q_{m,n_{i_x}+1}|x_{m,n_{i_x}+1}) = P_{m,n_{i_x}+1}(q_{m,n_{i_x}}|x_{m,n_{i_x}})$$

because  $x_{m,n_{i_x}+1} = x_{m,n_{i_x}}$ . Then, the expression of  $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n)$  follows (3.1) and the fact

$$P_{m,n+1}(q_{m,n}|x_{m,n}) = P_{m,n_{i_x}+1}(q_{m,n_{i_x}}|x_{m,n_{i_x}}), \text{ for all } m. \quad (3.32)$$

### 3.3.3 Obtaining weighted parameter vector $\mathbf{w}_{n+1}$

Finally, we consider obtaining the updated weight parameter vector  $\mathbf{w}_{n+1}$  given  $\mathbf{f}_n$  and  $P_{n+1}(\mathbf{q}_n|\mathbf{x}_n)$ . Recalling the purpose of introducing  $\mathbf{w}_n$  is to properly select the  $M$  most reliable sensors among all sensors in the classification model. Thus, the developed OL algorithm updates  $\mathbf{w}_{n+1}$  to promote sparse sensor selection. The update will be terminated when at most  $M$  elements of  $\mathbf{w}_{n+1}$  are not 0.

To make a fair reliability comparison, we initialize  $w_{1,1} = \dots = w_{M,1}$  and  $\lambda_{2,1} \geq 0$ . We update  $\mathbf{w}_{n+1}$  using one of the conditional sub-gradient methods that can enforce the sparse solution, e.g., the project sub-gradient method, to update each  $w_{m,n+1}$  individually. Thus, we first derived a general expression of the sub-gradient of  $R(s_n)$  with respect to  $w_{m,n}$  using the chain rule for any weighted marginalized kernel at sensor  $m$ , which is given by

$$\partial_{w_{m,n}} R(s_n) = \mu_{2,n} \partial_{w_{m,n}} \langle \mathbf{f}_n, \tilde{\Phi}_n(\mathbf{x}_n) \rangle + \lambda_{2,n}, \quad (3.33)$$

where  $\tilde{\Phi}_n(\mathbf{x}_n)$  defined by (3.20) is the updated marginalized mapping function at time  $n$ .  $\mu_{2,n}$  in (3.33) is given by

$$\mu_{2,n} = \partial_{\langle \mathbf{f}_n, \tilde{\Phi}_n(\mathbf{x}_n) \rangle} \ell(\langle \mathbf{f}_n, \tilde{\Phi}_n(\mathbf{x}_n) \rangle, y_n). \quad (3.34)$$

The term that multiplies  $\mu_{2,n}$  using (3.23) is given by

$$\begin{aligned} \partial_{w_{m,n}} \langle \mathbf{f}_n, \tilde{\Phi}_n(\mathbf{x}_n) \rangle &= \partial_{w_{m,n}} \sum_{i=1}^{n-1} \sum_{\mathbf{q}_i, \mathbf{q}_n \in \mathcal{Q}} \alpha_{i,n} P_{n+1}(\mathbf{q}_n | \mathbf{x}_n) \\ &\quad \times P_i(\mathbf{q}_i | \mathbf{x}_i) k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n). \end{aligned} \quad (3.35)$$

The computational power for solving (3.35) increases with  $|\mathcal{Q}|$ . Thus, the following proposition uses (3.17) to obtain  $\partial_{w_{m,n}} R(s_n)$  in order to reduce the computational power.

**Proposition 3.3.2.** *Given  $\mathbf{f}_n$ ,  $P_{n+1}(\mathbf{q}_n | \mathbf{x}_n)$ , and an i.i.d sequence  $\mathcal{S}$ ,  $\partial_{w_{m,n}} R(s_n)$  can be shown based on (3.17) as*

$$\begin{aligned} \partial_{w_{m,n}} R(s_n) &= \lambda_{2,n} + \sum_{i=1}^{n-1} \sum_{q_m \in \mathcal{Q}_m} \left( \mu_{2,n} \alpha_{i,n} w_{m,i} P_{m,i}(q_{m,i} | x_{m,i}) \right. \\ &\quad \left. \times P_{m,n+1}(q_{m,n} | x_{m,i}) \mathbb{1}(q_{m,i} = q_{m,n}, x_{m,i} = x_{m,n}) \right). \end{aligned} \quad (3.36)$$

*Proof.* The proof starts by considering (3.33) which can be written in terms of (3.34) and (3.35). Thus, we can end the proof by obtaining an expression using (3.17) for (3.35), which is

$$\begin{aligned} (3.35) &= \sum_{i=1}^{n-1} \sum_{q_{m,i} \in \mathcal{Q}_m} \alpha_{i,n} w_{m,i} P_{m,i}(q_{m,i} | x_{m,i}) \\ &\quad \times P_{m,n+1}(q_m | x_{m,n}) \mathbb{1}(q_{m,i} = q_{m,n}, x_{m,i} = x_{m,n}). \end{aligned} \quad (3.37)$$

Thus, we obtain (3.36) using (3.17), (3.33), and (3.37).  $\square$

Followed by this, we can employ a projected sub-gradient method to update  $\mathbf{w}_{n+1}$  while satisfying the constraint in (3.3), for all  $n$ . Additionally, given the objective of selecting  $M \leq M$  sensors to be kept on, the update is terminated once we reach a time instance  $n$  such that  $M - M$  elements of  $\mathbf{w}_{n+1}$  are zero. Thus, at time  $n$ , we update

$$w_{m,n+1} \begin{cases} \frac{\tilde{w}_m M}{\|\tilde{\mathbf{w}}\|_1}, & \text{if } \|\mathbf{w}_n\|_0 < M - M, \\ w_{m,n+1} = w_{m,n}, & \text{if } \|\mathbf{w}_n\|_0 = M - M, \end{cases} \quad (3.38)$$

to maintain the constraint (3.3), where the  $m$ th element of  $\tilde{\mathbf{w}}$  is

$$\tilde{w}_m = \max\{0, w_{m,n} - \eta_n^{\mathbf{w}} \partial_{w_{m,n}} R(s_n)\} \quad (3.39)$$

and  $\eta_n^{\mathbf{w}} > 0$  is learning rate with respect to  $\mathbf{w}_n$ . To accomplish  $M$  sensors are enabled, we can design the regularization constant  $\lambda_{2,n}$  such that at most one sensor is set to zero in (3.38) at each time  $n$ , i.e.,

$$\begin{aligned} \lambda_{2,n} = \min \bigg\{ & \{w_{m,n} - \partial_{w_{m,n}} \langle \mathbf{f}_n, \tilde{\Phi}_n(\mathbf{x}_n) \rangle \\ & \times \partial_{\langle \mathbf{f}_n, \tilde{\Phi}_n(\mathbf{x}_n) \rangle} \ell(\langle \mathbf{f}_n, \tilde{\Phi}_n(\mathbf{x}_n) \rangle, y_n) \}_{m=1}^M \bigg\}. \end{aligned} \quad (3.40)$$

There are various control settings that can be imposed here; for example, combining (3.38)–(3.40) where we set  $\eta_n^{\mathbf{w}} = 1$  will shorten number of time instance to achieve the similar goal. Additionally, other settings for  $\lambda_{2,n}$  which will result in  $\|\mathbf{w}_n\|_0 = M - M$  while yielding possibly different values of  $\mathbf{w}_{n+1}$  comparing to the described equations (3.38)–(3.40). In this paper, we focus on disabling a sensor per time to make the algorithm adjustable over time. For example, a sensor could be the most reliable sensor at time  $n$ , but it might become the less reliable sensor at time  $n + 1$  due to the external environment or interference which have not been made here.

Algorithm 2 summarizes the procedure developed in this section when  $M$  sensors are enabled. Given the fact that we turn to the non-differentiable, soft margin loss function from (3.9), we can express (3.22), (3.25), and (3.34) based on

$$\partial_{\langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle} \ell(\cdot, \cdot) = \begin{cases} -y_n, & \text{if } y_n \langle \mathbf{f}_n, \Phi_n(\mathbf{x}_n) \rangle \leq \rho, \\ 0, & \text{otherwise.} \end{cases} \quad (3.41)$$

### 3.4 Performance Analysis

In this section, we analyze the classification performance of the OL setting. Thus, we first provide reasonable assumptions for the OL and BL approaches similar to [39], [40], [49] (Section 3.4.1). Then, we study the relationship between the OL and BL approaches

(Section 3.4.2). Followed by this, we discuss the update behavior of the quantization rules across the sensors and weighted parameter vector (Section 3.4.3). Finally, the convergence analysis is made between the OL and the BL approaches (Section 3.4.4).

### 3.4.1 Assumptions

We define a set  $\Gamma$  that contains all possible  $(\gamma_n)_{n=1}^N$  for the OL setting and  $\gamma$  for the BL setting for all elements of  $(\gamma_n)_{n=1}^N$  and  $\gamma$ . Moreover, we define a **deterministic set**  $\Gamma_d \subset \Gamma$  such that the weight of the decision function and the weight parameter vector are jointly derived based on the quantization rules defined in (3.13) for both the OL and the BL setting. We assume that  $\mathbf{x}_n$  is the  $n$ th sample at the  $n$ th time of the OL setting while  $\mathbf{x}_n$  represents the  $n$ th sample of the BL setting.

Given a fix i.i.d. sequence  $\mathcal{S}$  and the weighted parameter vectors with respect to the OL setting  $\mathbf{w}_n$  and BL setting  $\mathbf{w}$  satisfy (3.3), there is  $\Psi > 0$  such that  $\|\Phi_n(\mathbf{x}_n)\|, \|\Phi(\mathbf{x}_n)\| \leq \Psi$  inspired by (3.16), for all  $n$  and  $\mathbf{x}_n \in \mathcal{X}$ . Note that  $\Phi_n(\mathbf{x}_n)$  and  $\Phi(\mathbf{x}_n)$  relate to the OL and BL setting, respectively. We assume that  $\mathbf{f}$  derived by the BL approach is bounded such that  $\|\mathbf{f}\| \leq F$  for any  $\mathbf{f} \in \mathcal{H}$ , where

$$F \geq \frac{c\Psi}{\lambda_1}, \text{ for } \lambda_1 > 0, \text{ for some } c > 0. \quad (3.42)$$

The value of  $c > 0$  is derived based on the loss function  $\ell(\cdot, \cdot)$  that satisfies the Lipschitz condition [49]

$$\|\ell(\xi_1, y) - \ell(\xi_2, y)\| \leq c\|\xi_1 - \xi_2\|, \quad (3.43)$$

where  $\xi_1, \xi_2 \in \mathbb{R}$  and  $y \in \mathcal{Y}$ , for some  $c$ . Hence, for the OL approach, given  $\|\mathbf{f}_1\| = 0$ , we bound  $\|\mathbf{f}_{n+1}\| \leq (1 - \lambda_1\eta_n)\|\mathbf{f}_n\| + \eta_n c\Psi$  considering (3.19) for all possible  $\mathbf{f}_n$  and  $n$ . Hence, we need to set  $\|\mathbf{f}_n\| \leq \frac{c\Psi}{\lambda_1}$  in order to maintain the fact  $\|\mathbf{f}_{n+1}\| \geq 0$ , for all  $n$ . As we set  $\|\Phi_n(\mathbf{x}_n)\| \leq \Psi$  and  $\|\mathbf{f}_n\| \leq \frac{c\Psi}{\lambda_1}$ , we obtain  $\|\partial_{\mathbf{f}_n} R(s_n)\| \leq 2c\Psi$ .

### 3.4.2 Connection Between OL and BL Approach

Inspired by [41], [42], [49], [106], the performance of the OL approach is analyzed theoretically based on the relationship between the average instantaneous regularized risk function related to the OL setting, i.e.,

$$R_{\text{avg}}(\mathcal{S}) = \frac{1}{N} \sum_{n=1}^N R(s_n) \quad (3.44)$$

and the regularized empirical risk function  $R_{\text{reg}}(\mathcal{S})$  defined in (3.11). Thus, we start this section by showing an inequality first.

As we consider (3.18), the inequality inspired by [49]

$$\begin{aligned} & \frac{1}{\eta_n} (\|\mathbf{f}_n - \mathbf{f}\|^2) - \frac{1}{\eta_{n+1}} \|\mathbf{f}_{n+1} - \mathbf{f}\|^2 \\ & \geq -4\eta_n c^2 \Psi^2 + 2R(s_n) - 2R_{\mathbf{f}}(s_n) \\ & \quad + \left( \frac{1}{\eta_n} - \frac{1}{\eta_{n+1}} \right) \|\mathbf{f}_{n+1} - \mathbf{f}\|^2 \end{aligned} \quad (3.45)$$

holds based on (3.43) and the convexity of  $R(\cdot)$  such that  $\langle \partial_{\mathbf{f}_n} R(s_n), \mathbf{f}_n - \mathbf{f} \rangle \geq R_{\mathbf{f}}(s_n) - R(s_n)$ , where

$$R_{\mathbf{f}}(s_n) = \ell(\langle \mathbf{f}, \Phi_n(\mathbf{x}_n) \rangle, y_n) + \frac{\lambda_1}{2} \|\mathbf{f}\|^2 + \lambda_{2,n} \|\mathbf{w}_n\|_1. \quad (3.46)$$

The only difference between  $R(s_n)$  and  $R_{\mathbf{f}}(s_n)$  is the weight of the decision function comparing (3.14) and (3.46).

**Lemma 1.** *Given an i.i.d sequence  $\mathcal{S}$ , suppose  $\gamma_n \in \Gamma$  is produced Algorithm 2 with  $\eta_n = \eta_1 n^{-0.5}$  and  $\eta_1 \in (0, 1]$ , for all  $n$ , then, for any  $\gamma \in \Gamma$  and some  $c > 0$ ,*

$$\begin{aligned} R_{\text{avg}}(\mathcal{S}) &= \sum_{n=1}^N \frac{cF \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\|}{2N} \\ &+ \sum_{n=1}^N \frac{c \|\lambda_{2,n} \|\mathbf{w}_n\|_1 - \lambda_2 \|\mathbf{w}\|_1\|}{2N} + R_{\text{reg}}(\mathcal{S}) + O(N^{-0.5}). \end{aligned} \quad (3.47)$$

*Proof.* We start the proof by first considering the RHS and left hand side (LHS) of (3.48), which is given by

$$\begin{aligned}
\frac{\|\mathbf{f}_1 - \mathbf{f}\|^2}{\eta_1} - \frac{\|\mathbf{f}_{n+1} - \mathbf{f}\|^2}{\eta_{n+1}} &\stackrel{(a)}{=} \sum_{n=1}^N \left[ \frac{\|\mathbf{f}_n - \mathbf{f}\|^2}{\eta_n} - \frac{\|\mathbf{f}_{n+1} - \mathbf{f}\|^2}{\eta_{n+1}} + \left( \frac{1}{\eta_n} - \frac{1}{\eta_{n+1}} \right) \|\mathbf{f}_{n+1} - \mathbf{f}\|^2 \right] \\
&\stackrel{(b)}{\geq} -8c^2\eta_1\Psi^2N^{0.5} + 2\sum_{n=1}^N R(s_n) - 2\sum_{n=1}^N R_{\mathbf{f}}(s_n) - \frac{4F^2N^{0.5}}{\eta_1} \\
&\stackrel{(c)}{\geq} \frac{-8c^2\eta_1\Psi^2}{N^{0.5}} + \frac{1}{N} \sum_{n=1}^N 2R(s_n) - 2R_{\text{reg}}(\mathcal{S}) - \frac{4F^2}{N^{0.5}\eta_1} \\
&\quad - c \sum_{n=1}^N \frac{\left( F \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\| + \|\lambda_{2,n}\|\mathbf{w}_n\|_1 - \lambda_2\|\mathbf{w}\|_1 \right)}{2N}.
\end{aligned} \tag{3.48}$$

Inspired by [49], we obtain (3.48a) by adding  $\sum_{n=1}^N \frac{\|\mathbf{f}_{n+1} - \mathbf{f}\|^2}{\eta_{n+1}} - \sum_{n=1}^N \frac{\|\mathbf{f}_{n+1} - \mathbf{f}\|^2}{\eta_n}$ . The first three terms on the RHS of (3.48b) are derived using  $\sum_{n=1}^N \eta_n \leq 2\eta N^{0.5}$  shown in [49] and (3.45). The fourth term on the RHS (3.48b) holds because  $\|\mathbf{f}_{n+1} - \mathbf{f}\|^2 \leq 4F^2$ ,

$$\sum_{n=1}^N \frac{1}{\eta_n} - \frac{1}{\eta_{n+1}} = \frac{1}{\eta_1} - \frac{1}{\eta_{N+1}} = \frac{1}{\eta_1} - \frac{(N+1)^{0.5}}{\eta_1},$$

and  $-N^{0.5} \leq 1 - (N+1)^{0.5}$ . We derive (3.48c) using

$$\begin{aligned}
&\frac{1}{N} \left( \sum_{n=1}^N R_{\mathbf{f}}(s_n) \right) - R_{\text{reg}}(\mathcal{S}) \\
&\stackrel{(a)}{\leq} \frac{1}{N} \sum_{n=1}^N \left\| \ell(\langle \mathbf{f}, \Phi_n(\mathbf{x}_n) \rangle, y_n) \right. \\
&\quad \left. - \ell(\langle \mathbf{f}, \Phi(\mathbf{x}_n) \rangle, y_n) + (\lambda_{2,n}\|\mathbf{w}_n\|_1 - \lambda_2\|\mathbf{w}\|_1) \right\| \\
&\stackrel{(b)}{\leq} \frac{c}{N} \sum_{n=1}^N \left( F \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\| \right. \\
&\quad \left. + \|\lambda_{2,n}\|\mathbf{w}_n\|_1 - \lambda_2\|\mathbf{w}\|_1 \right),
\end{aligned} \tag{3.49}$$

where (3.49a) follows Jensen's inequality and the weights of  $\langle \mathbf{f}, \Phi_n(\mathbf{x}_n) \rangle$  and  $\langle \mathbf{f}, \Phi(\mathbf{x}_n) \rangle$  are the same. (3.49b) follows (3.43) and Cauchy Schwartz inequality. Now, we show

$$\frac{\|\mathbf{f}_1 - \mathbf{f}\|^2}{\eta_1} - \frac{\|\mathbf{f}_{n+1} - \mathbf{f}\|^2}{\eta_{n+1}} \leq \frac{F^2}{\eta_1} - 0, \tag{3.50}$$

where (3.50) holds based on the facts  $\mathbf{f}_1 = \mathbf{0}$  and  $\|\mathbf{f}_{n+1} - \mathbf{f}\| \geq 0$ . Then, (3.47) is justified considering (3.48) and (3.50).  $\square$

Lemma 1 establishes the relationship between the OL and the BL approach. However, the derived solution in (3.47) is very loose which motivates us to further investigate Algorithm 2 and assumptions that can be used to tighten (3.47).

### 3.4.3 Quantization Rules and Weight Parameter Vector Analysis

To tighten up the bound shown in (3.47), the goal is to close the gap introduced by (3.47). The first term of (3.47) relates to the quantization rules and the weighed parameter vector while the second term relates to the weight parameter vector. Thus, we need to study how Algorithm 2 affects the updates of the quantization rules and the weighted parameter vector per time when the soft margin loss function (3.9) for the binary classification problem is considered.

#### Quantization Rules Analysis

The quantization rules across the sensors have significant impacts on the classification performance which will be shown in Section 3.5. However, the update behaviors of the quantization rules across the sensors have never been discussed even in the BL [39], [40] and OL approaches [38]. Thus, we are interested in studying the update behaviors of the quantization rules across the sensors based on (3.17) with respect to time  $n$ , for all  $n$ . Before we provide the detailed explanations, some facts need to be addressed.

