UNIVERSITY OF CALIFORNIA,
IRVINE

Massive MIMO with Low-Resolution ADCs

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical and Computer Engineering

by

Hessam Pirzadeh

Dissertation Committee:
Professor A. Lee Swindlehurst, Chair
Professor Ender Ayanoglu
Professor Yanning Shen

2021

DEDICATION

To my Teachers and my Family.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Professor Lee Swindlehurst. I would have not been able to proceed through my doctoral degree without his support and encouragement. Before every meeting, I had a messy mind and was full of stress but after the meeting, I was hopeful about the outcome of my research and motivated. His knowledge and vision in research enlightened my academic path, and his attitude toward life will have lasting impact through my life. I was so lucky to have him as my advisor. Thank you LS for making my doctoral degree the best period of my life.

I appreciate Prof. Ayanoglu's and Prof. Shen's willingness to serve as committee members. I should also thank Prof. Gonzalo Seco-Granados, Prof. Amine Mezghani, and Prof. Josef A. Nossek for their mentorship and support. I spent many hours with them coping with my technical research problems and I learned a lot. It has also been a pleasure to work with and learn from mentors at DOCOMO Innovations and Qualcomm. Finally, I would like to thank my group-mate, Shilpa Rao, for her helpful collaboration. The financial support of my research from the National Science Foundation under grant ECCS-1547155, ECCS-1824565, CCF-1703635, and the Hans Fischer Senior Fellowship from the Technische Universität München Institute for Advanced Study is gratefully acknowledged.

At last, Mom, my appreciation and gratitude to you for your endless support and kindness cannot be put into words.

CURRICULUM VITAE

Hessam Pirzadeh

EDUCATION

Doctor of Philosophy in Electrical and Computer Engineering **2021**
University of California Irvine *Irvine, California*

Master of Science in Electrical Engineering **2016**
Iran University of Science and Technology *Tehran, Iran*

Bachelor of Science in Engineering Science **2013**
University of Tehran *Tehran, Iran*

RESEARCH EXPERIENCE

Graduate Research Assistant **2016–2021**
University of California, Irvine *Irvine, California*

Engineering Intern **Summer 2020**
Qualcomm Inc. *Santa Clara, California*

Research Engineer Intern **Summer 2018**
DOCOMO Innovations Inc. *Palo Alto, California*

TEACHING EXPERIENCE

Teaching Assistant **2019–2020**
University of California, Irvine *Irvine, California*

JOURNAL PUBLICATIONS

1. H. Pirzadeh, G. Seco-Granados, S. Rao, and A. Lee Swindlehurst, **"Spectral Efficiency of One-bit Sigma-Delta Massive MIMO,"** *IEEE J. Sel. Areas in Commun.*, vol. 38, no. 9, pp. 2215-2226, Sep. 2020.

2. H. Pirzadeh, and A. Lee Swindlehurst, **"Spectral Efficiency of Mixed-ADC Massive MIMO,"** *IEEE Trans. Sig. Process.*, vol. 66, no. 13, pp. 3599-3613, July 2018.

3. K. Roth, H. Pirzadeh, A. Lee Swindlehurst, and J. A. Nossek,, **"A Comparison of Hybrid Beamforming and Digital Beamforming with Low-Resolution ADCs for Multiple Users and Imperfect CSI,,"** *IEEE J. Sel. Topics Sig. Process.*, vol. 12, no. 3, pp. 484-498, June 2018.

4. H. Pirzadeh, and A. Lee Swindlehurst, **"Spectral Efficiency under Energy Constraint for Mixed-ADC MRC Massive MIMO,"** *IEEE Sig. Process. Lett.*, vol. 24, no. 12, pp. 1847-1851, Dec. 2017.

5. H. Pirzadeh, S. M. Razavizadeh, and E. Björnson, , **"Subverting Massive MIMO by Smart Jamming,"** *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 20-23, Feb. 2016.

CONFERENCE PUBLICATIONS

1. H. Pirzadeh, G. Seco-Granados, A. Lee Swindlehurst, and J. A. Nossek, **"On the Effect of Mutual Coupling in One-Bit Spatial Sigma-Delta Massive MIMO Systems,,"** in *Proc. IEEE Intl Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, May 2020.

2. H. Pirzadeh, G. Seco-Granados, and A. Lee Swindlehurst, **"Mitigation of Jamming Attack in Massive MIMO With One-Bit FBB Sigma-Delta ADCs,,"** in *Proc. 53rd Asilomar Conf. Signals Syst. Comput.*, Nov. 2019.

3. H. Pirzadeh, and A. Lee Swindlehurst, **"Space-Constrained Mixed-ADC Massive MIMO,,"** in *Proc. IEEE Intl Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019.

4. H. Pirzadeh, C. Wang, and H. Papadopoulos, **"Machine-Learning-Assisted Outdoor Localization via Sector-Based Fog Massive MIMO,,"** in *IEEE Intl Conference on Communications (ICC)*, May 2019.

5. S. Rao, A. Lee Swindlehurst, and H. Pirzadeh, **"Massive MIMO Channel Estimation with 1-Bit Spatial Sigma-Delta ADCs,,"** in *IEEE Intl Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019.

6. K. Roth, H. Pirzadeh, A. Lee Swindlehurst, and J. A. Nossek,, **"Are the Data Rates Predicted by the Analytic Analysis of Receivers with Low Resolution ADCs Achievable?,"** in *Proc. European Conference on Networks and Communications (EuCNC)*, June 2018.

7. H. Pirzadeh, and A. Lee Swindlehurst, **"On the Optimality of Mixed-ADC Massive MIMO with MRC Detection,,"** in *Proc. Intl ITG Workshop Smart Antennas (WSA)*, March 2017.

8. H. Pirzadeh, and A. Lee Swindlehurst, **"Analysis of MRC for Mixed-ADC Massive MIMO,,"** in *Proc. Intl Workshop on Computational Advances in Multi-Channel Sensor Array Processing (CAMSAP)*, Dec. 2017.

PROFESSIONAL MEMBERSHIP

IEEE Communications and Signal Processing Societies.

ABSTRACT OF THE DISSERTATION

Massive MIMO with Low-Resolution ADCs

By

Hessam Pirzadeh

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Irvine, 2021

Professor A. Lee Swindlehurst, Chair

A key potential of massive multiple-input multiple-output (MIMO) systems which has made it interesting from a practical standpoint is its ability to substantially increase the network capacity with inexpensive, low-power components. Nevertheless, the static power consumption at the base station (BS) will increase proportionally to the number of antennas. Hence, considering hardware-aware design together with power consumption at the BS seems necessary in realizing practical massive MIMO systems.