Recalling  $n_{i_x} \in \{1, 2, \dots, N\}$  is the time index, where  $i_x$  represents the number of times that sensor  $m$  observes  $x$  when  $\mu_{1,n_{i_x}} \neq 0$  and  $\alpha_{i,n_{i_x}} \neq 0$  for all  $n_{i_x}$ . The values of  $\alpha_{i,n_{i_x}}$  and  $\mu_{1,n_{i_x}}$  based on (3.22) and (3.25) are affected by  $\rho$  defined in (3.9), for all  $n_{i_x}$ . Hence, for any  $x \in \mathcal{X}$ , as we consider (3.22), (3.25), and (3.41), the following relationship is given by

$$\text{sgn}(\alpha_{i,n_{i_x}} \mathbb{1}(x_{m,i} = x_{m,n_{i_x}})) = -\text{sgn}(\mu_{1,n_{i_x}}). \quad (3.51)$$



This indicates that

$$\text{sgn}\left(\partial_{P_{m,n_{i_x}}}(q_{m,n_{i_x}}|x_{m,n_{i_x}})R(s_{n_{i_x}})\right) < 0, \text{ for all } q_{m,n_{i_x}} \in \mathcal{Q}_m \quad (3.52)$$

Hence, the update of  $P_{m,n_{i_x}+1}(q_{m,n_{i_x}}|x_{m,n_{i_x}})$  occurs when  $\partial_{P_{m,n_{i_x}}}(q_{m,n_{i_x}}|x_{m,n_{i_x}})R(s_{n_{i_x}}) < 0$  for some  $q_{m,n_{i_x}} \in \mathcal{Q}_m$ . Additionally,  $P_{m,n_{i_x}}(q_{m,n_{i_x}} = q_m|x_{m,n_{i_x}})$  is increased by  $\eta_{n_{i_x}}^P - \eta_{n_{i_x}}^P P_{m,n_{i_x}}(q_{m,n_{i_x}} = q_m|x_{m,n_{i_x}}) > 0$  while  $P_{m,n_{i_x}}(q_{m,n_{i_x}} \neq q_m|x_{m,n_{i_x}})$  is reduced by  $\eta_{n_{i_x}}^P$  based on (3.29).

Before we introduce the following proposition, we first initialize

$$P_{m,n_1}(q_{m,n_1}|x_{m,n_1}) = P_m(q_m|x_m), \quad (3.53)$$

for  $n_1 \in \{1, 2, \dots, N\}$  and  $m = 1, 2, \dots, M$  such that

$$P_m(q_m|x_m) = \max_{q_m \in \mathcal{Q}_m} \{P_m(q_m|x_m)\} \quad (3.54)$$

for all  $x_m \in \mathcal{X}_m$ , and some  $q_m \in \mathcal{Q}_m$ . Then,  $q_m$  in (3.54) is guaranteed to be selected at each iteration based on (3.22), (3.29), (3.31), and (3.51)–(3.53). This motivate us to find the expressions of the quantization rules across the sensors over time

**Proposition 3.4.1.** *Given an i.i.d sequence  $\mathcal{S}$ , suppose quantization rules across the sensors are initialized based on (3.53) and (3.54), and  $\ell(\cdot, \cdot)$  defined in (3.9) with a properly selected  $\rho$  such that  $\{\partial_{P_{m,n_{i_x}}}(q_{m,n_{i_x}}|x_{m,n_{i_x}})R(s_{n_{i_x}}) < 0\}_{i_x=1}^{N_x}$ , then  $P_{m,n_{N_x}}(q_{m,n_{N_x}}|x_{m,n_{N_x}})$  is updated based on (3.17) as*

$$P_{m,n_{N_x}+1}(q_{m,n_{N_x}}|x_{m,n_{N_x}}) = \begin{cases} 1 - O((1 - \eta)^{N_x}), \\ \quad \text{if } q_{m,n_{N_x}} = q_m, \\ O((1 - \eta)^{N_x}), \\ \quad \text{otherwise,} \end{cases} \quad (3.55)$$

where a fixed  $q_m \in \mathcal{Q}_m$  is described by (3.54) and

$$\eta = 1 - \prod_{i_x=1}^{N_x} (1 - \eta_{n_{i_x}}^P)^{\frac{1}{N_x}}. \quad (3.56)$$

*Proof.* This proof starts with generalizing the expression of  $P_{m,n_{N_x+1}}(q_m, n_{N_x} \neq q_m | x_{m,n_{N_x}})$ , which is given by

$$\begin{aligned} P_{m,n_{N_x+1}}(q_m \neq q_m | x_{m,n_{N_x}}) &= \prod_{i_x=1}^{N_x} (1 - \eta_{n_{i_x}}^P) P_{m,n_1}(q_{m,n_1} \neq q_m | x_{m,n_1}) \\ &= (1 - \eta)^{N_x} P_{m,n_1}(q_{m,n_1} \neq q_m | x_{m,n_1}) \\ &= O((1 - \eta)^{N_x}) \end{aligned} \quad (3.57)$$

using (3.29), (3.31), (3.51)-(3.53), and (3.56). Followed by this, we use the summation described in (3.2) to obtain  $P_{m,n_{N_x+1}}(q_m, n_{N_x} = q_m | x_{m,n_{N_x}}) = 1 - O((1 - \eta)^{N_x})$ .  $\square$

Based on Proposition 3.4.1, as we set the learning rate  $\eta_{n_{i_x}}^P = 1$  for all  $n_{i_x}$ , the quantizer of sensor  $m$  will become a deterministic quantizer for all  $m$  when  $N_x > 1$ . If a fixed  $N_x$  is assumed to be large, e.g.,  $N_x \gg 1$ , then the quantizer of sensor  $m$  can be approximated as the deterministic quantizer when  $\eta_{n_{i_x}}^P \in (0, 1)$  for all  $n_{i_x}$  and  $m = 1, 2, \dots, M$ .

## Weighted Parameter Vector Analysis

The purpose of Algorithm 2 for solving  $\mathbf{w}_n$  is to handle the fact that the sensor output will vary in quality, and  $M - M$  sensors can be disabled with minimal effect on classification performance. Moreover, since the optimization problems and algorithms used for computing the desired solutions for the OL and BL approaches are different it is challenging to show that  $\mathbf{w}_n$  equals or approximately equals to  $\mathbf{w}$ . The values of  $\{\mathbf{w}_n\}_{n=1}^N$  are not unique which depend on the control settings discussed in Section 3.3.3. Hence, closing the gap in (3.47) under the assumptions made in this paper can be very challenging.

However, since the primary goal is to train the decision function and the quantization rules across the sensor, this allows us to study the case when the sensor selection strategy

is not used (i.e.,  $M = M$ ). Additionally, the algorithms proposed by [39] and [40] are the same when  $\lambda_2 = 0$  and fixed  $\mathbf{w} = \mathbf{1}$ . Hence, we can study  $R_{\text{avg}}(\mathcal{S})$  and  $R_{\text{reg}}(\mathcal{S})$  defined in [39] when  $\lambda_2 = \lambda_{2,n} = 0$  and  $\mathbf{w} = \mathbf{w}_n$  which indicates that

$$k_{\mathbf{w},n}(\cdot, \mathbf{q}_n) = k_{\mathbf{w}}(\cdot, \mathbf{q}_n), \text{ for } n = 1, 2, \dots, N, \quad (3.58)$$

where  $k_{\mathbf{w},n}(\cdot, \cdot)$  and  $k_{\mathbf{w}}(\cdot, \cdot)$  associate with the OL and the BL setting [39], respectively.

### 3.4.4 Convergence Analysis

Before the convergence bound is studied, some assumptions made in [39] need to be addressed. For the BL approach in [39] (same assumptions made in [40]), the relationship between the true risk function defined in LHS of (3.59) and empirical risk function of the BL approach (the first term on the RHS of (3.11)) have been analyzed in detail in [39], [40] using

$$\begin{aligned} \underset{\gamma_{\text{true}} \in \Gamma}{\text{argmin}} \mathbb{E}_{\mathbf{X}, Y} [\mathbb{E}_{\mathbf{Q}} [\ell(f(\mathbf{Q}), y) | \mathbf{X}, Y]] \\ = \underset{\gamma \in \Gamma_d}{\text{argmin}} \sum_{n=1}^N \frac{\ell(\langle \mathbf{f}, \Phi(\mathbf{x}_n) \rangle, y_n)}{N} + O(N^{-0.5}), \end{aligned} \quad (3.59)$$

where  $\gamma_{\text{true}}$  is the optimal solution of the true risk function assuming the fusion center and all sensors know  $P(\mathbf{X}, Y)$ .  $\gamma \in \Gamma_d$  indicates that

$$P(\mathbf{q}_n | \mathbf{x}_n) = \begin{cases} 1, & \text{if } \mathbf{q}_n = \mathbf{q}, \\ 0, & \text{otherwise,} \end{cases} \quad \text{for some } \mathbf{q} \in \mathcal{Q}. \quad (3.60)$$

**Theorem 3.4.1.** *Given  $\ell(\cdot, \cdot)$  in (3.9) with a properly selected  $\rho$ ,  $\eta_n = \eta_1 n^{-0.5}$ , an i.i.d sequence  $\mathcal{S}$  such that all sensors observe all possible  $x \in \mathcal{X}_m$  at least  $N_x + 1$  times at time  $N$ ,  $\mathbf{w}_n = \mathbf{w}$ , and  $\lambda_{2,n} = \lambda_2 = 0$ , for  $n$  and some fixed  $N < N$ , suppose the initialized quantization rules of the OL approach described in (3.53) and (3.54), and the deterministic quantization rules of the BL approach described in (3.13) follow (3.54), then*

$$R_{\text{avg}}(\mathcal{S}) = R_{\text{reg}}(\mathcal{S}) + O(h(N)), \quad (3.61)$$

such that  $h(N)$  is given by

$$h(N) = \max \left\{ N^{-0.5}, O((1 - \eta)^{N_x}) \right\}, \quad (3.62)$$

where  $\eta$  is defined by Proposition 3.4.1 and a very large  $N$ .

*Proof.* Given the fact that  $\lambda_{2,n} = \lambda_2 = 0$  and  $\mathbf{w}_n = \mathbf{w}$  for all  $n$ , (3.47) can be re-written based on Lemma 1 as

$$R_{\text{avg}}(\mathcal{S}) = \sum_{n=1}^N \frac{cF \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\|}{2N} + R_{\text{reg}}(\mathcal{S}) + O(N^{-0.5}). \quad (3.63)$$

The norm in (3.63) can be written

$$\begin{aligned} & \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\| \\ &= \left\| \sum_{\mathbf{q}_n \in \mathcal{Q}} k_{\mathbf{w},n}(\cdot, \mathbf{q}_n) (P_n(\mathbf{q}_n|\mathbf{x}_n) - P(\mathbf{q}_n|\mathbf{x}_n)) \right\|. \end{aligned} \quad (3.64)$$

where (3.64) follows (3.58). Note that  $P_n(\mathbf{q}_n|\mathbf{x}_n)$  and  $P(\mathbf{q}_n|\mathbf{x}_n)$  are quantization rules across the sensors with respect to the OL and BL approaches, respectively. Hence, we need to obtain the expression of  $P_n(\mathbf{q}_n|\mathbf{x}_n)$  in order to complete the proof, for all possible  $\mathbf{q}_n \in \mathcal{Q}$ ,  $\mathbf{x}_n \in \mathcal{X}$ , and  $n = 1, 2, \dots, N$ .

There exists a time instance  $N < N$  such that each sensor observes all possible  $x$  at least  $N_x$  times. After time  $n$  such that  $n \geq N$ , we can approximate the quantization rules across the sensors based on Proposition 3.4.1 and (3.2) as

$$P_n(\mathbf{q}_n|\mathbf{x}_n) = \begin{cases} 1 - O((1 - \eta)^{N_x}), & \text{if } \mathbf{q}_n = \mathbf{q}, \\ O((1 - \eta)^{N_x}), & \text{otherwise,} \end{cases} \quad (3.65)$$

where a fixed  $\mathbf{q} \in \mathcal{Q}$ . Now, as we consider

$$\begin{aligned} \sum_{n=1}^N \frac{cF \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\|}{2N} &= \sum_{n=1}^{N-1} \frac{cF \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\|}{2N} \\ &\quad + \sum_{n=N}^N \frac{cF \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\|}{2N}, \end{aligned} \quad (3.66)$$

we can approximate

$$\sum_{n=1}^N \frac{cF \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\|}{2N} \leq \frac{NcF\Psi}{2N} = O\left(\frac{1}{N}\right) \quad (3.67)$$

and

$$\sum_{n=N+1}^N \frac{cF \|\Phi_n(\mathbf{x}_n) - \Phi(\mathbf{x}_n)\|}{2N} = O((1 - \eta)^{N_x}) \quad (3.68)$$

because the initialized quantization rules of the OL and the deterministic quantization rules of the BL follow (3.54). Hence, (3.61) follows (3.67) and (3.68).  $\square$

Note that in the case when  $\eta = \eta_1^P = \eta_2^P = \dots = \eta_N^P = 1$ , (3.61) can be written based on Proposition 3.4.1 as

$$R_{\text{avg}}(\mathcal{S}) = R_{\text{reg}}(\mathcal{S}) + O(N^{-0.5}). \quad (3.69)$$

**Remark 1.** *Placing a deterministic quantizer at each sensor indicates that the size of  $|\mathcal{X}|$  is reduced to  $|\mathcal{Q}|$ . For any  $\mathbf{x} \in \mathcal{X}$ , the class label of  $\mathbf{x}$  is always the same as that of  $\mathbf{q}$ , where  $\mathbf{q}$  is the corresponding quantization output of  $\mathbf{x}$ . Thus, we are also interesting in analyzing the classification performance between a  $\log_2|\mathcal{Q}_m|$ -bit deterministic quantizer and the infinite-resolution deterministic quantizer. In this paper, the infinite-resolution deterministic quantizer indicates that the quantization operation can be neglected, i.e.,  $\mathcal{Q}_m = \mathcal{X}_m$ . Each sensor using an infinite-resolution deterministic quantizer also indicates that sensors are equally reliable and Algorithm 2 can be treated as native online regularized-risk minimization algorithm (NORMA) [49]. NORMA described by [49] only trains the weighted of the decision function that can be implemented in the proposed multi-sensor model. Given the fact that the classification performance of NORMA [49, Theorem 4] are approximately equal to the BL setting when the infinite-resolution deterministic quantizers are considered, we use NORMA as the upper bound based line in Section 3.5 assuming  $N$  is very large.*

### 3.5 Numerical Performance Analysis

In this section, we evaluated Algorithm 2 and analytic results based on two types of finite datasets which are 2-class randomly generated dataset and the 3-class Iris dataset introduced

by University of California, Irvine (UCI) repository dataset [109]. For both cases, we set the learning rate for  $\mathbf{f}_n$ ,  $\eta_n = 0.1n^{-.5}$ , learning rate for  $P_n(\cdot|\cdot)$ ,  $\eta_n^P = 0.1$ , learning rate for  $\mathbf{w}_n$ ,  $\eta_n^{\mathbf{w}} = 0.5$ , margin parameter  $\rho = 1$ , and  $\lambda_1 = 0.1$ . We set  $\mathbf{f}_1 = \mathbf{0}$  and  $\mathbf{w}_1 = \mathbf{1}$ . We initialize  $P_n(\cdot|\cdot)$  that follows (3.53) and (3.54). The simulations are done using the weighted count marginalized kernel. Inspired by [110], the quantization points are bounded in  $[-A, A]$  that are drawn from

$$\mathcal{Q}_m = \left\{ \frac{(2\tau + 1 - |\mathcal{Q}_m|)A}{|\mathcal{Q}_m|} \right\}_{\tau=0}^{|\mathcal{Q}_m|-1}. \quad (3.70)$$

Before we analyze the following figures, we need to define labels appeared in each simulation. We define MSOKSQ without updating  $\mathbf{w}_n$  and  $\log_2[|\mathcal{Q}_m|]$ -bit deterministic quantizer MSOKSQ when Algorithm 2 is performed with  $\mathbf{w}_n = \mathbf{1}$  for all  $n$  and any  $\mathcal{Q}_m$ . MSOKSQ without updating  $P_{m,n}(\cdot|\cdot)$  and  $\mathbf{w}_n$  can be considered as NORMA because  $P_{m,n}(\cdot|\cdot)$  and  $\mathbf{w}_n$  are not being updated by Algorithm 2 for all  $n$ . MSOKSQ optimally selects  $M$  sensors means that Algorithm 2 is performed given  $M$  sensors need to be enabled. MSOKSQ randomly selects  $M$  sensors indicates that  $M$  sensors are randomly enabled comparing to MSOKSQ optimally selects  $M$ , which can be done by reassigning the value of  $w_{m,n}$  produced by Algorithm 1 randomly for all  $m$ . NORMA introduced by [49] can be used for the proposed multi-sensor mode, where each sensor forwards its observation directly to the fusion center. It acts as the upper bound baseline when we analyze the proposed system model because each sensor can be considered to equip infinite-resolution deterministic quantizer, i.e.,  $\mathcal{Q}_m = \mathcal{X}_m$ , please refers to Remark 3.

### 3.5.1 Randomly Generated Data for Binary Classification Problem

Now, we consider the first finite dataset, i.e., the two-class randomly generated dataset. It is randomly generated based on  $|\mathcal{X}_m|$ , which is

$$x_m = \begin{cases} y - 0.75((1 + c_1)c_2 + c_2^2), & \text{when } y = 1, \\ y + 0.75((1 + c_1)c_2 + c_2^2), & \text{when } y = -1, \end{cases} \quad (3.71)$$

where  $c_1 \in (0, 1]$  a constant and  $c_2$  is uniformly drawn from  $(0, 1)$  such that  $x_m \in [-1, 1]$ , for  $m = 1, 2, \dots, M$ . Any element in  $\mathcal{X}_m$  belongs to the interval  $[-1, 1]$  such that  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_M$  are not necessarily equivalent to each other. Although there are various ways to generate the data, we consider (3.71) because we can produce separable and non-separable dataset when we properly select  $c_1$ . For example, at sensor  $m$ , Fig. 3.2 shows the construction of the two-class finite sets  $\mathcal{X}_m$  with 100 data points (50 data points/class). Followed by this, Fig. 3.3 shows the performance of (3.71) over  $N$ . In particular,  $\mathcal{X}_m$  with  $c_1 = 0.1$  is separable dataset while  $\mathcal{X}_m$  with  $c_1 = 1$  is non-separable dataset.

Fig. 3.4–Fig. 3.6 are simulations based on the following assumptions. We consider  $N = 600$  time instances. We assume that  $P_{m,n}(q_{m,n}|x_{m,n})$  is initialized based on (3.2) for all possible  $q_{m,n} \in \mathcal{Q}$  and  $x_{m,n} \in \mathcal{X}_m$ . Given the fact any element in  $\mathcal{X}$  belongs to  $[-1, 1]$ , each sensor equips with a 3-bit uniform quantizer, where  $\mathcal{Q}_m$  is defined as (4.7). Followed by this, it is straightforward to generalize  $\mathcal{Q}_m$  when 1-bit and 2-bit uniform quantizers are considered. At time  $n$ , each sensor makes its observations  $x_{m,n}$  that is assumed to be drawn from  $\mathcal{X}_m$  randomly, where  $|\mathcal{X}_m| = 20$  and 10 data points/class. Moreover, the class of  $x_{m,n}$ , e.g.,  $y_n \in \{-1, 1\}$  is assumed to be the same, for  $m = 1, 2, \dots, M$ . The fusion center first classifies the quantization outputs  $\mathbf{q}_n$  forwarded by the sensors using the decision function

$$\langle \mathbf{f}_n, k_{\mathbf{w}_{i,n}}(\cdot, \mathbf{q}_n) \rangle = \sum_{i=1}^{n-1} \alpha_{i,n} \sum_{\mathbf{q}_i \in \mathcal{Q}} P_i(\mathbf{q}_i | \mathbf{x}_i) k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n), \quad (3.72)$$

where  $k_{\mathbf{w}_{i,n}}(\mathbf{q}_i, \mathbf{q}_n)$  follows (3.4). Then, the fusion center updates the desired solution  $\gamma_n$  based on Section 3.3. The accuracy classification rate (ACR) is defined as a classification performance metric, which is

$$\text{ACR} = \sum_{i=1}^N \frac{1}{N} \mathbb{1} \left( \text{sgn} \left( \langle \mathbf{f}_i, k_{\mathbf{w}_{i,i}}(\cdot, \mathbf{q}_i) \rangle \right) = y_i \right). \quad (3.73)$$

In Fig. 3.4, we study how the number of sensors and optimal quantization rules affect the ACR performances. We consider the number of sensor  $m = 1, 5$ , and 10. We assume that the elements in a two-class finite dataset are non-separable, i.e.,  $c_1 = 0.1$  described by (3.71). The ACR for both MSOKSQ without updating  $\mathbf{w}_n$  and MSOKSQ without updating

$P_{m,n}(\cdot, \cdot)$  and  $\mathbf{w}_n$  increase with number of sensors. As the time instance increases, MSOKSQ without updating  $\mathbf{w}_n$  offers superior ACR performance than MSOKSQ without updating  $P_{m,n}(\cdot, \cdot)$  and  $\mathbf{w}_n$ . Moreover, the ACR performances of MSOKSQ without updating  $\mathbf{w}_n$  converge to the ACR performance of NORMA because the randomized quantizers become the deterministic quantizers. As each sensor uses a 1-bit deterministic quantizer, the ACR performance of MSOKSQ overlaps NORMA, refers Remark 2.

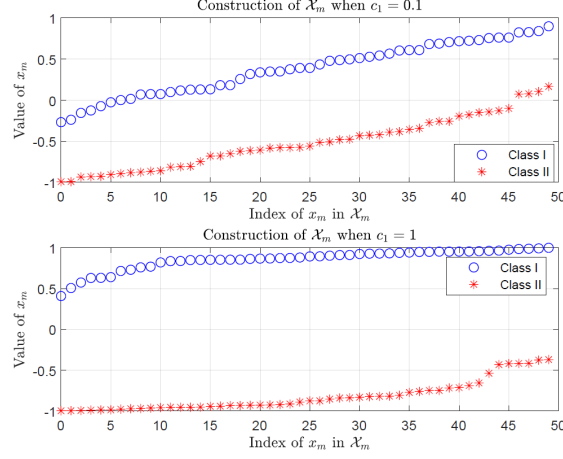
Fig. 3.5 analyzes the ACR performances for both optimal and random sensor selections when the total number of sensor is  $M = 11$ . In particular, we compare three different cases when  $M = 1, 5$ , and 10 number of sensors are enabled given sensors equip 3-bit quantizers. Although the ACR performances of MSOKSQ with optimal/random sensor selection increase with time, MSOKSQ optimal sensor selection out-performs MSOKSQ random sensor selection for  $M = 1, 5$ , and 10. Followed by this, the ACR performances of MSOKSQ optimal/random sensor selections converge to NORMA as the time instance increases for all  $M$ . This is because the randomized quantizers become the deterministic quantizers. Finally, as each sensor uses a 2-bit deterministic quantizer, the ACR performance of MSOKSQ overlaps NORMA, refers Remark 2.

Fig. 3.6 demonstrates the behavior of the quantization rule when  $x_{m,n} = 0.4329$  for all  $m$  and  $n$ . The assumptions of Fig. 3.6 are the same as Fig. 3.5. Given  $\mathcal{Q}_m$  is defined in (4.7), the quantization rule of sensor  $m$  is initialized such that

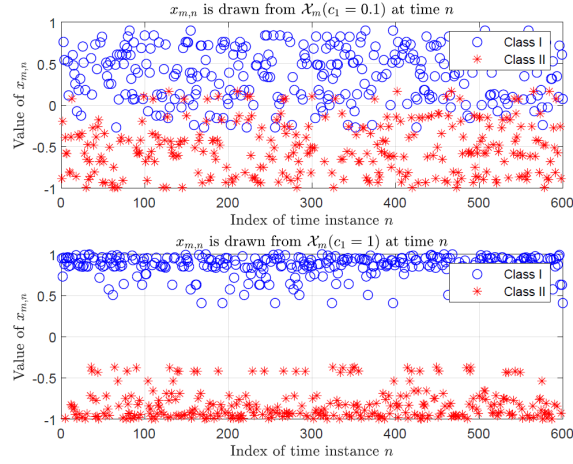
$$\begin{aligned} &P_{m,n_1}(q_{m,n_1} = 0.5 | x_{m,n_1} = 0.4329) \\ &= \max_{q_{m,n_{1_x}} \in \mathcal{Q}_m} P_{m,n_1}(q_{m,n_1} | x_{m,n_1} = 0.4329). \end{aligned} \quad (3.74)$$

Fig. 3.6 indicates that  $P_{m,n_1}(q_{m,n_1} = 0.5 | x_{m,n_1} = 0.4329)$  approaches to one as  $i_x$  increases while maintaining the summation constraint of the quantization rules proposed by (3.2). Note that  $i_x$  is the  $i_x$ th times sensor  $m$  observes 0.4329 when  $\partial_{P_{m,n_{i_x}}(q_{m,n_{i_x}} | x_{m,n_{i_x}} = 0.4329)} R(s_{n_{i_x}}) \neq 0$ . Thus, Fig. 3.6 validates Proposition 3.4.1.





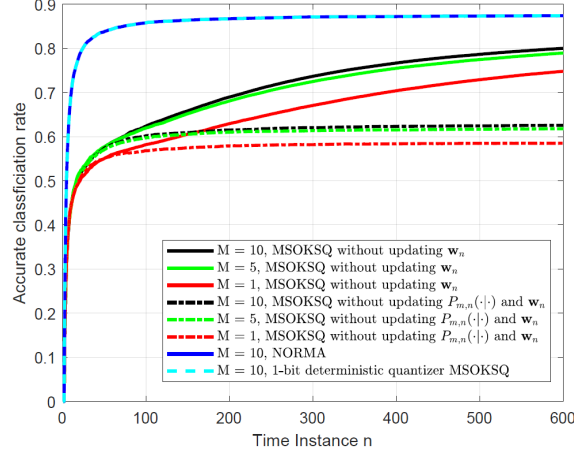
**Figure 3.2.** Construction of  $\mathcal{X}_m$  based on (3.71) with  $c_1 = 0.1$  and  $c_1 = 1$ .



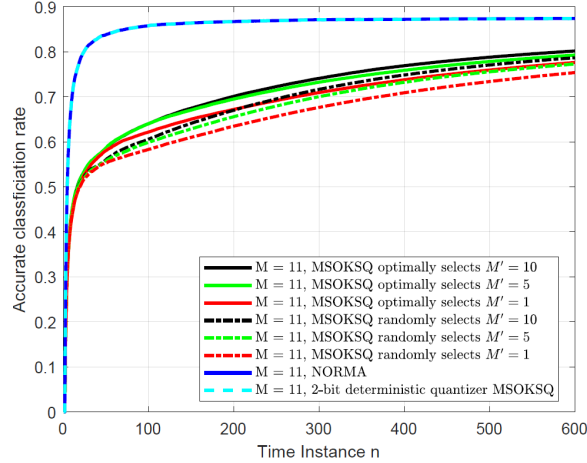
**Figure 3.3.** Demonstrating examples that sensor  $m$  observes data/time drawn from the non-separable ( $c = 0.1$ ) and separable dataset ( $c = 1$ ).

### 3.5.2 UCI Iris Data for Multi-Class Problem

Although this paper focuses on the binary classification problem, Algorithm 2 can also handle a multi-class classification problem. Here, we introduce a second type of the finite dataset which is the 3-class Iris dataset [109] whose labels at time  $n$ ,  $\mathbf{y}_n = [y_{1,n}, y_{2,n}]$  are drawn from  $\mathcal{Y} \in \{[-1, 1], [1, -1], [1, 1]\}$ . Moreover, given the fact each sensor uses a scalar quantizer, we assume that each  $x_m \in \mathcal{X}_m$  has the same class such that  $\mathcal{X}_1 = \mathcal{X}_2 = \dots = \mathcal{X}_M$ , where  $\mathcal{X}_m$  is the finite dataset representing the petal width of the Iris [109], for  $m = 1, 2, \dots, M$ . Followed by the petal width of the Iris, if we assume each sensor equips a

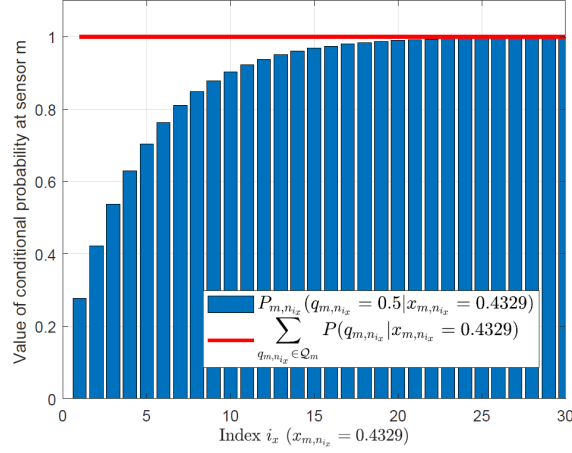


**Figure 3.4.** The ACR performance for Algorithm 2 between MSOKSQ without updating  $\mathbf{w}_n$  and MSOKSQ without updating  $\mathbf{w}_n$  and  $P_n(\cdot|\cdot)$  assuming a 3-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 1-bit deterministic quantizer given  $M = 10$ .



**Figure 3.5.** The ACR performance for Algorithm 2 between MSOKSQ with optimal/random sensor selection when  $M = 1, 5$ , and 10 sensors are enabled among  $M = 11$  and 3-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 2-bit deterministic quantizer given  $M = 11$ .

2-bit quantizer, then, the quantization points can be drawn from  $\{-3, -1, 1, 3\}$  [110]. Over  $N = 600$  time instances, each sensor makes its observations  $x_{m,n}$  that is assumed to be drawn from  $\mathcal{X}_m$  randomly per time instance, for all  $m$ . One way of solving this multi-class problem is to use the One Versus All (OVA) technique which has been extensively discussed



**Figure 3.6.** Demonstrating the value of  $P_{m,n_{i_x}}(q_{m,n_{i_x}} = 0.5 | x_{m,n_{i_x}} = 0.4329)$  for  $i_x = 1, 2, \dots, 30$ , where  $i_x$  is the  $i_x$ th times that sensor  $m$  observes 0.4329 when  $\partial_{P_{m,n_{i_x}}(q_{m,n_{i_x}} | x_{m,n_{i_x}} = 0.4329)} R(s_{n_{i_x}}) \neq 0$ .

$$\text{ACR} = \sum_{i=1}^n \frac{\mathbb{1} \left( \text{sgn} \left( \langle \mathbf{f}_{1,i}, k_{\mathbf{w},i}(\cdot, \mathbf{q}_i) \rangle \right) = y_{1,n}, \text{sgn} \left( \langle \mathbf{f}_{2,i}, k_{\mathbf{w},i}(\cdot, \mathbf{q}_i) \rangle \right) = y_{2,n} \right)}{n}, \quad (3.75)$$

in [106], [111], [112]. As a result, we need two decision functions, where the first and the second decision functions classify the first label,  $y_{1,n}$  and the second label,  $y_{2,n}$ , respectively. This means Algorithm 1 runs twice with respect to each decision function independently. The ACR for the multi-class problem is defined in (3.75), where  $\mathbf{f}_{1,i}$  and  $\mathbf{f}_{2,i}$  are the weights of the decision functions that will be use to determine  $\mathbf{y}_i$ , for  $i = 1, 2, \dots, n$ .

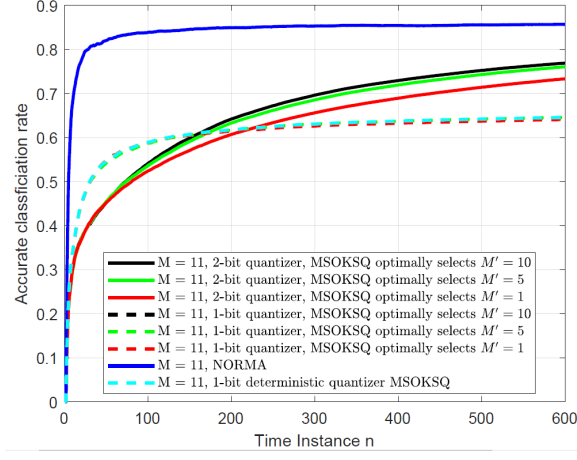
The purpose of providing Fig. 3.7 is to show how the number of quantization bits affects the ACR for the multi-class problem. We assume the total number of sensor  $M = 11$ . We compare the cases when each sensor produces 1-bit and 2-bit outputs assuming  $M = 1, 5$ , and 10 number of sensors are enabled. At the beginning, the ACR performance of the 1-bit quantizer exceeds that of the 2-bit quantizer because the quantization rules of the 1-bit quantizer becomes deterministic faster than the quantization rules of the 2-bit quantizer. However, a 1-bit quantizer maps  $x_{m,n}$  onto one of the two quantization points. Hence, as the time instance increases, Fig. 3.7 shows that the ACR performance of using a 2-bit quantizer/sensor out-performs a 1-bit quantizer/sensor as expected. Intuitively, the system requires at least 2-bit quantizers assuming each sensor uses same number of quantizer to

handle a 3-class problem. Moreover, as each sensor uses a 1-bit deterministic quantizer, the ACR performance of MSOKSQ overlaps MSOKSQ optimal sensor selection given  $M = 5, 10$ , refers Remark 2.

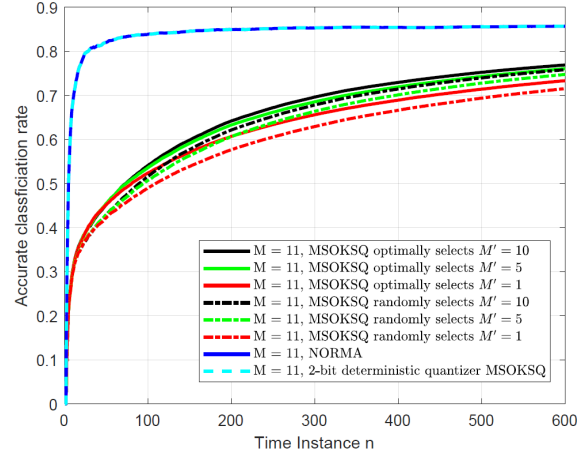
The purpose of Fig. 3.8 is same as Fig. 3.5 which shows the ACR performance of both optimal and random sensor selection when  $M = 11$ . We assume each sensor equips 2-bit quantizer. The assumptions made in Fig. 3.8 shares similar to Fig. 3.7. We compare three different cases when  $M = 1, 5$ , and 10 number of sensors are assumed to be enabled. Fig. 3.8 shows a similar trend to that of Fig. 3.5, as would be expected, please refers to Fig. 3.5. Moreover, as each sensor uses a 2-bit deterministic quantizer, the ACR performance of MSOKSQ overlaps NORMA, refers Remark 2.

The purpose of Fig. 3.9 much like Fig. 3.4 studies how the number of sensors and optimal quantization rules affect the ACR performance defined in (3.75). Similar to Fig. 3.4, we consider the number of sensors  $m = 1, 5$ , and 10. Fig. 3.9 shows a trend similar to that of Fig. 3.4 as expected, please refers the analysis in Fig. 3.4. Finally, as each sensor uses a 3-bit deterministic quantizer, the ACR performance of MSOKSQ overlaps NORMA, refers Remark 2.

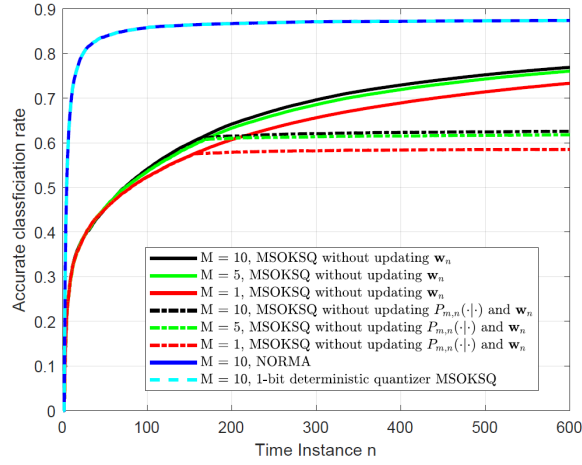
**Remark 2.** *Sensors are equally reliable when the deterministic quantizers are considered. As Fig. 3.4, Fig. 3.5, and Fig. 3.7– Fig. 3.9 show, if we properly select the size of the quantization points of the deterministic quantizers, e.g.,  $|\mathcal{Q}_m|$ , for all  $m$ , the ACR performance of the  $\log_2[|\mathcal{Q}_m|]$ -bit deterministic quantizer converge to NORMA.*



**Figure 3.7.** The ACR performance for Algorithm 2 between MSOKSQ with optimal and random sensor selection.  $M = 1, 5$ , and 10 sensors are assumed to be enabled and a 2-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 1-bit deterministic quantizer given  $M = 11$ .



**Figure 3.8.** The ACR performance for Algorithm 2 between MSOKSQ without updating  $\mathbf{w}_n$  and MSOKSQ without updating  $\mathbf{w}_n$  and  $P_n(\cdot|\cdot)$ .  $M = 1, 5$ , and 10 sensors are assumed to be enabled and a 2-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 2-bit deterministic quantizer when  $M = 11$ .



**Figure 3.9.** The ACR performance for Algorithm 2 between MSOKSQ with optimal and random sensor selection.  $M = 1, 5$ , and 10 sensors are assumed to be enabled and a 2-bit quantizer is used for each sensor. The ACR performance for between NORMA and MSOKSQ when each sensor uses a 3-bit deterministic quantizer when  $M = 11$ .

## 4. A FOLDING APPROACH FOR MULTIPLE ANTENNA ARRAYS USING LOW-RESOLUTION ADCS

### 4.1 System Model and Low-Resolution ADC Array Architecture

In this section, we explain our general system model. Then, we motivate the framework of constructing a high-resolution ADC by using multiple antennas with each antenna served by a low-resolution ADC.

#### 4.1.1 System Model

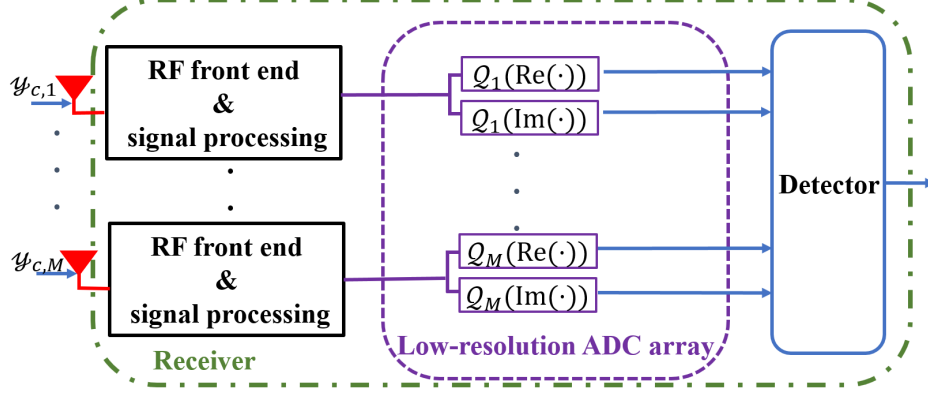
We consider an  $N$  transmit antenna by  $M$  receive antenna wireless MIMO system. We assume perfect synchronization **pan1**, [99] and a frequency flat, narrow-band channel with a complex baseband representation at the receiver of

$$\mathbf{y}_c = \mathbf{H}\mathbf{f}_c s_c + \mathbf{n}_c \quad (4.1)$$

$$= \mathbf{h}s_c + \mathbf{n}_c, \quad (4.2)$$

where  $\mathbf{y}_c = [y_{c,1} \cdots y_{c,M}]^T$ ,  $\mathbf{H} \in \mathbb{C}^{M \times N}$  is the channel matrix,  $\mathbf{f}_c \in \mathbb{C}^{N \times 1}$  is the transmit beamforming vector,  $\mathbf{h} = [h_1 \cdots h_M]^T$  is the effective channel vector, and  $\mathbf{n}_c \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$  is the complex additive (spatially and temporally) independent and identical white Gaussian noise. The transmit symbol  $s_c$  is assumed to be uniformly selected from a finite set  $\mathcal{S}_c \in \mathbb{C}$  which is designed to satisfy  $\mathbb{E}[|s_c|^2] = P$ . Since channels can be estimated using [69]–[74], we assume that the transmitter and the receiver have any necessary knowledge of the effective channels. Note that (4.2) means that the MIMO system model can be translated into an effective single-input multiple-output (SIMO) problem.