Among the various components responsible for power dissipation at the BS, the contribution of analog-to-digital converters (ADCs) is known to be dominant. Consequently, the idea of replacing the high-power high-resolution ADCs with power efficient low-resolution ADCs could be a viable approach to address power consumption concerns at the massive MIMO BSs.

In this thesis, we study two different architectures in design of massive MIMO systems with one-bit ADCs, namely, mixed-ADC architecture and spatial $\Sigma\Delta$ architecture. The basic premise behind the mixed-ADC architecture is to achieve the benefits of conventional massive MIMO systems by just exploiting small pairs of high-resolution ADCs. In spatial $\Sigma\Delta$ architecture, by subtracting the quantized output of one antennas radio frequency (RF) chain from the signal at an adjacent antenna, coupled with spatial oversampling, the quantization

noise is shaped to angular regions that the signal of interest is not present.

We start with finding the optimal distribution of high-resolution and one-bit ADCs in a base station with energy constraint to maximize the spectral efficiency. Then, we use the Bussgang decomposition to develop a linear minimum mean-squared error (LMMSE) channel estimator for mixed-ADC architecture based on the combined round-robin measurements and we derive a closed-form expression for the resulting mean-squared error (MSE). We also perform a spectral efficiency (SE) analysis of the mixed-ADC implementation for the maximum ratio combining (MRC) and zero-forcing (ZF) receivers, and obtain expressions for a lower bound on the SE that takes into account the channel estimation error and the loss of efficiency due to the round-robin training. Finally, the possible SE improvement that can be achieved by using an antenna selection algorithm is investigated.

Next we introduce the spatial $\Sigma\Delta$ architecture by adopting the oversampling and $\Sigma\Delta$ quantization approach in the time domain signal processing. We propose the appropriate design of spatial $\Sigma\Delta$ architecture by applying a scalar version of Bussgang approach. The results of the analysis indicate the significant gain of the $\Sigma\Delta$ approach compared with standard one-bit quantization for users that lie in the angular sector where the shaped quantization error spectrum is low. To alleviate the impact of strong interference in systems with one-bit quantizer, we extend the $\Sigma\Delta$ approach to present spatial feedback beamformer (FBB) $\Sigma\Delta$ architecture. We compare the symbol error rate of the FBB $\Sigma\Delta$ array with that of a system with high-resolution ADCs. The results show the superior performance of the one-bit FBB $\Sigma\Delta$ architecture which achieves performance equivalent to that of a system with only high resolution ADCs.

To complete our analysis, we study the impact of mutual coupling and space-constrained arrays on the performance of mixed-ADC and spatial $\Sigma\Delta$ architectures. This phenomenon can eliminate the need for round-robin training for channel estimation in mixed-ADC architectures. For spatial $\Sigma\Delta$ architecture which relies on spatial oversampling (antenna spacing

less than half a wavelength), the impact of mutual coupling may become significant as the antenna spacing decreases. Unlike temporal oversampling, there is a limit to the amount of spatial oversampling that can be achieved, due to the physical dimensions of the antennas. In the last section of this dissertation, we show that the one-bit $\Sigma\Delta$ array is particularly advantageous in space-constrained scenarios, and can still provide significant gains even in the presence of mutual coupling when the antennas are closely spaced. Through this thesis, we show that by exploiting the advanced capabilities of (MIMO) signal processing methods, performance of massive MIMO systems with coarse quantization can be significantly improved.

Chapter 1

Introduction

The seminal work of Marzetta introduced massive multiple-input multiple-output (MIMO) as a promising architecture for future wireless systems [1]. In the limit of an infinite number of base station (BS) antennas, it was shown that massive MIMO can substantially increase the network capacity. Another key potential of massive MIMO systems which has also made it interesting from a practical standpoint is its ability of achieving this goal with inexpensive, low-power components [2, 3]. However, preliminary studies on massive MIMO systems have for the most part only analyzed its performance under the assumption of perfect hardware [4, 5]. The impact of hardware imperfections and nonlinearities on massive MIMO systems has recently been investigated in [6, 7, 8, 9, 10]. Although it is well-known that the dynamic power in massive MIMO systems can be scaled down proportional to $\sqrt{M}$, where M denotes the number of BS antennas, the static power consumption at the BS will increase proportionally to M [11]. Hence, considering hardware-aware design together with power consumption at the BS seems necessary in realizing practical massive MIMO systems.

Among the various components responsible for power dissipation at the BS, the contribution of analog-to-digital converters (ADCs) is known to be dominant [12]. Particularly, since the

power consumption of ADCs grows linearly with the system bandwidth, for millimeter wave (mmW) systems where the spectrum ranges between $10 - 30$ GHz, using high-resolution ADCs is not a feasible approach. Consequently, different architectures are proposed in literature to cope with this problem. For example, hybrid beamforming (HB) which reduces the number of radio-frequency (RF) chains or fully digital beamforming (DB) in which the high-power high-resolution ADCs are replaced with power efficient low-resolution ADCs are viable candidates to address the power consumption concerns at the massive MIMO BSs [13, 14]. A comprehensive comparison between these two schemes is performed in [13]. It is shown that unlike usual claim, DB does not have a higher power consumption than HB in all scenarios. Depending on the power consumption model, number of ADC bits, number of antenna elements, technical characteristics of analog receiver components, and even environmental changes (such as channel and SNR), each scheme can have a better spectral or energy efficiency. Based on this study, DB always has a better performance than HB in terms of SE. However, depending on the number of ADC bits and SE/EE trade-offs, HB can provides slightly better EE in some scenarios. Generally, DB is preferable in an architecture with low-power ADCs and moderately high-power phase shifters (PSs) while HB provides a better performance in a scheme with cheap receivers fabricated with high-power ADCs and low-power PSs. Besides, it is worthwhile to note that since the characteristics of mmW systems are not completed yet, DB versatility is an advantageous feature compared with HB. In addition, reducing the number of ADC bits in DB leads to less fronthaul rate. Nevertheless, HB is a better option in terms of component dimension and cost. Shortly, the appropriate choice of beamforming directly depends on the selection of receiver components and model parameters. In this dissertation, we consider DB and study the performance of the proposed architectures.

The impact of utilizing low-resolution ADCs on the spectral efficiency (SE) and energy consumption of massive MIMO systems has been considered in [15, 16, 17, 18, 19, 20, 21, 22, 23]. In particular, studies on massive MIMO systems with purely one-bit ADCs show

that the high spatial multiplexing gain owing to the use of a large number of antennas is still achievable even with one-bit ADCs [15, 16]. However, many more antennas with one-bit ADCs (at least 2-2.5 times) are required to attain the same performance as in the high-resolution ADCs case.