In our paper, a low-resolution ADC array architecture is considered, which is shown in Fig. 4.1. At antenna  $m$ , we assume phase shifting dependent on the phase of the effective



**Figure 4.1.** Implementation of low-resolution ADC array at the receiver.

channel  $h_m$ , which can be done, for example, using an analog per-antenna phase alignment [113]–[117]. Mathematically, the processed baseband signal at antenna  $m$  is

$$\tilde{y}_{c,m} = \frac{h_m^* y_{c,m}}{|h_m|} = |h_m| s_c + n_{c,m}, \quad (4.3)$$

where  $|h_m|$  is the magnitude of the effective channel and we redefine the noise as  $n_{c,m}$  since phase rotation does not change the distribution. Then, the processed real and imaginary parts of the receive signal are fed to low-resolution ADCs to be quantized separately which are

$$\mathcal{Q}(\text{Re}(\tilde{\mathbf{y}}_c)) = [\mathcal{Q}_1(\text{Re}(\tilde{y}_{c,1})) \cdots \mathcal{Q}_M(\text{Re}(\tilde{y}_{c,M}))]^T, \quad (4.4)$$

$$\mathcal{Q}(\text{Im}(\tilde{\mathbf{y}}_c)) = [\mathcal{Q}_1(\text{Im}(\tilde{y}_{c,1})) \cdots \mathcal{Q}_M(\text{Im}(\tilde{y}_{c,M}))]^T, \quad (4.5)$$

to the detector. Note that  $\tilde{\mathbf{y}}_c = [\tilde{y}_{c,1} \cdots \tilde{y}_{c,M}]^T$ .

#### 4.1.2 Low-Resolution ADC Array Construction by Deconstructing a High-Resolution ADC

A low-resolution ADC array uses a small number of bits received from each of a (potentially large) number of antennas. Therefore, across the entire array, the number of bits



per sample may be quite large. The goal then is to design the analog processing and low-resolution ADC used at the different elements to be complementary when the array is considered as a whole.

There has been much research over the last few decades on centralized ADCs (i.e., a single ADC that is designed for general applications as in [67], [68]). Many of the highest-performance centralized ADCs can be implemented using a parallel structure [67], [68]. In this structure, the ADC circuit consists of multiple branches where each branch produces a single bit per sample, e.g., **pan1**, [96], [118]. Note that there is a strong analogy between this kind of structure and the architecture of low-resolution ADC arrays. In this work, we will show how ideas gained from the architecture of a specific kind of high-performance ADC, called an FADC, can be used to develop a new form of low-resolution ADC array.

Before providing a description of our proposed low-resolution ADC array, it is critical to recall the basics of ADCs in general. For this reason, the discussion in this subsection will start with a focus on the case of a *single or centralized* high-resolution ADC that produces  $B$ -bits per sample. Designing an ADC can be viewed as designing a quantizer that maps a real input to a quantization point from a finite set  $\mathcal{A} = \{a_0, a_1, \dots, a_{2^B-1}\}$ . Thus, each quantization point  $a_\ell$  in the set  $\mathcal{A}$  can be represented by a  $B$ -bit binary vector (or bit label) denoted by  $\mathbf{b}_\ell$ , where  $0 \leq \ell \leq 2^B - 1$ . The set of all possible binary outputs for the quantizer are in the set  $\mathcal{B} = \{\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{2^B-1}\}$ . We describe the quantization function of an ADC as  $\mathcal{Q}_{\text{ADC}}(\cdot) : \mathbb{R} \rightarrow \{-1, +1\}^B$ .

The simplest quantization scheme is uniform quantization [110] which we focus on throughout this paper. In uniform quantizers, the quantization points in the set  $\mathcal{A}$  are equally spaced in a bounded interval  $[-A, A]$  where  $A > 0$ . The minimum distance between quantization points is

$$d_{\min} \triangleq |a_{\ell+1} - a_\ell| = \frac{A}{2^{B-1}}, \text{ for } 0 \leq \ell \leq 2^B - 2. \quad (4.6)$$

The quantization points of the uniform quantizer are

$$a_\ell = \frac{(2\ell + 1 - 2^B)A}{2^B}, \text{ for } 0 \leq \ell \leq 2^B - 1. \quad (4.7)$$

The mapping of bit labels to quantization points can be done in a variety of ways. Natural labeling, which is used as the bit labeling throughout much of this paper, uses the base-two labels  $\mathbf{b}_\ell$  on  $a_\ell$ . The explicit expression for the mapping between  $a_\ell$  and  $\mathbf{b}_\ell$  is written

$$a_\ell = \frac{A}{2^B} \sum_{i=1}^B b_{\ell,i} 2^{B-i}, \quad (4.8)$$

where  $b_{\ell,i} \in \{-1, +1\}$  is the  $i^{\text{th}}$  element of  $\mathbf{b}_\ell$ . Eqn. (4.8) is obtained by first establishing the connection between  $2\ell + 1 - 2^B$  in (4.7) and its binary natural label  $\hat{\mathbf{b}}_\ell \in \{0, 1\}^B$  based on  $\beta$ -expansion [119]–[121]. Hence, it is straightforward to show

$$2\ell = \sum_{i=1}^B \hat{b}_{\ell,i} 2^{B-i+1}, \quad (4.9)$$

where  $\hat{b}_{\ell,i} \in \{0, 1\}$  is the  $i^{\text{th}}$  element of  $\hat{\mathbf{b}}_\ell$  and

$$2^B - 1 = \sum_{i=1}^B 2^{B-i}. \quad (4.10)$$

Then, as we apply (4.9) and (4.10) into the right hand side of (4.7), we have

$$a_\ell = \frac{A}{2^B} \sum_{i=1}^B (2\hat{b}_{\ell,i} - 1) 2^{B-i} = \frac{A}{2^B} \sum_{i=1}^B b_{\ell,i} 2^{B-i}, \quad (4.11)$$

where  $b_{\ell,i} = 2\hat{b}_{\ell,i} - 1 \in \{-1, +1\}$ . For example, if we use a two-bit uniform quantizer, which corresponds to  $B = 2$  and  $L = 2^B$ , with quantization points uniformly distributed between  $-A$  and  $A$ , then we have the quantization points  $a_0 = -3A/4$ ,  $a_1 = -A/4$ ,  $a_2 = A/4$ ,  $a_3 = 3A/4$  based on (4.7). A natural labeling of the quantization points is  $\mathbf{b}_0 = [-1 \ -1]$ ,  $\mathbf{b}_1 = [-1 \ 1]$ ,  $\mathbf{b}_2 = [1 \ -1]$ ,  $\mathbf{b}_3 = [1 \ 1]$  so that  $\mathbf{b}_\ell$  can recover  $a_\ell$  above using (4.8). Note that the quantization points can be labeled with bits in a variety of ways.

Many high-performance centralized high-resolution ADCs use a parallel implementation e.g., **pan1**, [67], [68], [96]. This parallel form means that the quantization function of the ADC can be written as  $\mathcal{Q}_{\text{ADC}}(\alpha) = [\mathcal{Q}_{\text{ADC},1}(\alpha) \cdots \mathcal{Q}_{\text{ADC},B}(\alpha)]^T \in \{-1, +1\}^B$ , where  $\mathcal{Q}_{\text{ADC},i}(\cdot) \in \{-1, +1\}$  is a one-bit quantizer for the  $i^{\text{th}}$  branch. Using this kind of parallel ADC thinking,

we move on to provide a detailed explanation of our folded low-resolution ADC array in Section 4.2.

## 4.2 Folded Low-Resolution ADC Array Design and Implementation for Multiple Antenna Communication

In this section, we first explain the design of a low-resolution ADC array based on the multi-branch architecture, with a specific focus on the FADC architecture. Then, we derive the desired parameters needed to facilitate folded low-resolution ADC array operation even when there is little to no spatial channel diversity. Finally, we discuss how our implementation affects the output statistically.

### 4.2.1 Basics of Folding

A folded low-resolution ADC array employs analog processing and a simple sign quantizer at each array element. To ease our exposition, we will concentrate on an explanation based on concepts from QAM. This QAM assumption alone differentiates our approach from the sign-based low-resolution ADC array, because our formulation can easily handle a constellation with symbols of varying magnitudes.

Further recall that any rectangular QAM constellation can be written in terms of two independently modulated pulse amplitude modulation (PAM) constellations. To simplify the explanation, we assume a square QAM constellation is used. This allows us to design the in-phase and quadrature analog processing at each of our array elements in the same way. The square QAM constellation is assumed to be formed of symbols  $\mathcal{S}_c = \{\text{Re}(s_c) + j\text{Im}(s_c) : \text{Re}(s_c), \text{Im}(s_c) \in \mathcal{S}\}$ , where  $\text{Card}(\mathcal{S}) = 2^M$ . Therefore, we explain the design of the folded low-resolution ADC array operating on a real vector  $\mathbf{y} = [y_1 \cdots y_M]^T$  throughout the paper, where  $\mathbf{y} \in \{\text{Re}(\tilde{\mathbf{y}}_c), \text{Im}(\tilde{\mathbf{y}}_c)\}$ .

The quantization function of the folded low-resolution ADC array in real space in (4.4) and (4.5) is denoted as  $\mathcal{Q}_{\text{FADC}}(\mathbf{y}) = [\mathcal{Q}_1(y_1) \cdots \mathcal{Q}_M(y_M)]^T$ , where  $\mathcal{Q}_{\text{FADC}}(\cdot) \in \{-1, +1\}^M$  and the one-bit quantizer  $\mathcal{Q}_m(\cdot) \in \{-1, +1\}$ . This means that the multiple one-bit quantizers spread across the array produce  $M$  bits per sample for the in-phase portion of the

signal and  $M$  bits per sample for the quadrature portion of the signal. As mentioned earlier, the symmetry of a square QAM constellation allows us to focus only on the in-phase portion without loss of generality. This indicates that we can think of the *entire* array as a high-resolution quantizer, with  $B = M$ , that is composed of  $M$  one-bit quantizer branches corresponding to each antenna element in the array. Therefore, the goal is to formulate the  $M$  different one-bit quantizers  $\mathcal{Q}_m(\cdot)$ , for all  $m$ , such that they are complementary to each other and provide a form of quantization diversity that will allow the receiver to fuse the quantization outputs together.

The proposed real one-bit quantizer  $\mathcal{Q}_m(\cdot)$  used at each array element consists of analog processing followed by a sign quantizer. The analog processing used at each antenna element contains a *folding* stage, which consists of a carefully chosen additive offset (or shift) and a modulo. Thus, in this thesis, we name the proposed real one-bit quantizer as a *one-bit modulo quantizer*. We discuss the analog processing at the more abstract signal level, but it should be noted that there is a long and well-developed history of shift and modulo circuit implementations [87]. Mathematically,  $y_m$  is folded at *the  $m^{\text{th}}$  array element located at antenna  $m$*  using the function  $F_m(\cdot) \in [-\beta_m, \Lambda_m - \beta_m)$ , which we refer to as a folding function, described by

$$F_m(y_m) = ([y_m + \gamma_m] \bmod \Lambda_m) - \beta_m, \quad (4.12)$$

where  $\Lambda_m > 0$  is the modulo size, and  $\gamma_m$  and  $\beta_m$  are real-valued constants.

Each one-bit modulo quantizer is designed to perform uniform quantization, which will mean that the sizes of  $[-\beta_m, 0)$  and  $[0, \Lambda_m - \beta_m)$  are equal, i.e.,  $\beta_m = \Lambda_m - \beta_m$ . Therefore, the relationship between  $\beta_m$  and  $\Lambda_m$  is

$$\beta_m = \frac{\Lambda_m}{2}. \quad (4.13)$$

As a result, we rewrite the folding function as

$$F_m(y_m) = [y_m + \gamma_m] \bmod \Lambda_m - \frac{\Lambda_m}{2} \quad (4.14)$$

throughout the remainder of the paper. After folding,  $F_m(y_m)$  is quantized by a sign quantizer so that  $\mathcal{Q}_m(\cdot)$  is described by

$$\mathcal{Q}_m(y_m) = \text{sgn}(F_m(y_m)). \quad (4.15)$$

As mentioned earlier, our goal is to select the low-resolution ADC, e.g.,  $\mathcal{Q}_m(\cdot)$  for all  $m$ , at the array elements in a complementary way. This will mean that the low-resolution quantization across the array has a kind-of diversity and that the folding functions must be jointly designed. The goal is thus to derive  $\Lambda_m$  and  $\gamma_m$  for all  $m$  in (4.14). As will be shown, our joint design has a close tie to the bit-to-quantization point mapping described in (4.8) in the following subsection.

#### 4.2.2 Design of Folded Low-Resolution ADC Array

We define the transmit symbol in the real space as  $s \in \{\text{Re}(s_c), \text{Im}(s_c)\}$ , which is uniformly drawn from a finite set  $\mathcal{S}$  such that

$$\mathcal{S} = \left\{ \left(2\ell + 1 - 2^M\right) \frac{d}{2} : 0 \leq \ell \leq 2^M - 1 \right\}, \quad (4.16)$$

where  $d$  denotes the minimum distance between any of the adjacent QAM constellation points. Each symbol  $s_\ell \in \mathcal{S}$  can be written using the corresponding natural labels  $\mathbf{b}_\ell$  according to

$$s_\ell = \frac{d}{2} \sum_{m=1}^M b_{\ell,m} 2^{M-m} \quad (4.17)$$

by exploiting (4.6)-(4.8) when  $B = M$  is assumed.

To find the desired parameters  $\gamma_m$  and  $\Lambda_m$  of  $F_m(\cdot)$  for all  $m$ , we first derive a closed-form solution of  $\Lambda_m$  using Proposition 1. Then, we use  $\Lambda_m$  and  $F_m(\cdot)$  to derive  $\gamma_m$ , for all  $m$ .

**Proposition 4.2.1.** *Assume  $|h_1|, \dots, |h_M|$  are arbitrary positive reals. A symbol  $s_\ell \in \mathcal{S}$  that is defined by (4.17) can be described according to*

$$s_\ell = \sum_{m=1}^M b_{\ell,m} \frac{\Lambda_m}{4|h_m|}, \quad (4.18)$$

where  $b_{\ell,m}$  is the  $m^{\text{th}}$  element of  $\mathbf{b}_\ell$  and  $\frac{\Lambda_1}{|h_1|} = \frac{2\Lambda_2}{|h_2|} = \frac{2^2\Lambda_{M-1}}{|h_{M-1}|} = \dots = \frac{2^{M-1}\Lambda_M}{|h_M|}$  if and only if

$$\Lambda_m = 2^{M-m+1}d|h_m|, \text{ for } 1 \leq m \leq M. \quad (4.19)$$

*Proof.* We first assume that (4.18) holds. Applying  $\frac{\Lambda_1}{|h_1|} = \frac{2\Lambda_2}{|h_2|} = \dots = \frac{2^{M-1}\Lambda_M}{|h_M|}$  into (4.18), we can reconstruct (4.18) as

$$s_\ell = \sum_{m=1}^M b_{\ell,m} 2^{M-m} \frac{\Lambda_M}{4|h_M|}. \quad (4.20)$$

Since both of (4.17) and (4.20) are two representations of the quantization point  $s_\ell$ , the relationship between the right hand sides of both equations are shown as

$$\frac{d}{2} \sum_{m=1}^M b_{\ell,m} 2^{M-m} = \sum_{m=1}^M b_{\ell,m} 2^{M-m} \frac{\Lambda_M}{4|h_M|}.$$

Comparing both sides of coefficients of  $b_{\ell,i}$  in the described relationship, (4.19) holds.

To prove the converse part of this lemma, we assume that (4.19) holds. Using the fact that  $\mathbf{b}_\ell$  follows natural labeling and substituting (4.19) into (4.17), the result leads directly to (4.18).  $\square$

Once we obtain  $\Lambda_m$ , the mapping of the natural bit labels to the constellations is accomplished when we set

$$\gamma_m = \begin{cases} \frac{\Lambda_m}{2}, & \text{when } m = 1, \\ 0, & \text{when } m \neq 1, \end{cases} \quad (4.21)$$

where the values of  $\gamma_m$  are not unique as long as the output of the folded low-resolution ADC array follows natural labeling. Note that the  $m = 1$  offset  $\gamma_1$  is different from the other values. This follows because the first bit represents the sign of the quantization point for the portion of the real line that the transmit signal is constrained to in the natural labeling mapping as shown in (4.18).

Therefore, the one-bit modulo quantization process can be summarized as follows. For  $1 \leq m \leq M$ ,  $y_m$  is folded into  $\left[-\frac{\Lambda_m}{2}, 0\right)$  when  $y_m \in \bigcup_{k \in \mathbb{Z}} \left[\left(k - \frac{1}{2}\right) \Lambda_m - \Delta_m, k \Lambda_m - \Delta_m\right)$ , where

$$\Delta_m = \gamma_m - \frac{\Lambda_m}{2}. \quad (4.22)$$

The corresponding quantization output of the one-bit modulo quantizer  $b_m = -1$ . Similarly, the input  $y_m$  is folded into  $\left[0, \frac{\Lambda_m}{2}\right)$  when  $y_m \in \bigcup_{k \in \mathbb{Z}} \left[k \Lambda_m - \Delta_m, \left(k + \frac{1}{2}\right) \Lambda_m - \Delta_m\right)$ . The corresponding quantization output  $b_m = 1$ . Thus, we denote  $\mathcal{R}_{m,1}$  and  $\mathcal{R}_{m,2}$  as the folding quantization regions such that

$$\mathcal{R}_{m,1} = \bigcup_{k \in \mathbb{Z}} \left[\left(k - \frac{1}{2}\right) \Lambda_m - \Delta_m, k \Lambda_m - \Delta_m\right), \quad (4.23)$$

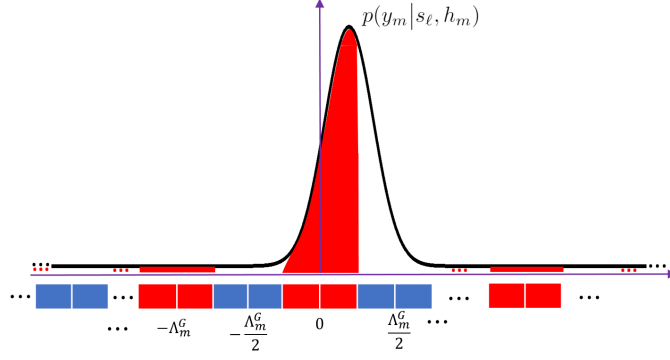
$$\mathcal{R}_{m,2} = \bigcup_{k \in \mathbb{Z}} \left[k \Lambda_m - \Delta_m, \left(k + \frac{1}{2}\right) \Lambda_m - \Delta_m\right). \quad (4.24)$$

The size of each interval that is obtained by  $\mathcal{R}_{m,1}$  or  $\mathcal{R}_{m,2}$  is equal to  $\frac{\Lambda_m}{2}$ . In addition, (4.23) and (4.24) indicate  $\mathcal{R}_{m,1} \cap \mathcal{R}_{m,2} = \emptyset$  and  $\mathcal{R}_{m,1} \cup \mathcal{R}_{m,2} = \mathbb{R}$ .

**Remark 3.** *The analog shift and modulo operations proposed in this thesis shape the magnitude of the observation of each antenna in a particular way before the sign-based quantization operation is performed at each array element. Thus, assumptions made by the sign-based low-resolution ADC array are compatible with the proposed designs. In addition, our description of folding is based on a square constellation. However, it can be used for other constellations, e.g., circular QAM or PSK constellations, by embedding the non-grid constellation into the QAM grid. This has the benefit of potentially generating soft information for use in error control decoding schemes. This embedding is beyond the scope of this paper.*

### 4.2.3 Statistical Analysis of Folded Low-Resolution ADC Array in Real Space

Having characterized the implementation of the folded low-resolution ADC array, we discuss how the one-bit modulo quantizer affects the output by studying the conditional probability in real space. For given  $s \in \mathcal{S}$  and  $h_m$ ,  $y_m$  is distributed as  $\mathcal{N}\left(|h_m|s, \frac{1}{2}\right)$ . This



**Figure 4.2.** Showing the red regions must be integrated over to compute the received bit probabilities, e.g.,  $\Pr(b_1 = 1 | s, h_1)$ .

can be combined with the spatial independence of the noise to compute the conditional probability

$$\Pr(\mathbf{b}|s, \mathbf{h}) = \prod_{m=1}^M \Pr(b_m | s, h_m). \quad (4.25)$$

This probability will be computed using  $p(y_m | s, h_m) = \frac{1}{\sqrt{\pi}} e^{-\left(y_m - |h_m|s\right)^2}$ .

For any given  $s$  and  $h_m$ , Section 4.2.1 indicates that the outputs  $b_m = -1$  when  $y_m \in \mathcal{R}_{m,1}$  and  $b_m = +1$  when  $y_m \in \mathcal{R}_{m,2}$ . Thus, the conditional probabilities are written as  $\Pr(b_m = -1 | s, h_m) = \int_{\mathcal{R}_{m,1}} p(y_m | s, h_m) dy_m$  and  $\Pr(b_m = +1 | s, h_m) = \int_{\mathcal{R}_{m,2}} p(y_m | s, h_m) dy_m$ . The folded quantization regions  $\mathcal{R}_{m,1}$  and  $\mathcal{R}_{m,2}$  that are shown in (4.23) and (4.24) are the union of an infinite number of non-overlapping intervals. Therefore, we define the conditional probabilities of the one-bit modulo quantizer at antenna  $m$  as

$$\Pr(b_m = -1 | s, h_m) = \sum_{k \in \mathbb{Z}} \int_{(k-\frac{1}{2})\Lambda_m - \Delta_m}^{k\Lambda_m - \Delta_m} p(y_m | s, h_m) dy_m, \quad (4.26)$$

$$\Pr(b_m = +1 | s, h_m) = \sum_{k \in \mathbb{Z}} \int_{k\Lambda_m - \Delta_m}^{(k+\frac{1}{2})\Lambda_m - \Delta_m} p(y_m | s, h_m) dy_m, \quad (4.27)$$

for all  $m$ . Note that  $\Pr(b_m | s, h_m)$  is obtained by summing the area of  $p(y_m | s, h_m)$  under  $\mathcal{R}_{m,1}$  or  $\mathcal{R}_{m,2}$ . Fig. 4.2 demonstrates the relationship between the areas corresponding to each bit value. In this figure's example, the red regions would be integrated over to compute the value of  $\Pr(b_1 = +1 | s, h_1)$ .



### 4.3 Folded Low-Resolution ADC Array Enhancements

In this section, we are interested in enhancing the achievable rate performance by making the quantization outputs more reliable. Thus, we first explain the implementation of the folded low-resolution ADC array using Gray labeling and discuss the reliability of the outputs. Then, we explore using the sign quantizer to modify the folded low-resolution ADC array into the sign left-most bit (LMB) folded low-resolution ADC array. Finally, we discuss antenna ordering strategies in the low and the moderate-to-high SNR regimes. Note that we analyze the modified design in the real space similar to Section 4.2.

#### 4.3.1 Folded Low-Resolution ADC Array using Gray Labeling

Gray labeling is extensively studied and widely used in systems **Agrell**, [122]–[124], [126], including in practical bit-to-constellation point mappings. We use the superscript  $G$  to designate Gray labeling. An  $M$ -bit binary vector follows Gray labeling when only one entry of  $\mathbf{b}_\ell^G$  and  $\mathbf{b}_{\ell+1}^G$  are different for  $\ell = 0, \dots, 2^M - 2$ . There are various ways to represent Gray labeling, and we focus on the Gray labeling that is obtained by a specific remapping of natural labeling [126]. This corresponds to the conversion

$$b_{\ell,m}^G = \begin{cases} b_{\ell,m}, & \text{when } m = 1, \\ -b_{\ell,m}b_{\ell,m-1}, & \text{when } m \neq 1. \end{cases} \quad (4.28)$$

The intuitive behavior of the folded low-resolution ADC array with Gray labeling is very similar to the development in Section 4.2. When using Gray labeling, each array element still uses a folding function followed by a sign quantizer. The folding function from array antenna element  $m$  is now  $F_m^G(y_m) = \left( [y_m + \gamma_m^G] \bmod \Lambda_m^G \right) - \frac{\Lambda_m^G}{2}$ , and the low-resolution ADC output from array antenna element  $m$  is  $Q_m^G(y_m) = \text{sgn}(F_m^G(y_m))$ .

The parameters  $\Lambda_m^G$  and  $\gamma_m^G$  can be found based on the relationship that is shown in (4.28). The modulo size at receive antenna  $m$  is

$$\Lambda_m^G = \begin{cases} \Lambda_m, & m = 1, \\ 2\Lambda_m, & m \neq 1, \end{cases} \quad (4.29)$$

and

$$\gamma_m^G = \begin{cases} \frac{\Lambda_m^G}{2}, & \text{when } m = 1, \\ -\frac{\Lambda_m^G}{4}, & \text{when } m = 2, \\ \frac{\Lambda_m^G}{4}, & \text{otherwise.} \end{cases} \quad (4.30)$$

Note that depending on how Gray labeling is defined, the construction of the folded low-resolution ADC array using Gray labeling is not unique.

The folding quantization regions  $\mathcal{R}_{m,1}^G$  and  $\mathcal{R}_{m,2}^G$  can be represented in a similar way as (4.23) and (4.24), respectively, by replacing the parameters  $\gamma_m$  and  $\Lambda_m$  with  $\gamma_m^G$  and  $\Lambda_m^G$ . Moreover, the size of each interval obtained from  $\mathcal{R}_{m,1}^G$  and  $\mathcal{R}_{m,2}^G$  are equal to  $\frac{\Lambda_m^G}{2}$ . Thus, the expression of  $\Pr(b_m^G|s, h_m)$  is obtained using (4.26) and (4.27) by replacing  $\mathcal{R}_{m,1}$  and  $\mathcal{R}_{m,2}$  with  $\mathcal{R}_{m,1}^G$  and  $\mathcal{R}_{m,2}^G$ .

In the moderate-to-high SNR regime, the quantization output of antenna  $m$  of a folded low-resolution ADC array using Gray labeling has a reliability that exceeds that of natural labeling, for  $1 < m \leq M$ . We can make such a determination by comparing the value of  $\Pr(b_m^G = b_{\ell,m}^G|s = s_\ell, h_m)$  and  $\Pr(b_m = b_{\ell,m}|s = s_\ell, h_m)$ . We denote  $\Pr(b_m^G = b_{\ell,m}^G|s = s_\ell, h_m)$  (or  $\Pr(b_m = b_{\ell,m}|s = s_\ell, h_m)$ ) as the conditional probability of correct reception. Note that the values of  $\Pr(b_m = b_{\ell,m}|s = s_\ell, h_m)$  and  $\Pr(b_m^G = b_{\ell,m}^G|s = s_\ell, h_m)$  are dominated by the area of  $p(y_m|s = s_\ell, h_m)$  under the interval centered about the mean in the moderate-to-high SNR regime. As the size of the interval with respect to Gray labeling is twice as large as natural labeling as shown in (4.29), it is straightforward to show  $\Pr(b_m^G = b_{\ell,m}^G|s = s_\ell, h_m) > \Pr(b_m = b_{\ell,m}|s = s_\ell, h_m)$  which indicates that the quantization output of antenna  $m$  of Gray labeling is more reliable than natural labeling, for  $1 < m \leq M$ .

### 4.3.2 Sign LMB Folded Low-Resolution ADC Array

The analog modulo operation *folds* the input, even when the bit that is detected is the sign bit, i.e.,  $b_1$ . Thus, we can modify the structure of the folded low-resolution ADC array by replacing the one-bit modulo quantizer of antenna 1 with a sign quantizer, e.g.,  $\mathcal{Q}_1(y_1) = \text{sgn}(y_1)$  while the one-bit modulo quantizer at antenna  $m$  is the same as (4.15) for all  $m \neq 1$ . We refer to this as the sign left-most bit (LMB) folded low-resolution ADC array. By doing so, the quantization output of antenna 1 is less likely to be in error which is shown in the following lemma. We define  $b_1^{\text{sgn}}$  as the output of the sign quantizer to distinguish the output of the one-bit modulo quantizer  $b_1$ .

**Lemma 2.** *For any  $s \in \mathcal{S}$  and  $h_1$ , the conditional probability of correct reception satisfies*

$$\Pr(b_1^{\text{sgn}} = b_{\ell,1} | s = s_\ell, h_1) - \Pr(b_1 = b_{\ell,1} | s = s_\ell, h_1) > 0. \quad (4.31)$$

*Proof.* We start the proof assuming  $|h_1|s_\ell \in [0, \frac{\Lambda_1}{2})$  for some  $\ell$ . Therefore, the correct quantization output of antenna 1 is  $+1$ , i.e.,  $b_1 = +1$ . We show

$$\begin{aligned} & \Pr(b_1^{\text{sgn}} = +1 | s = s_\ell, h_1) - \Pr(b_1 = +1 | s = s_\ell, h_1) \\ &= \int_0^\infty p(y_1 | s = s_\ell, h_1) dy_1 - \sum_{k \in \mathbb{Z}} \left( \int_{k\Lambda_1}^{(k+\frac{1}{2})\Lambda_1} p(y_1 | s = s_\ell, h_1) dy_1 \right) \\ &= \sum_{k \in \mathbb{Z}^+} \left( \int_{(k-\frac{1}{2})\Lambda_1}^{k\Lambda_1} p(y_1 | s = s_\ell, h_1) dy_1 - \int_{-k\Lambda_1}^{(-k+\frac{1}{2})\Lambda_1} p(y_1 | s = s_\ell, h_1) dy_1 \right) \\ &\stackrel{(a)}{=} \sum_{k \in \mathbb{Z}^+} \left( \int_{(k-\frac{1}{2})\Lambda_1}^{k\Lambda_1} p(y_1 | s = s_\ell, h_1) dy_1 - \int_{2|h_1|s_\ell + (k-\frac{1}{2})\Lambda_1}^{2|h_1|s_\ell + k\Lambda_1} p(y_1 | s = s_\ell, h_1) dy_1 \right) \\ &\stackrel{(b)}{>} 0, \end{aligned} \quad (4.32)$$

where (4.32a) follows the fact that  $p(y_1 | s = s_\ell, h_1)$  is the Gaussian probability that is symmetric about the mean  $|h_1|s_\ell$  and (4.32b) holds because  $p(y_1 | s = s_\ell, h_1)$  is decreasing for  $y > |h_1|s_\ell$ . The proof assuming  $|h_1|s_\ell \in (-\frac{\Lambda_1}{2}, 0]$  is similar for some  $\ell$ .  $\square$

### 4.3.3 Antenna Ordering Strategies

In designing the folding functions so far, we have implicitly assumed that the folding function  $F_m(\cdot)$  corresponds to antenna  $m$ . We can adapt the analog processing as a function of the effective channel to make the output more reliable because any real system will have antenna gain variation across the antennas. Therefore, we relabel the antennas as  $i_1, \dots, i_M$  such that  $|h_{i_1}| \geq |h_{i_2}| \geq \dots \geq |h_{i_M}|$ .

#### Antenna ordering in the low SNR regime

All bits suffer from a high error probability in the low SNR regime. Therefore, standard spatial diversity analysis shows that folding functions should be assigned to antennas to maximize the reliability of the most reliable output bit. The probability analysis in (4.26), (4.27) show that  $F_1(\cdot)$  corresponds to the most reliable bit. Maximizing this reliability means that the folding function  $F_1(\cdot)$  should be assigned to antenna  $i_1$ . Note that performance can be even further increased by assigning  $F_1(\cdot)$  to *all* antennas.

#### Antenna ordering in the moderate-to-high SNR regime

The goal is to maximize the number of bits conveyed through the channel, meaning to design the system to receive as many reliable bits as possible. It is straightforward to show that the quantization outputs  $b_M, \dots, b_1$  are ordered by reliability where  $b_M$  is the least reliable bit and  $b_1$  is the most reliable when we assume  $|h_1| = |h_2| = \dots = |h_M|$ . This ordering for the equal channel gain assumption follows from the fact that the value of  $\Pr(b_m = b_{\ell,m} | s = s_\ell, h_m)$  is dominated by the area of  $p(y_m | s = s_\ell, h_m)$  under the interval centered about the mean  $|h_m|s_\ell$ . The size of the interval at each antenna follows the relationship  $\frac{\Lambda_1}{2} > \frac{\Lambda_2}{2} > \dots > \frac{\Lambda_M}{2}$ .

If our goal is to get all bits  $b_1, \dots, b_M$  through reliably, we can maximize the reliability of the least reliable bit. This corresponds to mapping the  $m^{\text{th}}$  most reliable bit to the  $m^{\text{th}}$  least reliable channel. From a folding function perspective, this means  $F_m(\cdot)$  is assigned to  $i_{M+1-m}$ . This leads to modulo size  $\Lambda_{i_m} = 2^m d |h_{i_m}|$ , for  $i_m = i_1, \dots, i_M$ . Note that the sign

LMB folded low-resolution ADC array, which is described in Section 4.3.2, can adapt the antenna ordering strategies in both cases as well.

**Remark 4.** *Most practical systems exploit the statistical channel model to develop algorithms so that the reconstructed signal has less quantization error [24], [26], [81], [82], [99], [100]. However, different from [24], [26], [81], [82], [99], [100], we modify the parameters of the analog shift modulo operations to make each bit of the bit stream more reliable. It is unnecessary to consider the reconstructed signal because this would potentially lead to a loss of information [127].*

**Remark 5.** *Although finding an optimal input distribution is not the goal of this paper, it can be done using similar method as in [77]. This is because the folded low-resolution ADC array can be considered as a replacement of the centralized high-resolution ADC.*

## 4.4 System Throughput Analysis

We analyze the system throughput using different system models, quantizers, and QAM constellations. The complex domain mutual information can be considered as the combination of the independently derived mutual information in the real and imaginary spaces. Thus, we analyze the mutual information in the real space, which can be mapped to the complex space.

### 4.4.1 Mutual Information Preliminary

The general form of the mutual information in the real space is [127]

$$\mathcal{I}(\mathbf{b}, s|\mathbf{h}) = \mathcal{H}(\mathbf{b}|\mathbf{h}) - \mathcal{H}(\mathbf{b}|s, \mathbf{h}) \quad (4.33)$$

where  $\mathcal{H}(\mathbf{b}|\mathbf{h})$  is the entropy given by

$$\mathcal{H}(\mathbf{b}|\mathbf{h}) = - \sum_{\ell=0}^{2^M-1} \Pr(\mathbf{b} = \mathbf{b}_\ell|\mathbf{h}) \log[\Pr(\mathbf{b} = \mathbf{b}_\ell|\mathbf{h})]. \quad (4.34)$$

Using the fact that  $s$  is uniformly selected from  $\mathcal{S}$ , we write  $\Pr(\mathbf{b}|\mathbf{h})$  as [127]

$$\Pr(\mathbf{b}|\mathbf{h}) = \sum_{s \in \mathcal{S}} \frac{1}{2^M} \Pr(\mathbf{b}|s, \mathbf{h}), \quad (4.35)$$

where  $s$  is independent of  $\mathbf{h}$ .

Followed by this, we express  $\mathcal{H}(\mathbf{b}|s, \mathbf{h})$  as

$$\mathcal{H}(\mathbf{b}|s, \mathbf{h}) = - \sum_{\ell=0}^{2^M-1} \sum_{s \in \mathcal{S}} \frac{\Pr(\mathbf{b} = \mathbf{b}_\ell|s, \mathbf{h}) \log[\Pr(\mathbf{b} = \mathbf{b}_\ell|s, \mathbf{h})]}{2^M}. \quad (4.36)$$

#### 4.4.2 System Throughput Analysis in High SNR Regime

In this section, we begin with the definition of  $\varepsilon_e(P)$ .

**Definition 4.4.1.** *Let  $f(\cdot)$  is a real-valued function. We write  $f(P) = \varepsilon_e(P)$  if there exists  $u \in \mathbb{R}$  and  $v > 0$  such that  $|f(P)| \leq ue^{-vP}$ .*

Note that this means

$$\varepsilon_e(P) = c(P)\varepsilon_e(P), \quad (4.37)$$

for any finite order real polynomial  $c(P)$ .

**Lemma 3.** *In the high SNR regime, i.e.,  $P \rightarrow \infty$ , for any  $s \in \mathcal{S}$  and  $h_m$ , the conditional probability of error can be written*

$$\Pr(b_m \neq b_{\ell,m}|s = s_\ell, h_m) = \varepsilon_e(P), \quad (4.38)$$

*and the conditional probability of correct reception is written*

$$\Pr(b_m = b_{\ell,m}|s = s_\ell, h_m) = 1 - \varepsilon_e(P). \quad (4.39)$$

*Proof.* We start this proof by first finding an upper bound of all the conditional probabilities of error. Then, we use the definition of  $\varepsilon_e(P)$ , i.e., Definition 4.4.1, to generalize (4.38). Finally, (4.39) is obtained based on

$$\Pr(b_m = b_{\ell,m}|s = s_\ell, h_m) = 1 - \Pr(b_m \neq b_{\ell,m}|s = s_\ell, h_m) \quad (4.40)$$

The upper bound of all the conditional probability of error can be described by

$$\begin{aligned} \Pr(b_m \neq b_{\ell,m}|s = s_\ell, h_m) &\stackrel{(a)}{<} \Pr(y_m \in \tilde{\mathcal{R}}_m|s = s_\ell, h_m) \\ &= 2Q\left(\sqrt{\frac{2(d|h_m|)^2}{4}}\right) \\ &\stackrel{(b)}{\leq} 2e^{-\left(\frac{d|h_m|}{2}\right)^2}, \end{aligned} \quad (4.41)$$

where  $\Pr(y_m \in \tilde{\mathcal{R}}_m|s = s_\ell, h_m)$  is the probability of any  $y_m$  lies in  $\tilde{\mathcal{R}}_m$ , using

$$y_m \in \tilde{\mathcal{R}}_m = \left(-\infty, |h_m| \left(\frac{-d}{2} + s_\ell\right)\right) \cup \left[|h_m| \left(\frac{d}{2} + s_\ell\right), \infty\right),$$

given  $|h_m|s_\ell \in \left[|h_m|(\frac{-d}{2} + s_\ell), |h_m|(\frac{d}{2} + s_\ell)\right)$  based on (26). This is inspired by the error probability calculation for the inner PAM constellation points in an AWGN channel [128], [129]. Considering (4.26), (4.27), the inequality of (4.41a) follows the fact that  $\tilde{\mathcal{R}}_m \cup \mathcal{R}_m = \tilde{\mathcal{R}}_m$  and the area of  $p(y_m|s = s_\ell, h_m)$  under  $\mathcal{R}_m$  resulting  $\Pr(b_m \neq b_{m,\ell}|s = s_\ell, h_m)$ , for all  $m$ . Note that  $\mathcal{R}_m \in \{\mathcal{R}_{m,1}, \mathcal{R}_{m,2}\}$ , where  $\mathcal{R}_{m,1}$  and  $\mathcal{R}_{m,2}$  are defined by (4.23) and (4.24), respectively. Eqn. (4.41b) holds using the Chernoff bound of the  $Q$ -function [130].