One of the main causes of SE degradation in purely one-bit massive MIMO systems is the error due to the coarse quantization that occurs during the channel estimation phase. While at low SNR the loss due to one-bit quantization is only about 2 dB, at higher SNRs performance degrades considerably more and leads to an error floor [15]. The SE degradation can be reduced by improving the quality of the channel estimation prior to signal detection. One approach for doing so is to exploit so-called *mixed-ADC architectures* during the channel estimation phase, in which a combination of low- and high-resolution ADCs are used side-by-side. Mixed-ADC implementations were introduced in [24, 25] and their performance was studied from an information theoretic perspective via generalized mutual information.

The basic premise behind the mixed-ADC architecture is to achieve the benefits of conventional massive MIMO systems by just exploiting $N \ll M$ pairs of high-resolution ADCs. An SE analysis of mixed-ADC massive MIMO systems with maximum ratio combining (MRC) detection for Rayleigh and Rician fading channels was carried out in [26] and [27], respectively. The SE and energy efficiency of mixed-ADC systems compared with systems composed of one-bit ADCs was studied in [28] for MRC detection, and conditions were derived under which each architecture provided the highest SE for a given power consumption. The advantage of using a mixed-ADC architecture in designing Bayes-optimal detectors for MIMO systems with low-resolution ADCs is reported in [29]. Although the nonlinearity of the quantization process increases the complexity of the optimal detectors, it is shown that adding a small number of high-resolution ADCs to the system allows for less complex detectors with only a slight performance degradation. Moreover, the benefit of using mixed-ADC architectures in massive MIMO relay systems and cloud-RAN deployments is elaborated in

[30, 31].

Another approach to mitigate the performance loss due to using low-resolution ADCs is temporal oversampling with a corresponding increase in complexity and power consumption [32, 33, 34, 35, 36]. Oversampled one-bit quantization has a long history in digital signal processing, particularly using the so-called Sigma-Delta ($\Sigma\Delta$) approach, which quantizes the difference (Δ) between the signal and its previously quantized value, and then integrates (Σ) the resulting output [37, 38, 39]. This has the effect of shaping the quantization noise to higher frequencies, while the signal occupies the low end of the spectrum due to the oversampling. Higher-order $\Sigma\Delta$ modulators can be constructed that provide increased shaping of the quantization noise from low to high frequencies. Compared with a standard one-bit ADC, a $\Sigma\Delta$ ADC requires additional digital circuitry to implement the integration, but very little additional RF hardware. $\Sigma\Delta$ ADCs have been commonly used in process control and instrumentation applications, and more recently in the implementation of multi-channel beamformers for ultrasound imaging systems.

The concept of $\Sigma\Delta$ modulation can also be applied in the spatial as well as the temporal domain. In a *spatial* $\Sigma\Delta$ implementation, the difference signal is formed by subtracting the quantized output of one antennas RF chain from the signal at an adjacent antenna. Coupled with spatial oversampling (*e.g.,* a uniform linear array with elements separated by less than one half wavelength), the quantization noise is shaped to higher *spatial* frequencies, and significantly reduced for signals arriving in a sector around broadside (0°). Applying a phase shift to the feedback signal allows one to move the band of low quantization error to different angular regions.

Relatively little research has focused on the spatial $\Sigma\Delta$ architecture. Prior related work has dealt with phased-array beamforming [40, 41], generalized structures for interference cancellation [42], and circuit implementations [43, 44]. Applications of the idea to massive MIMO were first presented in [45, 46], and more recently algorithms have been developed

for channel estimation [47] and transmit precoding using $\Sigma\Delta$ DACs [48].

1.1 Contributions

In this dissertation, we study the performance of massive MIMO systems with mixed-ADC and spatial $\Sigma\Delta$ architecture. We propose an optimal architecture for mixed-ADC MRC massive MIMO systems under energy constraint. In addition, we examine the channel estimation performance and the resulting uplink SE of mixed-ADC architectures with and without round-robin training, and compare them with implementations that employ uniform ADC quantization across all antennas. The main goals are to determine when, if at all, the benefits of using the round-robin approach with ADC/antenna switching outweigh the cost of increasing the training overhead, and furthermore to examine the question of whether or not one should employ a mixed-ADC architecture in the first place.

In addition, we adopt the oversampling and $\Sigma\Delta$ quantization approach in the time domain signal processing into spatial domain by introducing spatial $\Sigma\Delta$ architecture.

More precisely, the contributions of the dissertation can be summarized as follows:

- We find the optimal distribution of high-resolution and one-bit ADCs in a base station with energy constraint to maximize the spectral efficiency.
- We present an extension of the round-robin training approach that incorporates both high-resolution and one-bit measurements for the channel estimation.
- We use the Bussgang decomposition [49] to develop a linear minimum mean-squared error (LMMSE) channel estimator based on the combined round-robin measurements and we derive a closed-form expression for the resulting mean-squared error (MSE).

- We perform a spectral efficiency analysis of the mixed-ADC implementation for the MRC and ZF receivers, and obtain expressions for a lower bound on the SE that takes into account the channel estimation error and the loss of efficiency due to the round-robin training.

- We analyze the possible SE improvement that can be achieved by using an antenna selection algorithm that connects the high-resolution ADCs to the subset of antennas with the highest channel gain.

- We propose the appropriate design of spatial $\Sigma\Delta$ architecture by applying a scalar version of Bussgang approach. The results of the analysis indicate the significant gain of the $\Sigma\Delta$ approach compared with standard one-bit quantization for users that lie in the angular sector where the shaped quantization error spectrum is low.

- We propose *spatial feedback beamformer (FBB)* $\Sigma\Delta$ *architecture* to alleviate the impact of strong interference.

- We study the impact of mutual coupling on the performance of mixed-ADC and spatial $\Sigma\Delta$ architecture.

1.2 Organization

The following chapters cover the above contributions. Chapter 2 considers spectral efficiency of mixed-ADC MRC massive MIMO systems under energy constraints. We study the sum SE of a mixed-ADC massive MIMO system with MRC detection while taking into account channel estimation error. We derive a closed-form expression for the SE of a mixed-ADC massive MIMO system. This expression gives insight into how mixing one-bit and high-resolution ADCs affects the sum SE of a massive MIMO system. Then, we formulate an optimization problem to determine the best distribution of full and one-bit resolution ADCs

that maximizes the sum SE. Interestingly, the analytical and numerical analyses show that, in realistic scenarios, using a large number of antennas with only one-bit ADCs is the optimal approach.

In Chapter 3, we study the performance of mixed-ADC architecture with round-robin training and MRC/ZF receiver. Using an accurate model for the quantization noise based on the Bussgang decomposition, we propose and study the performance of a true mixed-ADC approach that uses both high-resolution and one-bit ADCs jointly in estimating the channel. It is shown that the mixed-ADC architecture improves the SE for both low and high SNRs depending on the length of the channel coherence interval. We also discuss some of the issues related to implementing an ADC switch or multiplexer in hardware that allows different ADCs to be assigned to different antennas.