Since  $d$  is written in terms of  $P$ , there exists  $v > 0$  depending on  $|h_1|, \dots, |h_M|$  such that  $\Pr(b_m \neq b_{\ell,m}|s = s_\ell, h_m) \leq ue^{-vP}$ , for some  $u \in \mathbb{R}$  and all  $m$ . Then, (4.38) and (4.39) are derived using Definition 4.4.1 and (4.40).  $\square$

**Theorem 4.4.1.** *In the high SNR regime, for any given  $h_m$  and  $s \in \mathcal{S}$ , the real-valued system throughput satisfies*

$$\mathcal{I}(\mathbf{b}, s | \mathbf{h}) = M - \varepsilon_e(P). \quad (4.42)$$

*Proof.* The proof starts with finding the expressions of  $\Pr(\mathbf{b} = \mathbf{b}_{\tilde{\ell}}|s = s_{\ell}, \mathbf{h})$ , for all  $\ell$ ,  $\tilde{\ell} = 0, \dots, 2^M - 1$ . Then, we derive the expressions of  $\mathcal{H}(\mathbf{b}|\mathbf{h})$  and  $\mathcal{H}(\mathbf{b}|s, \mathbf{h})$  to get  $\mathcal{I}(\mathbf{b}, s|\mathbf{h})$  based on  $\Pr(\mathbf{b} = \mathbf{b}_{\tilde{\ell}}|s = s_{\ell}, \mathbf{h})$ .

Thus, using (4.25), we first show

$$\Pr(\mathbf{b} = \mathbf{b}_{\ell}|s = s_{\ell}, \mathbf{h}) = \prod_{m=1}^M \Pr(b_m = b_{\ell,m}|s = s_{\ell}, h_m) = 1 - \varepsilon_e(P), \quad (4.43)$$

where (4.43) follows (4.37) and Lemma 3. Followed by this, we generalize

$$\Pr(\mathbf{b} = \mathbf{b}_{\tilde{\ell}}|s = s_{\ell}, \mathbf{h}) = \begin{cases} 1 - \varepsilon_e(P), & \text{for } \tilde{\ell} = \ell, \\ \varepsilon_e(P), & \text{for } \tilde{\ell} \neq \ell, \end{cases} \quad (4.44)$$

by first considering (4.43) and (4.45)

Then, we show

$$\Pr(\mathbf{b} \neq \mathbf{b}_{\ell}|s = s_{\ell}, \mathbf{h}) = (1 - \varepsilon_e(P))^{M_1} \varepsilon_e(P)^{M_2} = \varepsilon_e(P) \quad (4.45)$$

using (4.37) and Lemma 3, where  $M_1 + M_2 = M$ .

As we obtain (4.44), we show  $\Pr(\mathbf{b} = \mathbf{b}_{\ell}|\mathbf{h})$  before we derive the expression of  $\mathcal{H}(\mathbf{b}|\mathbf{h})$ .

In order to obtain

$$\Pr(\mathbf{b} = \mathbf{b}_{\ell}|\mathbf{h}) = \frac{1}{2^M} - \varepsilon_e(P), \quad (4.46)$$

we rewrite  $\Pr(\mathbf{b} = \mathbf{b}_{\ell}|\mathbf{h})$  in (4.35) as

$$\Pr(\mathbf{b} = \mathbf{b}_{\ell}|\mathbf{h}) \stackrel{(a)}{=} \frac{1}{2^M} \left( \Pr(\mathbf{b} = \mathbf{b}_{\ell}|s = s_{\ell}, \mathbf{h}) + \sum_{s \in \mathcal{S} \setminus \{s_{\ell}\}} \Pr(\mathbf{b} = \mathbf{b}_{\ell}|s, \mathbf{h}) \right) \quad (4.47)$$

Note that (4.47a) follows the fact that  $s_{\ell}$  has a unique binary representation  $\mathbf{b}_{\ell}$ . As we apply (4.44) into (4.47a), (4.46) is obtained using (4.37).



Then, the entropy  $\mathcal{H}(\mathbf{b}|\mathbf{h})$  is written in terms of (4.46) as shown

$$\begin{aligned}\mathcal{H}(\mathbf{b}|\mathbf{h}) &= - \sum_{\ell=0}^{2^M-1} \left( \frac{1}{2^M} - \varepsilon_e(P) \right) \log \left[ \frac{1}{2^M} - \varepsilon_e(P) \right] \\ &\stackrel{(a)}{=} M - \varepsilon_e(P),\end{aligned}\tag{4.48}$$

where (4.48a) follows (4.37) and

$$\begin{aligned}\log [1 - \varepsilon_e(P)] &\stackrel{(a)}{=} \sum_{i=0}^{\infty} \frac{(-1)^i (-\varepsilon_e(P))^{i+1}}{(i+1)\ln[2]} \\ &\stackrel{(b)}{=} \sum_{i=0}^L \frac{(-1)^i (-\varepsilon_e(P))^{i+1}}{(i+1)\ln[2]} + \varepsilon_e(P) \\ &\stackrel{(c)}{=} \varepsilon_e(P).\end{aligned}\tag{4.49}$$

Note that (4.49a) follows the Taylor series expansion. As we assume that there is a large  $L$  such that

$$\sum_{i=0}^{\infty} \frac{(-1)^i (\varepsilon_e(P))^{i+1}}{(i+1)\ln[2]} - \sum_{i=0}^L \frac{(-1)^i (\varepsilon_e(P))^{i+1}}{(i+1)\ln[2]} \leq u e^{-vP},\tag{4.50}$$

for some  $u \in \mathbb{R}$  and  $v > 0$  depend on  $L$ ,  $P$ , and  $|h_m|$  for all  $m$ , we extend right hand side of (4.49a) to obtain (4.49b) using (4.50) and Definition 4.4.1. This is because  $\log [1 - \varepsilon_e(P)]$  is bounded considering (4.46). Finally, we obtain (4.49c) using (4.37).

Before we approximate the expression of  $\mathcal{H}(\mathbf{b}|s, \mathbf{h})$ , some facts need to be addressed. Since Definition 4.4.1 indicates that  $\varepsilon_e(P) \in \mathbb{R}$ , we cannot use (4.45) to approximate the closed-form solution directly because  $\log [\varepsilon_e(P)]$  is not valid if  $\varepsilon_e(P) \leq 0$ . Thus, we need to discuss  $\Pr(\mathbf{b} \neq \mathbf{b}_\ell | s = s_\ell, \mathbf{h})$  before we provide the detailed derivation. The facts that  $0 \leq \Pr(\mathbf{b} \neq \mathbf{b}_\ell | s = s_\ell, \mathbf{h})$  and (4.45) indicate that

$$\Pr(\mathbf{b} \neq \mathbf{b}_\ell | s = s_\ell, \mathbf{h}) = |\varepsilon_e(P)|,\tag{4.51}$$

where  $|\varepsilon_e(P)| = \varepsilon_e(P)$  considering Definition 4.4.1. Thus, we use (4.51) to show

$$\Pr(\mathbf{b} \neq \mathbf{b}_\ell | s = s_\ell, \mathbf{h}) \log [\Pr(\mathbf{b} \neq \mathbf{b}_\ell | s = s_\ell, \mathbf{h})] = \frac{|\varepsilon_e(P)| \ln [|\varepsilon_e(P)|]}{\ln [2]} = \varepsilon_e(P)\tag{4.52}$$

where (4.52) follows  $|\varepsilon_e(P)| = \varepsilon_e(P)$  and (4.37).

Followed by this, we show

$$\begin{aligned}\mathcal{H}(\mathbf{b}|s, \mathbf{h}) &= - \sum_{\ell=0}^{2^M-1} \sum_{s \in \mathcal{S}} \frac{1}{2^M} \Pr(\mathbf{b} = \mathbf{b}_\ell | s, \mathbf{h}) \log \left[ \Pr(\mathbf{b} = \mathbf{b}_\ell | s, \mathbf{h}) \right] \\ &\stackrel{(a)}{=} \varepsilon_e(P).\end{aligned}\tag{4.53}$$

By expanding the summation of  $s$ , we obtain (4.53a) using (4.37) and (4.49), and (4.52). Then, (4.42) follows (4.33) and (4.37).  $\square$

**Remark 6.** *The conditional probability  $\Pr(b_m|s, h_m)$  depends on how the processed signal is quantized at each antenna. Therefore, the mutual information expression is dependent on the one-bit quantization structure at each antenna. This makes it intractable to generate exact mutual information expression for arbitrary SNR values and channel realizations. We numerically study the mutual information in Section 4.5 and discuss how the one-bit modulo architecture affects the mutual information for different SNRs.*

## 4.5 Numerical Analysis

In this section, we analyze an  $N$  transmit antenna by  $M$  receive antenna wireless MIMO system. We mainly consider two diverse channel models for our simulations. We first consider a single-path uniform planar array (UPA) channel meant to model mmWave channel [56], [57], [131]. To show the effect of rich multi-path (e.g., channels often found in sub-6 GHz bands), we consider the extreme case when the number of paths grows large by modeling the channel as a spatially i.i.d Rayleigh fading channel. The single-path UPA structure mmWave channel matrix is  $\mathbf{H} = g \mathbf{e}_r \mathbf{e}_t^*$ , where the channel gain  $g$  is distributed as  $\mathcal{CN}(0, 1)$ ,  $\mathbf{e}_r$  and  $\mathbf{e}_t$  are array response vectors corresponding to the receiver and the transmitter, respectively [56], [57], [131].

To understand the impact of the transmit antenna array, we also look at two diverse transmit beamformers. We consider non-adaptive beamforming where the multiple transmit antennas are used for transmit power enhancement by considering the beamforming vector  $\mathbf{f}_u = \frac{1}{\sqrt{N}}[1 \ \cdots \ 1]^T$ . We also look at optimal spatially adaptive beamforming using maximum

ratio transmission (MRT). The beamforming vector  $\mathbf{f}_{svd} = \mathbf{v}_1$ , where  $\mathbf{v}_1$  is the first column of the right singular value of  $\mathbf{H}$ .

The folded low-resolution ADC array simulations consider 4, 16, 64, and 256-QAM constellations, which correspond to  $M = 1, 2, 3$ , and 4, respectively. The sign-based low-resolution ADC array simulations consider 4, 16, 64, and 256-PSK constellations, which also correspond to  $M = 1, 2, 3$ , and 4, respectively. We assume that the transmit symbol is selected uniformly from  $2^{2M}$ -QAM and  $2^{2M}$ -PSK constellation. As  $M$  denotes the number of antennas, there are  $2M$  bits in the constellations (the real and the imaginary part of the constellation). Monte Carlo simulations are used to study the achievable rate of the designs of the folded low-resolution ADC array in Section 4.2.2 and Section 4.3.1, the designs of the sign LMB folded low-resolution ADC array in Section 4.3.2, and the sign-based low-resolution ADC array [79]. We simulate the theoretical achievable rate using (4.26), (4.27), and (4.33).

In Fig. 4.3, we show the achievable rates for the sign LMB folded low-resolution ADC array using both labelings and the sign-based low-resolution ADC array [79]. The channel was assumed to follow the single-path UPA mmWave channel with a transmitter performing beamforming  $\mathbf{f}_u$ . It is expected that the achievable rate of the design of the sign LMB folded low-resolution ADC array outperforms the sign-based low-resolution ADC array when  $M = 2, 3$  because the sign LMB folded low-resolution ADC array exploits the QAM constellation. The sign LMB folded low-resolution ADC array using Gray labeling offers superior performance because the modulo sizes used in Gray labeling are larger or equal than those of natural labeling. The relationships among the modulo sizes are shown as (4.29).

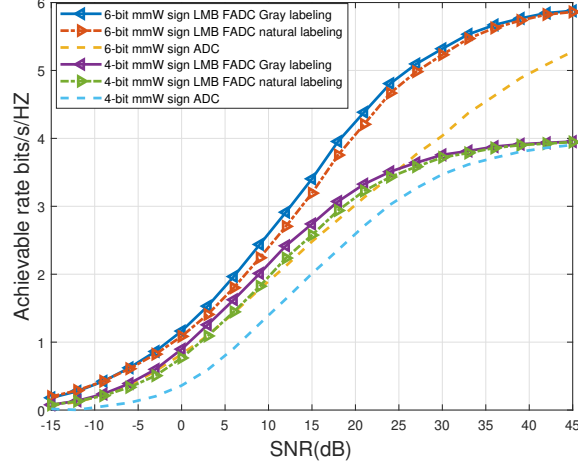
Fig. 4.4 shows the achievable rates for the sign LMB folded low-resolution ADC array using both labelings and the sign-based low-resolution ADC array. The simulations again used the single-path UPA mmWave channel and transmit beamformer  $\mathbf{f}_u$ . We assume that  $N = 16$ ,  $M = 1, 2, 3$ , and 4. We set the SNR to 42dB to emulate the high SNR regime in Fig. 4.3. The achievable rates of the sign LMB folded low-resolution ADC array using both labelings increase linearly with  $M$ . In contrast to the folded low-resolution ADC array, the achievable rate of the sign-based low-resolution ADC array increases logarithmically with  $M$ . Thus, Theorem 4.4.1 is validated.

In Fig. 4.5, we analyze the achievable rates of the sign LMB folded low-resolution ADC arrays for both labelings in the low SNR regime. We assume that  $N = 16$  and  $M = 2, 3$ . The channel was assumed to follow the single-path UPA mmWave channel with a transmitter performing beamforming with  $\mathbf{f}_u$ . The achievable rate ratio is defined as the simulated-derived achievable rate of the sign LMB folded low-resolution ADC array over the analytically-derived approximate achievable rate. The analytically-derived approximate achievable rate is obtained by considering the most reliable bit only using the antenna ordering strategy, which is written

$$\mathcal{I}(b_{i_1}, s \mid h_{i_1}) = \frac{P|h_{i_1}|^2}{\pi \ln [2]} + o(P), \quad (4.54)$$

where  $i_1 = \underset{m \in \{1, \dots, M\}}{\operatorname{argmax}} |h_m|$  and  $s$  is uniformly drawn from  $\mathcal{S}$ . The proof shares similarity with [79]. Fig. 4.5 shows the ratio between simulated-derived achievable rate and analytically-derived approximate achievable rate approximates to one which indicates that the output of the one-bit modulo quantizer is independent of its input. In addition, the achievable rate is dominated by the most significant bit.

In Fig. 4.6.(a) and Fig. 4.6.(b), we assume that  $N = 16$ ,  $M = 2$ , and  $\mathbf{f}_u$  is used. We compare the achievable rates for the designs of the sign LMB folded low-resolution ADC array and the folded low-resolution ADC array using the single-path UPA mmWave channel model in Fig. 4.6.(a) and the Rayleigh fading channels in Fig. 4.6.(b). In Fig. 4.6.(a), the achievable rates of the designs of the sign LMB folded low-resolution ADC arrays outperform the designs of the folded low-resolution ADC array at about -9 dB. In Fig. 4.6.(b), the achievable rates of the designs of the sign LMB folded low-resolution ADC arrays outperforms the designs of the folded low-resolution ADC arrays at about -4dB. In the low SNR regime, the achievable rate is dominated by the most reliable bit despite the design difference with respect to both labelings. As the SNR keeps increasing, the sign LMB folded low-resolution ADC array using Gray labelings provides better performance than the sign LMB folded low-resolution ADC array using natural labeling. Moreover, the achievable rate of the folded low-resolution ADC array using either labeling underperforms the sign LMB folded low-resolution ADC array with the same labeling as shown in both figures.

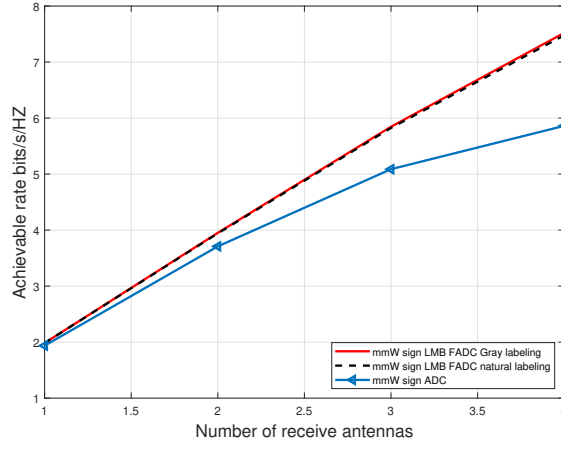


**Figure 4.3.** Achievable rates for sign LMB folded low-resolution ADC arrays using both labelings and sign-based low-resolution ADC array with single-path UPA structure mmWave channel when  $M = 2, 3$  and  $N = 16$ .

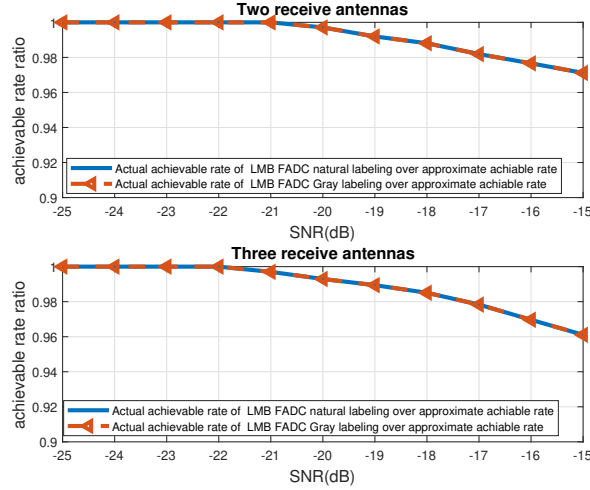
Fig. 4.7 shows the impact of antenna ordering strategies on achievable rate. The channel was assumed to follow Rayleigh fading channels with a transmitter performing beamforming  $\mathbf{f}_u$ . We assume that  $N = 16$  and  $M = 3$ . At low SNR, it is clear that assigning the sign quantizer to the antenna with the strongest magnitude channel, offers throughput enhancement. As the SNR increases, we assign  $\Lambda_{i_m} = 2^m d |h_{i_m}|$  to the antenna with the  $i_m^{\text{th}}$  strongest magnitude channel, and this approach offers an achievable rate improvement.

In Fig. 4.8, we assume that  $M = 1, 2, 3$  and  $N = 16$ . We analyze how beamforming affects the achievable rate given a design of the sign LMB folded low-resolution ADC array using Gray labeling. In particular, we compare the performance when we use  $\mathbf{f}_u$  and  $\mathbf{f}_{svd}$ . As we use single-path UPA structure for the mmWave channels, the performance of applying  $\mathbf{f}_{svd}$  is better than the performance of applying  $\mathbf{f}_u$  as expected.

Intuitively, for a fixed beamforming vector  $\mathbf{f}_{svd}$ , number of receive antenna  $M$ , and sign LMB folded low-resolution ADC array, the achievable rate can be further enhanced by increasing the number of transmit antennas. Fig. 4.9 using  $N = 16$  and  $64$ ,  $M = 1, 2$ , and  $3$ . The rest of the assumptions are the same as in Fig. 4.8. As we increase  $N = 16$  to  $N = 64$ , the achievable rate behaves as expected.

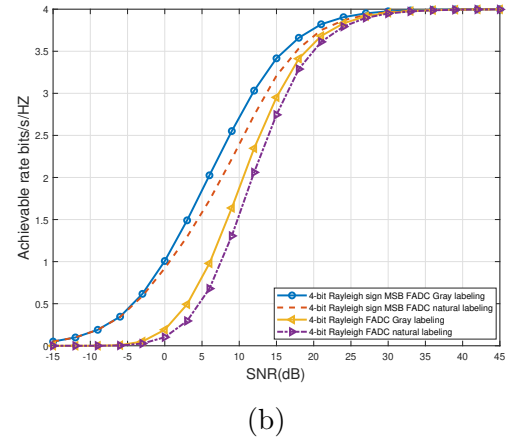
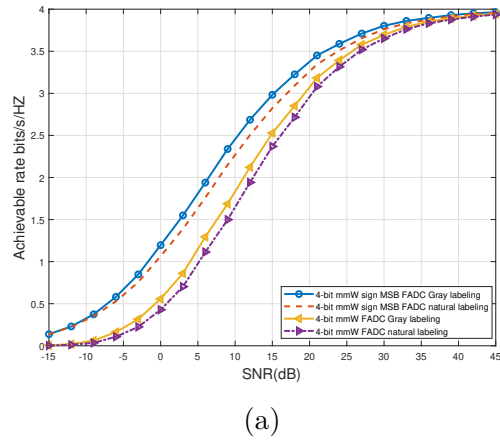


**Figure 4.4.** Achievable rates for sign LMB folded low-resolution ADC array using both labelings and sign-based low-resolution ADC array with single-path UPA structure mmWave channel when  $M = 1, 2, 3, 4$ ,  $N = 16$ .

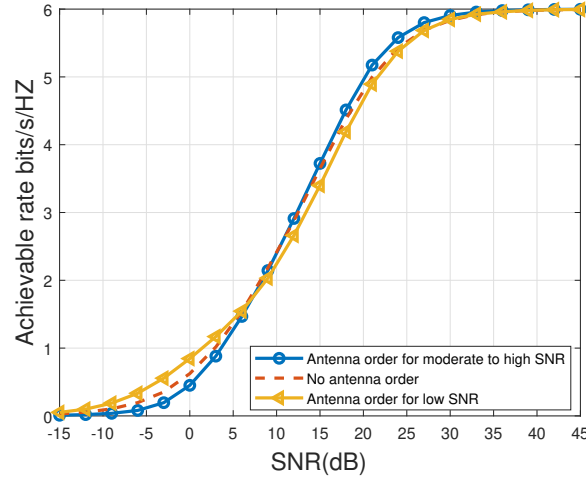


**Figure 4.5.** Ratio of achievable rate over approximate achievable rate with single-path UPA mmWave channel when  $M = 2, 3$  and  $N = 16$ .

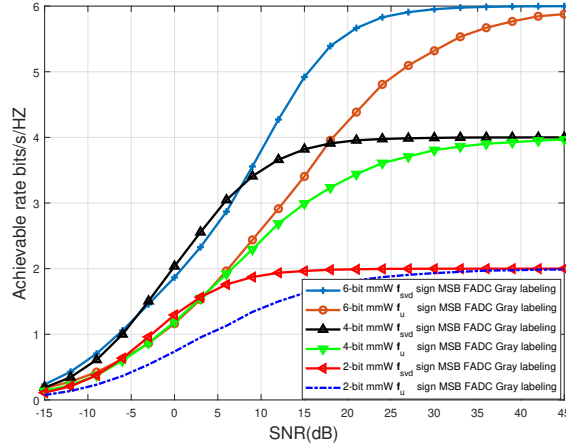
**Remark 7.** In our paper, we propose an architecture that can utilize any of the widely published centralized FADC implementations found in [67] by placing the centralized ADC branches at different antenna elements. Unlike a single antenna system, our system model exploits spatial diversity using multiple antennas. Like nearly all other low-resolution ADC arrays, the quantization at each array element makes it impossible to blindly combine the bits using any of the typical receive combiners. If standard receive combining is desired, it is



**Figure 4.6.** Achievable rates for designs of sign LMB folded low-resolution ADC arrays v.s. folded low-resolution ADC arrays when  $M = 2$  and  $N = 16$ . (a) and (b) use single-path UPA mmWave channel and Rayleigh fading channel, respectively.



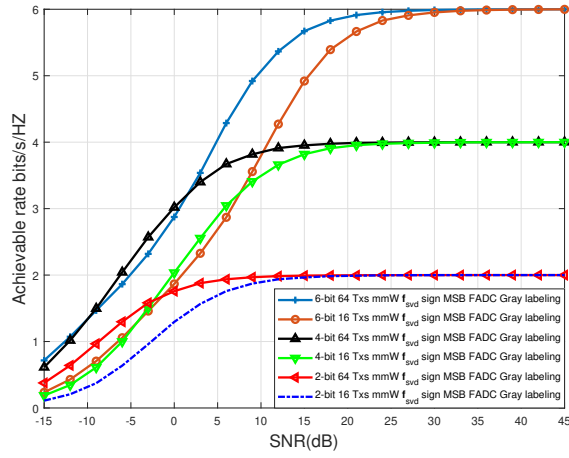
**Figure 4.7.** Achievable rates for LMB folded low-resolution ADC arrays with antenna ordering for moderate-to-high SNR, antenna ordering for low SNR, and no antenna ordering with Rayleigh fading channel when  $M = 3$  and  $N = 16$ .



**Figure 4.8.** UPA structure mmWave fading channel model achievable rates of sign LMB folded low-resolution ADC arrays using Gray labeling comparing  $\mathbf{f}_u$  v.s.  $\mathbf{f}_{svd}$  when  $M = 1, 2, 3$  and  $N = 16$ .

*better to perform this processing in the analog domain and then use a single high-resolution ADC.*





**Figure 4.9.** UPA mmWave channel model achievable rates of sign LMB folded low-resolution ADC arrays using Gray labeling for  $N = 16$  and  $N = 64$  given  $M = 1, 2$ , and  $3$  and beamforming vector  $\mathbf{f}_{svd}$ .

## 5. SUMMARY

In this thesis, we study two topics that can resolve the wireless communication demands, which are WSNs and mmWave systems. In the WSNs, our first topics develop an RAIF algorithm based on a proposed vector quantizer to enhance the system throughput in the quantized-and-forward relaying. Then, the second topic introduces an online supervised learning technique to enhance the classification performance of the classification problems. In mmWave systems, we develop a folded low-resolution ADC array to resolve issues such as the high power consumption and cost in the traditional communication system.

For our first topic, we proposed an end-to-end wireless system model that uses cooperative distributed relay networking. A lattice-based quantization strategy with low computational complexity was introduced to enhance the achievable rate performance. We also derived a rate-allocation scheme that made the system more sensitive to channel variations. Numerical simulations were used to show the advantages of the proposed algorithm comparing to prior works.

For our second topic, we proposed a novel framework that combines quantization, sensor selection strategy, classification to maximize the ACR to develop an online learning technique which is called MSOKSQ. We analyzed how the quantization rules are updated based on the proposed algorithm and the convergence bound that of the OL approaches theoretically. We provided numerical results to back our propositions and lemmas. Moreover, the simulation results validate our theoretical results. In our future work, we are interested in understanding the role the proposed system model in the field of wireless communication. Ideally, a realistic model for wireless communication should be combined with MSOKSQ. In particular, we use proposed algorithm to design the transmit beam-former, quantizer designs, and error control coding.

For our third topic, we proposed the folded low-resolution ADC array for multiple antenna wireless communication. We showed how the sign quantizer can be modified using properly designed folding functions to substantially enhance performance. We also discussed a number of practical enhancements including antenna ordering and optimized modulo regions. Simulation results validate our analysis. We would like to point out that our formulation

is a little different than that used in traditional digital communications modulation design. Our architecture's antennas do not directly detect the transmit symbol. Instead, we are designing an array that acts as a quantizer and uses bit assignments that are independent of those used by the transmitter. Thus, the proposed design can be implemented in distributed wireless communication [24], [26], [66], [99], [100], [132]–[134] as well. We believe that the use of high-resolution centralized ADC architectures in multiple antenna low-resolution arrays offers numerous research and implementation opportunities. There has been a tremendous amount of work over the last few decades on high-performance and low-power ADC design. If this research can be leveraged to develop new low-resolution arrays, there could be many performance and implementation benefits.

## REFERENCES

- [1] Google, “Project Loon,” <http://www.google.com/loon/>.
- [2] X. Tang and Y. Hua, “Optimal design of non-regenerative mimo wireless relays,” *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1398–1407, 2007.
- [3] O. Muñoz-Medina, J. Vidal, and A. Agustin, “Linear transceiver design in nonregenerative relays with channel state information,” *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2593–2604, 2007.
- [4] Y. Rong, X. Tang, and Y. Hua, “A unified framework for optimizing linear nonregenerative multicarrier mimo relay communication systems,” *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4837–4851, Dec. 2009, ISSN: 1053-587X. DOI: [10.1109/TSP.2009.2027779](https://doi.org/10.1109/TSP.2009.2027779).
- [5] W. Guan and H. Luo, “Joint mmse transceiver design in non-regenerative mimo relay systems,” *IEEE Communications Letters*, vol. 12, no. 7, pp. 517–519, 2008.
- [6] H. Bolcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, “Capacity scaling laws in mimo relay networks,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006, ISSN: 1536-1276. DOI: [10.1109/TWC.2006.1638664](https://doi.org/10.1109/TWC.2006.1638664).
- [7] A. S. Behbahani, R. Merched, and A. M. Eltawil, “Optimizations of a mimo relay network,” *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5062–5073, Oct. 2008, ISSN: 1053-587X. DOI: [10.1109/TSP.2008.929120](https://doi.org/10.1109/TSP.2008.929120).
- [8] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004, ISSN: 0018-9448. DOI: [10.1109/TIT.2004.838089](https://doi.org/10.1109/TIT.2004.838089).
- [9] T. Cover and A. E. Gamal, “Capacity theorems for the relay channel,” *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979, ISSN: 0018-9448. DOI: [10.1109/TIT.1979.1056084](https://doi.org/10.1109/TIT.1979.1056084).
- [10] J. N. Laneman and G. W. Wornell, “Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003, ISSN: 0018-9448. DOI: [10.1109/TIT.2003.817829](https://doi.org/10.1109/TIT.2003.817829).
- [11] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, 2005.

- [12] A. E. Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3413–3431, Oct. 2007, ISSN: 0018-9448. DOI: [10.1109/TIT.2007.904838](https://doi.org/10.1109/TIT.2007.904838).
- [13] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai, "Uplink macro diversity of limited backhaul cellular network," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3457–3478, 2009.
- [14] S. H. Lim, Y. H. Kim, A. E. Gamal, and S. Y. Chung, "Noisy network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011, ISSN: 0018-9448. DOI: [10.1109/TIT.2011.2119930](https://doi.org/10.1109/TIT.2011.2119930).
- [15] M. Yuksel and E. Erkip, "Multiple-antenna cooperative wireless systems: A diversity–multiplexing tradeoff perspective," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3371–3393, 2007.
- [16] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011, ISSN: 0018-9448. DOI: [10.1109/TIT.2011.2110110](https://doi.org/10.1109/TIT.2011.2110110).
- [17] X. Wu and L. L. Xie, "On the optimal compressions in the compress-and-forward relay schemes," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2613–2628, May 2013, ISSN: 0018-9448. DOI: [10.1109/TIT.2013.2241818](https://doi.org/10.1109/TIT.2013.2241818).
- [18] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011, ISSN: 0018-9448. DOI: [10.1109/TIT.2011.2165816](https://doi.org/10.1109/TIT.2011.2165816).
- [19] U. Niesen and P. Whiting, "The degrees of freedom of compute-and-forward," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5214–5232, Aug. 2012, ISSN: 0018-9448. DOI: [10.1109/TIT.2012.2197720](https://doi.org/10.1109/TIT.2012.2197720).
- [20] Z. Fang, X. Yuan, and X. Wang, "Towards the asymptotic sum capacity of the mimo cellular two-way relay channel," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4039–4051, Aug. 2014, ISSN: 1053-587X. DOI: [10.1109/TSP.2014.2332971](https://doi.org/10.1109/TSP.2014.2332971).
- [21] S. N. Hong and G. Caire, "Compute-and-forward strategies for cooperative distributed antenna systems," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5227–5243, Sep. 2013, ISSN: 0018-9448. DOI: [10.1109/TIT.2013.2265695](https://doi.org/10.1109/TIT.2013.2265695).
- [22] Y. Tan and X. Yuan, "Compute-compress-and-forward: Exploiting asymmetry of wireless relay networks," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 511–524, Jan. 2016, ISSN: 1053-587X. DOI: [10.1109/TSP.2015.2481876](https://doi.org/10.1109/TSP.2015.2481876).

- [23] O. Ordentlich and U. Erez, “Precoded integer-forcing universally achieves the mimo capacity to within a constant gap,” *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 323–340, 2015.
- [24] O. Ordentlich and U. Erez, “Integer-forcing source coding,” *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 1253–1269, Feb. 2017, ISSN: 0018-9448. DOI: [10.1109/TIT.2016.2629491](https://doi.org/10.1109/TIT.2016.2629491).
- [25] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. Cambridge University Press, 2014.
- [26] I. E. Bakoury and B. Nazer, “Integer-forcing architectures for uplink cloud radio access networks,” *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 67–75, Oct. 2017. DOI: [10.1109/ALLERTON.2017.8262720](https://doi.org/10.1109/ALLERTON.2017.8262720).
- [27] A. A. I. Ibrahim, J. Guo, D. J. Love, N. Yazdani, and K. W. Forsythe, “Receiver design and bit allocation for a multi-user distributed relay network performing vector quantization,” pp. 430–435, Nov. 2016. DOI: [10.1109/MILCOM.2016.7795365](https://doi.org/10.1109/MILCOM.2016.7795365).
- [28] X. Tang and Y. Hua, “Optimal design of non-regenerative mimo wireless relays,” *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007. DOI: [10.1109/TWC.2007.348336](https://doi.org/10.1109/TWC.2007.348336).
- [29] J. Guo, A. A. I. Ibrahim, A. Agaskar, D. J. Love, and N. Yazdani, “Implementation of rate-adaptive integer forcing compression in distributed wireless relay networking,” in *Proc. of the 52nd Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2018, pp. 1–6. DOI: [10.1109/CISS.2018.8362202](https://doi.org/10.1109/CISS.2018.8362202).
- [30] A. A. I. Ibrahim, J. Guo, D. J. Love, N. Yazdani, and K. W. Forsythe, “Receiver design and bit allocation for a multi-user distributed relay network performing vector quantization,” in *Proc. of IEEE Military Communications Conference*, Nov. 2016, pp. 430–435. DOI: [10.1109/MILCOM.2016.7795365](https://doi.org/10.1109/MILCOM.2016.7795365).
- [31] O. Ordentlich and U. Erez, “Integer-forcing source coding,” in *Proc. IEEE International Symposium on Information Theory*, Jun. 2014, pp. 181–185. DOI: [10.1109/ISIT.2014.6874819](https://doi.org/10.1109/ISIT.2014.6874819).
- [32] X. Wu and L. Xie, “On the optimal compressions in the compress-and-forward relay schemes,” *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2613–2628, May 2013. DOI: [10.1109/TIT.2013.2241818](https://doi.org/10.1109/TIT.2013.2241818).

- [33] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, “Wireless network information flow: A deterministic approach,” *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011. DOI: [10.1109/TIT.2011.2110110](https://doi.org/10.1109/TIT.2011.2110110).
- [34] M. Yuksel and E. Erkip, “Multiple-antenna cooperative wireless systems: A diversity–multiplexing tradeoff perspective,” *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3371–3393, Oct. 2007. DOI: [10.1109/TIT.2007.904972](https://doi.org/10.1109/TIT.2007.904972).
- [35] S. H. Lim, Y. Kim, A. El Gamal, and S. Chung, “Noisy network coding,” *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011. DOI: [10.1109/TIT.2011.2119930](https://doi.org/10.1109/TIT.2011.2119930).
- [36] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai, “Uplink macro diversity of limited backhaul cellular network,” *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3457–3478, Aug. 2009. DOI: [10.1109/TIT.2009.2023732](https://doi.org/10.1109/TIT.2009.2023732).
- [37] Q. Liang, X. Cheng, S. C. Huang, and D. Chen, “Opportunistic sensing in wireless sensor networks: Theory and application,” *IEEE Transactions on Computers*, vol. 63, no. 8, pp. 2002–2010, Aug. 2014, ISSN: 2326-3814. DOI: [10.1109/TC.2013.85](https://doi.org/10.1109/TC.2013.85).
- [38] J. Guo, R. G. Raj, and D. J. Love, “An online kernel scalar quantization scheme for signal classification,” in *Proc. of 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 3882–3886.
- [39] X. Nguyen, M. J. Wainwright, and M. I. Jordan, “Nonparametric decentralized detection using kernel methods,” *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4053–4066, Nov. 2005. DOI: [10.1109/TSP.2005.857020](https://doi.org/10.1109/TSP.2005.857020).
- [40] M. Wang, T. Lin, F. Ye, N. Li, and J. Ren, “A 1.2 v 1.0-GS/s 8-bit voltage-buffer-free folding and interpolating ADC,” in *Proc. 2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2012, pp. 274–277.
- [41] A. Koppel, A. S. Bedi, K. Rajawat, and B. M. Sadler, “Optimally compressed non-parametric online learning: Tradeoffs between memory and consistency,” *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 61–70, 2020.
- [42] A. Koppel, S. Paternain, C. Richard, and A. Ribeiro, “Decentralized online learning with kernels,” *IEEE Transactions on Signal Processing*, vol. 66, no. 12, pp. 3240–3255, 2018.
- [43] R. Viswanathan and P. K. Varshney, “Distributed detection with multiple sensors part i. fundamentals,” *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, 1997.

- [44] A. Nasipuri and S. Tantaratana, “Nonparametric distributed detector using wilcoxon statistics,” *Signal Processing*, vol. 57, no. 2, pp. 139–146, Mar. 1997, ISSN: 0165-1684. DOI: [https://doi.org/10.1016/S0165-1684\(96\)00191-0](https://doi.org/10.1016/S0165-1684(96)00191-0).
- [45] S. A. Kassam, “Nonparametric signal detection,” *Advances in Statistical Signal Processing*, Greenwich, CT: JAI, 1993. DOI: [10.1109/TIT.2007.904972](https://doi.org/10.1109/TIT.2007.904972).
- [46] M. M. Al- Ibrahim and P. Varshney, “Nonparametric sequential detection based on multisensor data,” in *Proc. 23rd Annu. Conf. Inf. Sci. Syst.*, 1989, pp. 157–162. DOI: [10.1109/MILCOM.2016.7795365](https://doi.org/10.1109/MILCOM.2016.7795365).
- [47] J. Predd, S. Kulkarni, and H. Poor, “Distributed learning in wireless sensor networks: Application issues and the problem of distributed inference,” *IEEE Signal Process.*, vol. 23, no. 4, pp. 56–69, Jul. 2006.
- [48] J. Predd, S. Kulkarni, and H. Poor, “Consistency in models for distributed learning under communication constraints,” *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 52–63, Jan. 2006.
- [49] J. Kivinen, A. J. Smola, and R. C. Williamson, “Online learning with kernels,” *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2165–2176, Aug. 2004. DOI: [10.1109/TSP.2004.830991](https://doi.org/10.1109/TSP.2004.830991).
- [50] C. Gentile, “A new approximate maximal margin classification algorithm,” *J. Mach. Learn. Res.*, vol. 2, pp. 213–242, Mar. 2002, ISSN: 1532-4435. [Online]. Available: <http://dl.acm.org/citation.cfm?id=944790.944811>.
- [51] O. Dekel, S. Shalev-shwartz, and Y. Singer, “The forgetron: A kernel-based perceptron on a fixed budget,” in *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. C. Platt, Eds., MIT Press, 2006, pp. 259–266. [Online]. Available: <http://papers.nips.cc/paper/2806-the-forgetron-a-kernel-based-perceptron-on-a-fixed-budget.pdf>.
- [52] F. Orabona, J. Keshet, and B. Caputo, “The projectron: A bounded kernel-based perceptron,” in *Proc. of the 25th International Conference on Machine Learning*, ser. ICML ’08, Helsinki, Finland: ACM, 2008, pp. 720–727, ISBN: 978-1-60558-205-4. DOI: [10.1145/1390156.1390247](https://doi.org/10.1145/1390156.1390247). [Online]. Available: <http://doi.acm.org/10.1145/1390156.1390247>.
- [53] R. Raj, “An online stochastic kernel machine for robust signal classification,” in *Proc. IEEE Asilomar Conference on Signals, Systems, and Computers*, Nov. 2019.