In Chapter 4, we study the uplink spectral efficiency of a massive MIMO base station (BS) that employs one-bit spatial $\Sigma\Delta$ quantization, and compare it with the performance achievable by systems with infinite resolution and standard one-bit quantization. We provide some background on temporal $\Sigma\Delta$ modulation and introduce the spatial $\Sigma\Delta$ architecture. We develop an equivalent linear model and characterize this architecture. The results of the analysis indicate the significant gain of the $\Sigma\Delta$ approach compared with standard one-bit quantization for users that lie in the angular sector where the shaped quantization error spectrum is low.

The impact of mutual coupling on the mixed-ADC and $\Sigma\Delta$ architectures is proposed in Chapter 5. Finally, Chapter 6 concludes the thesis.

Notation: We use boldface letters to denote vectors, and capitals to denote matrices. The (i, j)-th element of matrix $\boldsymbol{A}$ and the i-th element of vector $\boldsymbol{a}$ are denoted by $[\boldsymbol{A}]_{ij}$ and a_i, respectively. The symbols $(.)^*$, $(.)^T$, $(.)^H$, $(.)^\dagger$, $(.)^\star$, and $\odot$ represent conjugate, transpose, conjugate transpose, pseudo-inverse, optimal value, and Hadamard product, respectively. A

circularly-symmetric complex Gaussian (CSCG) random vector with zero mean and covariance matrix $\boldsymbol{R}$ is denoted $\boldsymbol{n} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R})$. $\boldsymbol{R_{uv}}$ denotes crosscorrelation of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. $\mathrm{Ci}(x) \triangleq \gamma + \log(x) + \int_0^x \frac{\cos(t)-1}{t} dt$, and $\mathrm{Si}(x) \triangleq \int_0^x \frac{\sin(t)}{t} dt$ denote cosine and sine integrals where γ is the Euler-Mascheroni constant. The symbol $\|.\|$ represents the Euclidean norm. The operator $vec(.)$ vectorizes a matrix and $vec^{-1}(.)$ performs the reverse operation. The identity matrix is denoted by $\boldsymbol{I}$, vector of all ones by $\boldsymbol{1}$, the expectation operator by $\mathbb{E}[.]$, and the variance operator by $\mathcal{V}[.]$. We use $\mathrm{diag}(\boldsymbol{C})$, $\mathrm{diag}(\boldsymbol{x})$, and $\mathrm{diag}(x_1, \cdots, x_M)$ as the diagonal matrix formed from the diagonal entries of the square matrix $\boldsymbol{C}$, elements of vector $\boldsymbol{x}$, and scalars $x_1, \cdots, x_M$, respectively. For a complex value, $x = x_r + jx_i$, we define $x_r = \mathfrak{Re}[x]$, $x_i = \mathfrak{Im}[x]$, and $\arcsin(x) \triangleq \arcsin(x_r) + j\arcsin(x_i)$

Chapter 2

Spectral Efficiency under Energy Constraint for Mixed-ADC MRC Massive MIMO

To provide a compromise between SE loss and power consumption, so-called *mixed-ADC* architectures have been proposed [24, 26, 27, 29]. In a mixed-ADC BS, a portion of the antenna RF chains are connected to high-resolution ADCs, while the remainder are connected to one-bit ADCs. Prior work has shown that mixed-ADC architectures can provide significant gains in SE with considerably reduced power consumption. In [24, 27], the spectral and energy efficiency trade-off of the mixed-ADC massive MIMO uplink is investigated, for both cases involving perfect and estimated channel state information (CSI). Both [24, 27] assumed a "round-robin" approach to obtain the estimated CSI, in which the high-resolution ADCs are connected sequentially to each antenna over several training intervals to estimate the channel. This significantly increases the training overhead and reduces the ultimate spectral efficiency, but allows one to determine the "strongest" channels to select for the high-resolution ADCs in the data transmission phase.

In this chapter, we take a different approach and assume that the high-resolution ADCs are connected to an arbitrary subset of the antennas to reduce the training overhead (no round-robin training), and we maximize the SE subject to a fixed power budget. We show that for a maximum ratio combiner, the optimal architecture uses either all one-bit or all full-resolution ADCs, depending on the relative power consumed by each type of RF chain. We further show that for typical power consumption parameters, an all-one-bit system is likely the optimal approach in practice.

2.1 System Model

We model the uplink of a single-cell multi-user MIMO system consisting of K single-antenna users and a BS equipped with M antennas. The $M \times 1$ signal received at the BS from the K users is given by

$$\boldsymbol{y} = \sqrt{p}\mathbf{H}\boldsymbol{x} + \boldsymbol{n}, \tag{2.1}$$

where p represents the transmit power, $\mathbf{H} \in \mathbb{C}^{M \times K}$ is the channel matrix whose elements are distributed independently and identically as $\mathcal{CN}(0,1)$, the symbol vector is $\boldsymbol{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$, and the additive noise is $\boldsymbol{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_M)$.

We consider a mixed-ADC architecture at the BS in which $M_0 = \mu M$ antennas are connected to high-resolution ADCs while $M_1 = (1-\mu)M$ antennas are fed to one-bit ADCs. Since M_0 and M_1 are integers, μ is restricted to certain rational values, but for the moment we ignore this constraint. As a result, by partitioning the channel matrix $\mathbf{H}$, we can rewrite (2.1) as

$$\begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} = \sqrt{p} \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \boldsymbol{x} + \boldsymbol{n}, \tag{2.2}$$

where $\mathbf{H}_0 \in \mathbb{C}^{M_0 \times K}$ ($\mathbf{H}_1 \in \mathbb{C}^{M_1 \times K}$) contains the channel coefficients from the users to the M_0 (M_1) antennas connected to high-resolution (one-bit) ADCs. Therefore, the received signal at the BS after one-bit quantization is

$$\boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{y}_0 \\ Q(\boldsymbol{y}_1) \end{bmatrix}, \tag{2.3}$$

where the element-wise one-bit quantization operation $Q(\cdot)$ replaces each input entry with the quantized value $\frac{1}{\sqrt{2}}(\pm 1 \pm j)$, depending on the sign of the real and imaginary parts. We can represent the non-linear quantization using a statistically equivalent linear model based on the Bussgang decomposition as (see [15] for details)

$$Q(\boldsymbol{y}_1) = \mathbf{A}\boldsymbol{y}_1 + \boldsymbol{q}, \tag{2.4}$$

where $\mathbf{A}$ is a linear gain and $\boldsymbol{q}$ denotes the statistically equivalent quantization noise which is uncorrelated with $\boldsymbol{y}_1$. As explained in [15], the Bussgang approach provides a more accurate analysis than the AQNM in the case of one-bit quantization.

We assume a block-fading model where the channel remains constant in a coherence interval of length T symbols and changes independently between different intervals. At the beginning of each coherence interval, the users send their η-tuple mutually orthogonal pilot sequences ($K \leq \eta \leq T$) to the BS for channel estimation. The remaining $T - \eta$ symbols are dedicated to uplink data transmission.

2.1.1 Training Phase

The pilot sequences are stacked into an $\eta \times K$ matrix $\mathbf{\Phi}$, where the kth column of $\mathbf{\Phi}$, $\boldsymbol{\phi}_k$, is the kth user's pilot sequence and $\mathbf{\Phi}^H\mathbf{\Phi} = \boldsymbol{I}_K$. Therefore, the $M \times \eta$ received signal at the BS

in the training phase is

$$\mathbf{Y}_t = \begin{bmatrix} \mathbf{Y}_{t_0} \\ \mathbf{Y}_{t_1} \end{bmatrix} = \sqrt{\eta p} \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \boldsymbol{\Phi}^{\boldsymbol{T}} + \begin{bmatrix} \mathbf{N}_0 \\ \mathbf{N}_1 \end{bmatrix}, \tag{2.5}$$

where $\mathbf{N}_0$ ($\mathbf{N}_1$) is an $M_0 \times \eta$ ($M_1 \times \eta$) matrix with i.i.d. $\mathcal{CN}(0, \sigma_n^2)$ elements. Although it has been shown that $\eta > K$ can improve performance in one-bit massive MIMO systems [15], we assume $\eta = K$ in the sequel for analytical simplification. To apply the Bussgang decomposition as in (2.4), we first vectorize (2.5)

$$\begin{bmatrix} vec\left(\mathbf{Y}_{t_0}\right) \\ vec\left(\mathbf{Y}_{t_1}\right) \end{bmatrix} = \sqrt{Kp} \begin{bmatrix} \tilde{\boldsymbol{\Phi}}_0 vec\left(\mathbf{H}_0\right) \\ \tilde{\boldsymbol{\Phi}}_1 vec\left(\mathbf{H}_1\right) \end{bmatrix} + \begin{bmatrix} vec\left(\mathbf{N}_0\right) \\ vec\left(\mathbf{N}_1\right) \end{bmatrix}, \tag{2.6}$$

where $\tilde{\boldsymbol{\Phi}}_i = \boldsymbol{\Phi} \otimes \mathbf{I}_{M_i}$, $i \in \{0, 1\}$. After quantization, from (2.4) we have

$$\begin{bmatrix} \tilde{\mathbf{y}}_{t_0} \\ \tilde{\mathbf{y}}_{t_1} \end{bmatrix} = \sqrt{Kp} \begin{bmatrix} \tilde{\boldsymbol{\Phi}}_0 vec\left(\mathbf{H}_0\right) \\ \mathbf{A}_t \tilde{\boldsymbol{\Phi}}_1 vec\left(\mathbf{H}_1\right) \end{bmatrix} + \begin{bmatrix} vec\left(\mathbf{N}_0\right) \\ \mathbf{A}_t vec\left(\mathbf{N}_1\right) + \boldsymbol{q}_t \end{bmatrix}, \tag{2.7}$$

with $\mathbf{A}_t = \sqrt{\frac{2}{\pi} \frac{1}{Kp+\sigma_n^2}} \mathbf{I}_{M_1 K}$ [15]. Accordingly, the linear minimum mean squared error (LMMSE) estimate of $\mathbf{H}$ for a mixed-ADC architecture is [15, 50]

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{H}}_0 \\ \hat{\mathbf{H}}_1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{Kp}}{Kp+\sigma_n^2} \; vec^{-1}\left(\tilde{\boldsymbol{\Phi}}_0^H \tilde{\mathbf{y}}_{t_0}\right) \\ \sqrt{\frac{2}{\pi} \frac{Kp}{Kp+\sigma_n^2}} \; vec^{-1}\left(\tilde{\boldsymbol{\Phi}}_1^H \tilde{\mathbf{y}}_{t_1}\right) \end{bmatrix}. \tag{2.8}$$

This estimation leads to

$$\sigma_0^2 = \frac{1}{1 + \frac{\sigma_n^2}{pK}} \quad \sigma_1^2 = \frac{2}{\pi} \frac{1}{1 + \frac{\sigma_n^2}{pK}} \tag{2.9}$$

where σ_0^2 and σ_1^2 are the variances of the independent zero mean elements of $\hat{\mathbf{H}}_0$ and $\hat{\mathbf{H}}_1$, respectively.

2.1.2 Data Transmission Phase

In the data transmission phase, using (2.3)-(2.4) the received signal after quantization is [50]

$$\begin{bmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{bmatrix} = \sqrt{p} \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{A}_d \mathbf{H}_1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} \boldsymbol{n}_0 \\ \mathbf{A}_d \boldsymbol{n}_1 + \boldsymbol{q}_d \end{bmatrix}, \tag{2.10}$$

where $\mathbf{A}_d = \sqrt{\frac{2}{\pi} \frac{1}{Kp+\sigma_n^2}} \mathbf{I}_{M_1} = \alpha_d \mathbf{I}_{M_1}$ and $\boldsymbol{q}_d$ is the statistically equivalent quantization noise with covariance matrix $\mathbf{C}_{\boldsymbol{q}_d}$. For data detection, the BS employs the MRC detector $\boldsymbol{W} \in \mathbb{C}^{M \times K}$ assuming that the channel estimate is the true channel. Note that the quantization model considered in (2.4) does not preserve the power of the input of the quantizer since the output is forced to be ± 1. Thus we implement MRC as follows to offset this effect:

$$\mathbf{W} = \begin{bmatrix} \hat{\mathbf{H}}_0 \\ \mathbf{A}_d^{-1} \hat{\mathbf{H}}_1 \end{bmatrix}. \tag{2.11}$$

Therefore, the resulting signal at the BS is

$$\hat{\boldsymbol{x}} = \mathbf{W}^H \begin{bmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{bmatrix} = \sqrt{p} \begin{bmatrix} \hat{\mathbf{H}}_0 \\ \hat{\mathbf{H}}_1 \end{bmatrix}^H \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} \hat{\mathbf{H}}_0 \\ \hat{\mathbf{H}}_1 \end{bmatrix}^H \begin{bmatrix} \boldsymbol{n}_0 \\ \boldsymbol{n}_1 + \mathbf{A}_d^{-1} \boldsymbol{q}_d \end{bmatrix}. \tag{2.12}$$