- [54] A. R. da Silva, M. H. T. Martins, B. P. S. Rocha, A. A. F. Loureiro, L. Ruiz, and H. C. Wong, “Decentralized intrusion detection in wireless sensor networks,” in *Proc. of the 1st ACM Int. Workshop on Qual. of Serv. Secur. Wireless Mobile Networks*, Montreal, Quebec, Canada, 2005, pp. 16–23, ISBN: 1595932410. DOI: [10.1145/1089761.1089765](https://doi.org/10.1145/1089761.1089765).
- [55] N. A. Goodman and D. Bruyere, “Optimum and decentralized detection for multistatic airborne radar,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 806–813, 2007.
- [56] J. Song, J. Choi, and D. J. Love, “Common codebook millimeter wave beam design: Designing beams for both sounding and communication with uniform planar arrays,” *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1859–1872, Apr. 2017, ISSN: 0090-6778. DOI: [10.1109/TCOMM.2017.2665497](https://doi.org/10.1109/TCOMM.2017.2665497).
- [57] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, “Millimeter wave beamforming for wireless backhaul and access in small cell networks,” *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013, ISSN: 0090-6778. DOI: [10.1109/TCOMM.2013.090513.120848](https://doi.org/10.1109/TCOMM.2013.090513.120848).
- [58] M. R. Castellanos, V. Raghavan, J. H. Ryu, O. H. Koymen, J. Li, D. J. Love, and B. Peleato, “Channel-reconstruction-based hybrid precoding for millimeter-wave multi-user MIMO systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 383–398, 2018. DOI: [10.1109/JSTSP.2018.2819135](https://doi.org/10.1109/JSTSP.2018.2819135).
- [59] N. N. Moghadam, G. Fodor, M. Bengtsson, and D. J. Love, “On the energy efficiency of MIMO hybrid beamforming for millimeter-wave systems with nonlinear power amplifiers,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7208–7221, 2018. DOI: [10.1109/TWC.2018.2865786](https://doi.org/10.1109/TWC.2018.2865786).
- [60] S. Noh, M. D. Zoltowski, and D. J. Love, “Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5689–5701, 2017. DOI: [10.1109/TWC.2017.2713357](https://doi.org/10.1109/TWC.2017.2713357).
- [61] Y. Lu, C. Cheng, J. Yang, and G. Gui, “Improved hybrid precoding scheme for mmwave large-scale MIMO systems,” *IEEE Access*, vol. 7, pp. 12 027–12 034, 2019. DOI: [10.1109/ACCESS.2019.2892136](https://doi.org/10.1109/ACCESS.2019.2892136).
- [62] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, “Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink,” *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4624–4634, Sep. 2017. DOI: [10.1109/TSP.2017.2715006](https://doi.org/10.1109/TSP.2017.2715006).

- [63] M. Shao, W. Ma, Q. Li, and A. L. Swindlehurst, “One-bit sigma-delta MIMO precoding,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 1046–1061, 2019. DOI: [10.1109/JSTSP.2019.2938687](https://doi.org/10.1109/JSTSP.2019.2938687).
- [64] H. Pirzadeh, G. Seco-Granados, S. Rao, and A. L. Swindlehurst, “Spectral efficiency of one-bit sigma-delta massive MIMO,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 2215–2226, 2020. DOI: [10.1109/JSAC.2020.3000887](https://doi.org/10.1109/JSAC.2020.3000887).
- [65] D. Ying, F. W. Vook, T. A. Thomas, and D. J. Love, “Hybrid structure in massive MIMO: Achieving large sum rate with fewer RF chains,” in *Proc. 2015 IEEE International Conference on Communications (ICC)*, 2015, pp. 2344–2349. DOI: [10.1109/ICC.2015.7248675](https://doi.org/10.1109/ICC.2015.7248675).
- [66] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, “Prospective multiple antenna technologies for beyond 5G,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1637–1660, 2020. DOI: [10.1109/JSAC.2020.3000826](https://doi.org/10.1109/JSAC.2020.3000826).
- [67] B. Murmann, *ADC performance survey 1997-2020*, [Online]. <http://www.stanford.edu/murmann/adcsurvey.html>.
- [68] R. H. Walden, “Analog-to-digital converter survey and analysis,” *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, Apr. 1999, ISSN: 0733-8716. DOI: [10.1109/49.761034](https://doi.org/10.1109/49.761034).
- [69] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, “Channel estimation in millimeter wave mimo systems with one-bit quantization,” in *Proc. of the 48th Asilomar Conference on Signals, Systems and Computers*, Nov. 2014, pp. 957–961. DOI: [10.1109/ACSSC.2014.7094595](https://doi.org/10.1109/ACSSC.2014.7094595).
- [70] J. Sung, J. Choi, and B. L. Evans, “Narrowband channel estimation for hybrid beamforming millimeter wave communication systems with one-bit quantization,” in *Proc. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 3914–3918.
- [71] A. Mezghani and A. L. Swindlehurst, “Blind estimation of sparse broadband massive MIMO channels with ideal and one-bit ADCs,” *IEEE Transactions on Signal Processing*, vol. 66, no. 11, pp. 2972–2983, Jun. 2018. DOI: [10.1109/TSP.2018.2821640](https://doi.org/10.1109/TSP.2018.2821640).
- [72] 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.

- [73] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, “Throughput analysis of massive MIMO uplink with low-resolution ADCs,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017. DOI: [10.1109/TWC.2017.2691318](https://doi.org/10.1109/TWC.2017.2691318).
- [74] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, “Channel estimation and performance analysis of one-bit massive mimo systems,” *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017. DOI: [10.1109/TSP.2017.2706179](https://doi.org/10.1109/TSP.2017.2706179).
- [75] G. Zeitler, A. C. Singer, and G. Kramer, “Low-precision a/d conversion for maximum information rate in channels with memory,” *IEEE Transactions on Communications*, vol. 60, no. 9, pp. 2511–2521, Sep. 2012, ISSN: 0090-6778. DOI: [10.1109/TCOMM.2012.071312.110682](https://doi.org/10.1109/TCOMM.2012.071312.110682).
- [76] O. Orhan, E. Erkip, and S. Rangan, “Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance,” in *2015 Information Theory and Applications Workshop (ITA)*, Feb. 2015, pp. 191–198. DOI: [10.1109/ITA.2015.7308988](https://doi.org/10.1109/ITA.2015.7308988).
- [77] A. Mezghani and J. A. Nossek, “On ultra-wideband mimo systems with 1-bit quantized outputs: Performance analysis and input optimization,” pp. 1286–1289, Jun. 2007, ISSN: 2157-8095. DOI: [10.1109/ISIT.2007.4557400](https://doi.org/10.1109/ISIT.2007.4557400).
- [78] 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, Dec. 2009, ISSN: 0090-6778. DOI: [10.1109/TCOMM.2009.12.080559](https://doi.org/10.1109/TCOMM.2009.12.080559).
- [79] 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, Oct. 2015, ISSN: 1053-587X. DOI: [10.1109/TSP.2015.2455527](https://doi.org/10.1109/TSP.2015.2455527).
- [80] J. Mo and R. W. Heath, “High SNR capacity of millimeter wave MIMO systems with one-bit quantization,” in *Proc. of the Information Theory and Applications Workshop (ITA)*, Feb. 2014, pp. 1–5. DOI: [10.1109/ITA.2014.6804238](https://doi.org/10.1109/ITA.2014.6804238).
- [81] N. Shlezinger, Y. C. Eldar, and M. R. D. Rodrigues, “Hardware-limited task-based quantization,” in *Proc of the IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2019, pp. 1–5. DOI: [10.1109/SPAWC.2019.8815422](https://doi.org/10.1109/SPAWC.2019.8815422).

- [82] N. Shlezinger, Y. C. Eldar, and M. R. D. Rodrigues, “Asymptotic task-based quantization with application to massive MIMO,” *IEEE Transactions on Signal Processing*, vol. 67, no. 15, pp. 3995–4012, Aug. 2019, ISSN: 1941-0476. DOI: [10.1109/TSP.2019.2923149](https://doi.org/10.1109/TSP.2019.2923149).
- [83] D. S. Palguna, D. J. Love, T. A. Thomas, and A. Ghosh, “Millimeter wave receiver design using low precision quantization and parallel delta-sigma architecture,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6556–6569, Oct. 2016, ISSN: 1536-1276. DOI: [10.1109/TWC.2016.2585640](https://doi.org/10.1109/TWC.2016.2585640).
- [84] A. Madanayake, N. Akram, S. Mandal, J. Liang, and L. Belostotski, “Improving ADC figure-of-merit in wideband antenna array receivers using multidimensional space-time delta-sigma multiport circuits,” in *Proc. 2017 10th International Workshop on Multidimensional (nD) Systems (nDS)*, 2017, pp. 1–6.
- [85] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhatieb, and G. C. Trichopoulos, “Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond,” *IEEE Access*, vol. 7, pp. 78 729–78 757, 2019.
- [86] I. Galton and H. T. Jensen, “Delta-sigma modulator based A/D conversion without oversampling,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 42, no. 12, pp. 773–784, Dec. 1995, ISSN: 1057-7130. DOI: [10.1109/82.476175](https://doi.org/10.1109/82.476175).
- [87] B. D. Smith, “An unusual electronic A/D conversion method,” *IRE Transactions on Instrumentation*, vol. PGI-5, pp. 155–160, Jun. 1956. DOI: [10.1109/IRE-I.1956.5007017](https://doi.org/10.1109/IRE-I.1956.5007017).
- [88] B. Gilbert, “A monolithic microsystem for analog synthesis of trigonometric functions and their inverses,” *IEEE Journal of Solid-State Circuits*, vol. 17, no. 6, pp. 1179–1191, Dec. 1982, ISSN: 0018-9200. DOI: [10.1109/JSSC.1982.1051878](https://doi.org/10.1109/JSSC.1982.1051878).
- [89] R. van de Grift, I. W. J. M. Rutten, and M. van der Veen, “An 8-bit video ADC incorporating folding and interpolation techniques,” *IEEE Journal of Solid-State Circuits*, vol. 22, no. 6, pp. 944–953, Dec. 1987, ISSN: 0018-9200. DOI: [10.1109/JSSC.1987.1052842](https://doi.org/10.1109/JSSC.1987.1052842).
- [90] W. T. Colleran and A. A. Abidi, “A 10-b, 75-MHz two-stage pipelined bipolar A/D converter,” *IEEE Journal of Solid-State Circuits*, vol. 28, no. 12, pp. 1187–1199, Dec. 1993.

- [91] B. Nauta and A. G. W. Venes, "A 70-ms/s 110-mw 8-b CMOS folding and interpolating A/D converter," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1302–1308, Dec. 1995, ISSN: 0018-9200. DOI: [10.1109/4.482155](https://doi.org/10.1109/4.482155).
- [92] P. Vorenkamp and R. Roovers, "A 12-b, 60-msample/s cascaded folding and interpolating ADC," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 12, pp. 1876–1886, Dec. 1997.
- [93] H. Pan and A. A. Abidi, "Signal folding in A/D converters," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 1, pp. 3–14, Jan. 2004, ISSN: 1549-8328. DOI: [10.1109/TCSI.2003.821278](https://doi.org/10.1109/TCSI.2003.821278).
- [94] S. D'Amico, G. Cocciolo, A. Spagnolo, M. De Matteis, and A. Baschirotto, "A 7.65-mW 5-bit 90-nm 1-GS/s folded interpolated ADC without calibration," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 2, pp. 295–303, 2014.
- [95] R. C. Taft, P. A. Francese, M. R. Tursi, O. Hidri, A. MacKenzie, T. Hohn, P. Schmitz, H. Werker, and A. Glenney, "A 1.8 V 1.0 GS/s 10b self-calibrating unified-folding-interpolating ADC with 9.1 ENOB at nyquist frequency," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 12, pp. 3294–3304, 2009.
- [96] J. van Valburg and R. J. van de Plassche, "An 8-b 650-MHz folding ADC," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 12, pp. 1662–1666, 1992.
- [97] Y. Nakajima, A. Sakaguchi, T. Ohkido, N. Kato, T. Matsumoto, and M. Yotsuyanagi, "A background self-calibrated 6b 2.7 GS/s ADC with cascade-calibrated folding-interpolating architecture," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 4, pp. 707–718, 2010.
- [98] K. R. Nary, R. Nubling, S. Beccue, W. T. Colleran, J. Penney, and Keh-Chung Wang, "An 8-bit, 2 gigasample per second analog to digital converter," in *GaAs IC Symposium IEEE Gallium Arsenide Integrated Circuit Symposium 17th Annual Technical Digest 1995*, 1995, pp. 303–306. DOI: [10.1109/GAAS.1995.529016](https://doi.org/10.1109/GAAS.1995.529016).
- [99] O. Ordentlich, G. Tabak, P. K. Hanumolu, A. C. Singer, and G. W. Wornell, "A modulo-based architecture for A/D conversion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 5, pp. 825–840, Oct. 2018. DOI: [10.1109/JSTSP.2018.2863189](https://doi.org/10.1109/JSTSP.2018.2863189).
- [100] J. Guo, A. A. I. Ibrahim, A. Agaskar, D. J. Love, and N. Yazdani, "Implementation of rate-adaptive integer forcing compression in distributed wireless relay networking," in *2018 52nd Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2018, pp. 1–6. DOI: [10.1109/CISS.2018.8362202](https://doi.org/10.1109/CISS.2018.8362202).

- [101] K. Rosen, *Elementary Number Theory and Its Applications (6th. Edition)*. Boston, MA, USA: Pearson/Addison-Wesley, 2005, ISBN: 9780321237071.
- [102] O. Ordentlich and U. Erez, “A simple proof for the existence of “good” pairs of nested lattices,” *IEEE Transactions on Information Theory*, vol. 62, no. 8, pp. 4439–4453, 2016.
- [103] U. Erez and R. Zamir, “Achieving  $1/2 \log(1 + \text{snr})$  on the AWGN channel with lattice encoding and decoding,” *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, 2004.
- [104] A. K. Lenstra, H. W. Lenstra, and L. Lovasz, “Factoring polynomials with rational coefficients,” *Mathematische Annalen*, vol. 261, no. 4, pp. 515–534, 1982.
- [105] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, Fourth. Boston: McGraw Hill, 2002.
- [106] B. Scholkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2001, ISBN: 0262194759.
- [107] K. Tsuda, T. Kin, and K. Asai, “Marginalized kernels for biological sequences,” *Bioinformatics (Oxford, England)*, vol. 18 Suppl 1, S268–75, Feb. 2002. DOI: [10.1093/bioinformatics/18.suppl\\_1.S268](https://doi.org/10.1093/bioinformatics/18.suppl_1.S268).
- [108] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [109] C. Blake and C. Merz, Univ. Calif. Irvine Repository of Machine Learning Databases, Irvine, CA, 1998.
- [110] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA, USA: Kluwer Academic Publishers, 1991, ISBN: 0-7923-9181-0.
- [111] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006, ISBN: 0387310738.
- [112] R. Rifkin and A. Klautau, “In defense of one-vs-all classification,” English, *J. Mach. Learn. Res.*, vol. 5, pp. 101–141, 2004, ISSN: 1532-4435; 1533-7928/e.
- [113] A. S. Nagra and R. A. York, “Distributed analog phase shifters with low insertion loss,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 9, pp. 1705–1711, Sep. 1999. DOI: [10.1109/22.788612](https://doi.org/10.1109/22.788612).



- [114] M. Chua and K. W. Martin, "1 GHz programmable analog phase shifter for adaptive antennas," in *Proc. of the IEEE 1998 Custom Integrated Circuits Conference (Cat. No.98CH36143)*, May 1998, pp. 71–74. DOI: [10.1109/CICC.1998.694909](https://doi.org/10.1109/CICC.1998.694909).
- [115] Y. Lee, C. Wang, and Y. Huang, "A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 305–317, Jan. 2015. DOI: [10.1109/TSP.2014.2370947](https://doi.org/10.1109/TSP.2014.2370947).
- [116] V. Venkateswaran and A. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010. DOI: [10.1109/TSP.2010.2048321](https://doi.org/10.1109/TSP.2010.2048321).
- [117] Junyi Wang, Zhou Lan, Chang-woo Pyo, T. Baykas, Chin-sean Sum, M. A. Rahman, Jing Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave wpan systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009. DOI: [10.1109/JSAC.2009.091009](https://doi.org/10.1109/JSAC.2009.091009).
- [118] S. H. Lewis and P. R. Gray, "A pipelined 5-Msample/s 9-bit analog-to-digital converter," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 6, pp. 954–961, Dec. 1987, ISSN: 0018-9200. DOI: [10.1109/JSSC.1987.1052843](https://doi.org/10.1109/JSSC.1987.1052843).
- [119] A. Rényi, "Representations for real numbers and their ergodic properties," *Acta Mathematica Hungarica*, vol. 8, no. 3-4, pp. 477–493, 1957, ISSN: 0236-5294.
- [120] W. Parry, "On the  $\beta$ -expansions of real numbers," *Acta Mathematica Hungarica*, vol. 11, pp. 401–416, Sep. 1960. DOI: [10.1007/BF02020954](https://doi.org/10.1007/BF02020954).
- [121] G. He, J. Belfiore, X. Liu, Y. Ge, R. Zhang, I. Land, Y. Chen, R. Li, J. Wang, G. Yang, and W. Tong, " $\beta$ -expansion: A theoretical framework for fast and recursive construction of polar codes," *CoRR*, vol. abs/1704.05709, 2017. arXiv: [1704.05709](https://arxiv.org/abs/1704.05709). [Online]. Available: <http://arxiv.org/abs/1704.05709>.
- [122] Yuen, "Analog-to-gray code conversion," *IEEE Transactions on Computers*, vol. C-27, no. 10, pp. 971–973, Oct. 1978. DOI: [10.1109/TC.1978.1674979](https://doi.org/10.1109/TC.1978.1674979).
- [123] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998, ISSN: 0018-9448. DOI: [10.1109/18.669123](https://doi.org/10.1109/18.669123).

- [124] E. Zehavi, “8-PSK trellis codes for a Rayleigh channel,” *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 873–884, May 1992, ISSN: 0090-6778. DOI: [10.1109/26.141453](https://doi.org/10.1109/26.141453).
- [125] E. Agrell, J. Lassing, E. G. Strom, and T. Ottosson, “Gray coding for multilevel constellations in Gaussian noise,” *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 224–235, Jan. 2007, ISSN: 0018-9448. DOI: [10.1109/TIT.2006.887070](https://doi.org/10.1109/TIT.2006.887070).
- [126] M. Mano and M. Ciletti, *Digital Design*. Pearson, 2017, ISBN: 9780134549897. [Online]. Available: <https://books.google.com/books?id=OhMdvGAACAAJ>.
- [127] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006, ISBN: 0471241954.
- [128] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005, ISBN: 0521837162.
- [129] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing (4th Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006, ISBN: 0131873741.
- [130] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ninth Dover printing, tenth GPO printing. New York: Dover, 1964.
- [131] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005, ISBN: 0-5218-4527-0.
- [132] J. Choi, D. J. Love, and T. P. Bidigare, “Coded distributed diversity: A novel distributed reception technique for wireless communication systems,” *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1310–1321, Mar. 2015, ISSN: 1053-587X. DOI: [10.1109/TSP.2015.2389766](https://doi.org/10.1109/TSP.2015.2389766).
- [133] J. Choi, D. J. Love, D. R. Brown, and M. Boutin, “Quantized distributed reception for MIMO wireless systems using spatial multiplexing,” *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3537–3548, 2015. DOI: [10.1109/TSP.2015.2424193](https://doi.org/10.1109/TSP.2015.2424193).
- [134] D. R. Brown III, U. Madhow, M. Ni, M. Rebholz, and P. Bidigare, “Distributed reception with hard decision exchanges,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3406–3418, 2014. DOI: [10.1109/TWC.2014.051314.131763](https://doi.org/10.1109/TWC.2014.051314.131763).