Note that premultiplication by $\mathbf{A}_d^{-1}$ only makes the quantization a power-preserving operation and does not alter the information of the quantizer output. Therefore, the kth element of $\hat{\boldsymbol{x}}$ is

$$\hat{x}_k = \sqrt{p} \hat{\boldsymbol{h}}_k^H \boldsymbol{h}_k x_k + \sum_{i=1, i \neq k}^{K} \sqrt{p} \hat{\boldsymbol{h}}_k^H \boldsymbol{h}_i x_i + \hat{\boldsymbol{h}}_k^H \boldsymbol{n} + \hat{\boldsymbol{h}}_{1_k}^H \mathbf{A}_d^{-1} \boldsymbol{q}_d, \tag{2.13}$$

where $\hat{\boldsymbol{h}}_k$, $\boldsymbol{h}_k$, and $\hat{\boldsymbol{h}}_{1_k}$ are the kth column of $\hat{\mathbf{H}}$, $\mathbf{H}$, and $\hat{\mathbf{H}}_1$, respectively. The BS treats $\hat{\boldsymbol{h}}_k^H \boldsymbol{h}_k$ as the desired channel and the other terms of (5.20) as worst-case Gaussian noise when decoding the signal. Consequently, a lower bound for the ergodic achievable SE at the kth user is [51]

$$\mathcal{S}_k = \mathcal{R}\left(\mathrm{SQINR}_k\right), \tag{2.14}$$

where $\mathcal{R}(\theta) \triangleq (1-\eta/T)\log_2(1+\theta)$ and SQINR_k is the effective signal-to-quantization-interference-and-noise ratio at the kth user given by

$$\mathrm{SQINR}_k = \frac{p\left|\mathbb{E}\left\{\hat{\boldsymbol{h}}_k^H \boldsymbol{h}_k\right\}\right|^2}{p\sum_{i=1}^{K}\mathbb{E}\left\{\left|\hat{\boldsymbol{h}}_k^H \boldsymbol{h}_i\right|^2\right\} - p\left|\mathbb{E}\left\{\hat{\boldsymbol{h}}_k^H \boldsymbol{h}_k\right\}\right|^2 + \sigma_n^2\mathbb{E}\left\{\|\hat{\boldsymbol{h}}_k\|^2\right\} + \alpha_d^{-2}\mathbb{E}\left\{\hat{\boldsymbol{h}}_{1_k}^H \mathbf{C}_{\boldsymbol{q}_d}\hat{\boldsymbol{h}}_{1_k}\right\}}. \tag{2.15}$$

The following theorem presents an expression for the sum SE of a mixed-ADC architecture with MRC detection.

Theorem 2.1. *The SE of the k-th user in a mixed-ADC massive MIMO system with MRC detection is*

$$\mathcal{S}_k = \mathcal{R}\left(\frac{M}{K+\frac{\sigma_n^2}{p}}\frac{\left(\mu\sigma_0^2+(1-\mu)\sigma_1^2\right)^2}{\mu\sigma_0^2+(1-\mu)\frac{\pi}{2}\sigma_1^2}\right). \tag{2.16}$$

Proof. The proof follows by calculating the expected values in (2.15) and using (2.9), following the reasoning in [15]. ∎

In the next section, we find the optimal values of M and μ by considering the power consumption at the BS to maximize the sum SE in (2.16). We will denote the fixed power consumed at the BS due to the local oscillator, baseband processors, etc, by $\mathcal{P}_{\mathrm{FIX}}$ and the power consumed by a single RF chain with high-resolution and one-bit ADCs by $\mathcal{P}_0$ and $\mathcal{P}_1$,

respectively. Hence, the total consumed power at the BS is

$$\mathcal{P}_{\text{tot}} = \mathcal{P}_{\text{FIX}} + M_0\mathcal{P}_0 + M_1\mathcal{P}_1 = \mathcal{P}_{\text{FIX}} + M\Big(\mu\mathcal{P}_0 + (1-\mu)\,\mathcal{P}_1\Big). \tag{2.17}$$

Since $\mathcal{P}_0 \gg \mathcal{P}_1$, a small value for μ means a larger number of antennas and hence an increase in the MIMO spatial multiplexing gain, but a loss of SE due to increased quantization noise. Increasing μ reduces the SE loss due to quantization, but also decreases the size of the antenna array. The problem of finding the optimal value of μ that maximizes the SE for a given total power constraint $\mathcal{P}$ is addressed next.

2.2 Optimal Mixed-ADC Design

In this section, we determine what fraction of high-resolution/one-bit ADCs should be installed at the BS to maximize the sum SE, $\mathcal{S} = \sum_{k=1}^{K} \mathcal{S}_k$, subject to a given power budget $\mathcal{P}$. The optimization problem is expressed as

$$\mathfrak{P}: \begin{array}{ll} \underset{\mu}{\text{maximize}} & \mathcal{S} \\ \text{subject to} & \mathcal{P}_{\text{tot}} \le \mathcal{P} \\ & 0 \le \mu \le 1. \end{array} \tag{2.18}$$

The following theorem provides the optimal solution for μ, assuming μ can take on any value in the interval $[0, 1]$. Once $\mu^\star$ is obtained, M_0 and M_1 are taken to be the integers closest to $\mu^\star M$ and $(1-\mu^\star)M$ that result in $\mathcal{P}_{\text{tot}} \le \mathcal{P}$.

Theorem 2.2. *Assume $\mathcal{P}_0 = \rho\mathcal{P}_1$ and denote $\rho_{\text{th}} = \frac{\pi^2}{4}$. Then, the maximum spectral*

efficiency is achieved by

$$\mu^\star = \begin{matrix} 0 & \rho \geq \rho_{\text{th}}, \\ 1 & \rho < \rho_{\text{th}}. \end{matrix} \tag{2.19}$$

Proof. Denote the argument of $\mathcal{R}$ in (2.16) as $\gamma(\mu)$. Since $\mathcal{R}(\gamma)$ is a continuous and strictly increasing function of γ, the first inequality constraint becomes equality, and the optimization problem is equivalent to

$$\mathfrak{P}: \begin{matrix} \underset{\mu}{\text{maximize}} & \gamma(\mu) \\ \text{subject to} & 0 \leq \mu \leq 1. \end{matrix} \tag{2.20}$$

The proof is carried out by analyzing the behavior of the first derivative of

$$\bar{\gamma}(\mu) = \frac{\left(\mu\sigma_0^2 + (1-\mu)\sigma_1^2\right)^2}{\left(\mu\mathcal{P}_0 + (1-\mu)\mathcal{P}_1\right)\left(\mu\sigma_0^2 + (1-\mu)\frac{\pi}{2}\sigma_1^2\right)}, \tag{2.21}$$

in $0 \leq \mu \leq 1$ for different values of $\rho > 1$. The first derivative of $\bar{\gamma}$ can be written as

$$\frac{d\bar{\gamma}}{d\mu} = f(\mu) = \beta\Big((a+b)\mu - a\Big), \tag{2.22}$$

where β is a positive constant in $0 \leq \mu \leq 1$, $a = \frac{\pi}{2}\rho\sigma_1^4 + (1-\pi)\sigma_0^2\sigma_1^2$ and $b = (\frac{\pi}{2}-2)\rho\sigma_0^2\sigma_1^2 + \sigma_0^4$.

- For $1 < \rho \leq \rho_0 = \frac{(\pi-1)\sigma_0^2\sigma_1^2 - \sigma_0^4}{\sigma_1^2\left((\frac{\pi}{2}-2)\sigma_0^2 + \frac{\pi}{2}\sigma_1^2\right)}$, we have $a < 0$, $b > 0$, and $a + b < 0$. Thus, f is a line with negative slope and a root at $\mu > 1$ which amounts to $\frac{d\bar{\gamma}}{d\mu} > 0$ and hence $\mu^\star = 1$.
- In the interval of $\rho_0 < \rho \leq \rho_1 = (2 - 2/\pi)\frac{\sigma_0^2}{\sigma_1^2}$, we have $a \leq 0$, $b > 0$, and $a + b > 0$. As a result, f has a positive slope and a root at $\mu < 0$. Then, $\frac{d\bar{\gamma}}{d\mu} > 0$, i.e., $\mu^\star = 1$.
- For $\rho_1 < \rho < \rho_2 = \frac{1}{2-\pi/2}\frac{\sigma_0^2}{\sigma_1^2}$, $a > 0$ and $b > 0$. Consequently, f has a positive slope with a root at $0 < \mu < 1$ which means either $\mu^\star = 0$ or $\mu^\star = 1$ in this interval. Since, $\bar{\gamma}(0) = \frac{2}{\pi}\frac{\sigma_1^2}{\mathcal{P}_1}$ and $\bar{\gamma}(1) = \frac{\sigma_0^2}{\rho\mathcal{P}_1}$, Then, $\mu^\star = 1$ if $\rho < \frac{\pi}{2}\frac{\sigma_0^2}{\sigma_1^2}$ and $\mu^\star = 0$ if $\rho > \frac{\pi}{2}\frac{\sigma_0^2}{\sigma_1^2}$.

- For $\rho \geq \rho_2$, it can be shown that $b \leq 0$, $a > 0$ and $a + b > 0$. Therefore f has a positive slope with root at $\mu \geq 1$, which leads to $\frac{d\bar{\gamma}}{d\mu} \leq 0$. Thus, $\bar{\gamma}(\mu)$ is a strictly decreasing function and $\mu^\star = 0$.

Combining the above four intervals together and $\frac{\pi}{2}\frac{\sigma_0^2}{\sigma_1^2} = \frac{\pi^2}{4}$ from (2.9) results in (2.19) and the proof is complete. ■

Theorem 2.2 states that, for a given power budget, the optimal approach is to deploy the BS with either purely high-resolution or one-bit ADCs. The proof implies that when $\rho = \rho_{\text{th}}$, the performance of a system with M_0 antennas connected to high-resolution ADCs is the same as that of with $M_1 = \rho_{\text{th}} M_0$ antennas connected to one-bit ADCs. This is consistent with the results of [15, 50], where it was shown that for an MRC receiver, the SE loss due to one-bit quantization can be offset by deploying $\pi^2/4$ more antennas at the BS.

The question of whether to use all one-bit or all high-resolution ADCs boils down to whether or not the cost of a high-resolution RF chain is more than $\rho_{\text{th}} = \pi^2/4 \simeq 2.5$ times that of an one-bit RF chain. In the next section we show that this will likely be the case in realistic scenarios, and thus a BS equipped with only one-bit ADCs is the most cost-effective solution.

2.3 Numerical Results

To find a typical value for ρ, we use as an example the RF power consumption model considered in [27]:

$$P_{\text{BS}} = P_{\text{LO}} + P_{\text{BB}} + M\left(P_{\text{LNA}} + P_{\text{H}} + 2P_{\text{M}}\right) + 2M_0\left(P_{\text{AGC}} + P_{\text{ADC}}^H\right) + 2M_1 P_{\text{ADC}}^L, \quad (2.23)$$

where P_{LO}, P_{BB}, P_{LNA}, P_{H}, P_{M}, P_{AGC}, P_{ADC}^H, and P_{ADC}^L denote power consumption of the local oscillator, baseband processing, low noise amplifier, $\pi/2$ hybrid and local oscillator buffer,

mixer, adaptive gain controller, high-resolution ADCs, and one-bit ADCs, respectively. To employ the result of Theorem 2.2, we can re-write (2.23) in the following form

$$P_{\mathrm{BS}} - P_{\mathrm{LO}} - P_{\mathrm{BB}} = \\ M_0 \left(P_{\mathrm{LNA}} + P_{\mathrm{H}} + 2P_{\mathrm{M}} + 2P_{\mathrm{AGC}} + 2P_{\mathrm{ADC}}^{H} \right) + M_1 \left(P_{\mathrm{LNA}} + P_{\mathrm{H}} + 2P_{\mathrm{M}} + 2P_{\mathrm{ADC}}^{L} \right). \quad (2.24)$$

As a result,

$$\rho = \frac{\mathcal{P}_0}{\mathcal{P}_1} = \frac{P_{\mathrm{LNA}} + P_{\mathrm{H}} + 2P_{\mathrm{M}} + 2P_{\mathrm{AGC}} + 2P_{\mathrm{ADC}}^{H}}{P_{\mathrm{LNA}} + P_{\mathrm{H}} + 2P_{\mathrm{M}} + 2P_{\mathrm{ADC}}^{L}} \sim 6.8, \quad (2.25)$$

where we have used the following power consumption values from [27, 12]: $P_{\mathrm{LO}} = 22.5$ mW, $P_{\mathrm{BB}} = 200$ mW, $P_{\mathrm{LNA}} = 5.4$ mW, $P_{\mathrm{H}} = 3$ mW, $P_{\mathrm{M}} = 0.3$ mW, $P_{\mathrm{AGC}} = 2$ mW, $P_{\mathrm{ADC}}^{H} = 25.6$ mW, and $P_{\mathrm{ADC}}^{L} = 0.2$ mW. It is apparent from (2.25) that $\rho > \rho_{\mathrm{th}} \sim 2.5$. The model in (2.23) may even be pessimistic for the one-bit case, since one-bit quantization can simplify the LNA and subsequent baseband processing. Thus, we observe that using only one-bit ADCs likely provides the highest SE for a given power budget assuming MRC processing. It is worthwhile to note that this conclusion may change for different power consumption models. However, every model can be written in the form of (2.17) and, hence, our result in Theorem 2.19 is general and can be evaluated based on different models for power consumption. In addition, this conclusion may change for different approaches such as using a zero-forcing receiver or antenna selection for the high-resolution ADCs, as these approaches tend to better exploit such ADCs, albeit with higher cost, complexity and training overhead.

Fig. 2.1 shows the sum SE with respect to SNR $\triangleq p/\sigma_n^2$ for various values of μ under the power constraint $P_{\mathrm{BS}} - P_{\mathrm{LO}} - P_{\mathrm{BB}} = 200\mathcal{P}_1$, and assuming $\eta = K = 10$ and $T = 200$. As expected, the sum SE is maximized when the BS is equipped with a large number of antennas connected to one-bit quantizers. Note that the simulated rates (symbols) match the theoretical expression in (2.16) (lines) quite accurately.

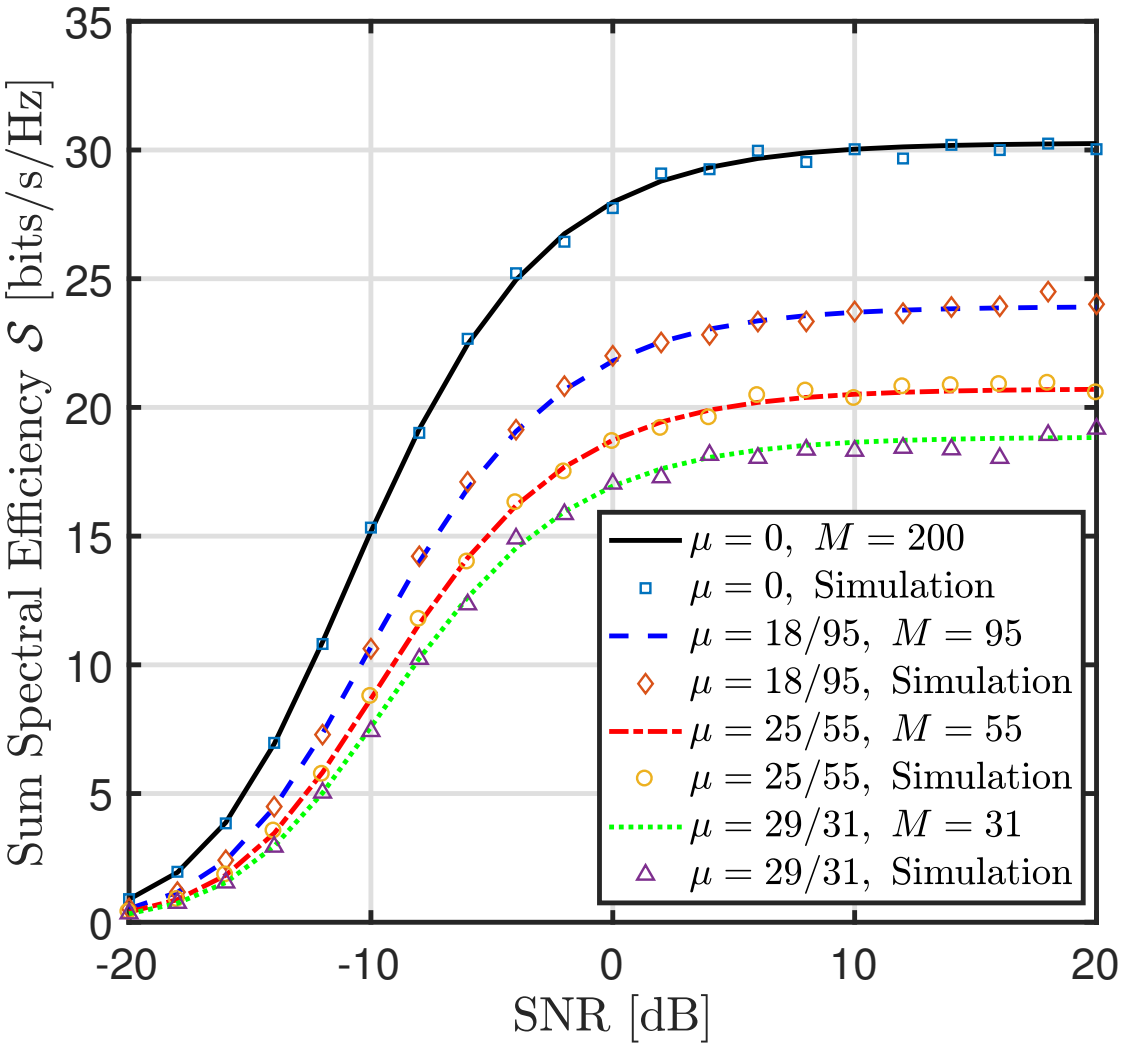

Figure 2.1: Sum spectral efficiency $\mathcal{S}$ versus the average SNR for $P_{\text{BS}} - P_{\text{LO}} - P_{\text{BB}} = 200\mathcal{P}_1$, $K = 10$, and $T = 200$. Solid lines denote the theoritical expression in (2.16) and symbols denote Monte-Carlo simulation for 1000 trials.

2.4 Summary

In this chapter, we have studied the spectral efficiency of the mixed-ADC massive MIMO uplink under energy constraint and assuming MRC processing. We derived a closed form expression for the SE assuming the channel is estimated using both high-resolution and one-bit ADCs, without using the training-intensive round-robin approach. We then maximized the SE under a constraint on the power budget, and showed that in fact the mixed-ADC approach is not optimal. Typical power consumption models indicate that a system with all one-bit ADCs provides the highest SE for a given uplink power budget.

Chapter 3

Spectral Efficiency of Mixed-ADC Massive MIMO

In this chapter, we focus on channel estimation and spectral efficiency performance of mixed-ADC massive MIMO systems. This architecture is depicted in Fig. 3.1. Most existing work in the mixed-ADC massive MIMO literature has assumed either perfect channel state information (CSI) or imperfect CSI with "round-robin" training. In the round-robin training approach [24, 25, 27], the training data is repeated several times and the high-resolution ADCs are switched among the RF chains so that every antenna can have a "clean" snapshot of the pilots for channel estimation. This obviously requires a larger portion of the coherence interval to be devoted to training rather than data transmission. More precisely, for M antennas and N pairs of high-resolution ADCs, M/N pilot signals are required in the single-user scenario to estimate all M channel coefficients with high-resolution ADCs. This issue is pointed out in [24] for the single user scenario and its impact is taken into account. This training overhead will be exacerbated in the multiuser scenario where orthogonal pilot sequences should be assigned to the users. In this case, the training period becomes $(M/N)\eta$, where η represents the length of the pilot sequences (at least as large as the number of user

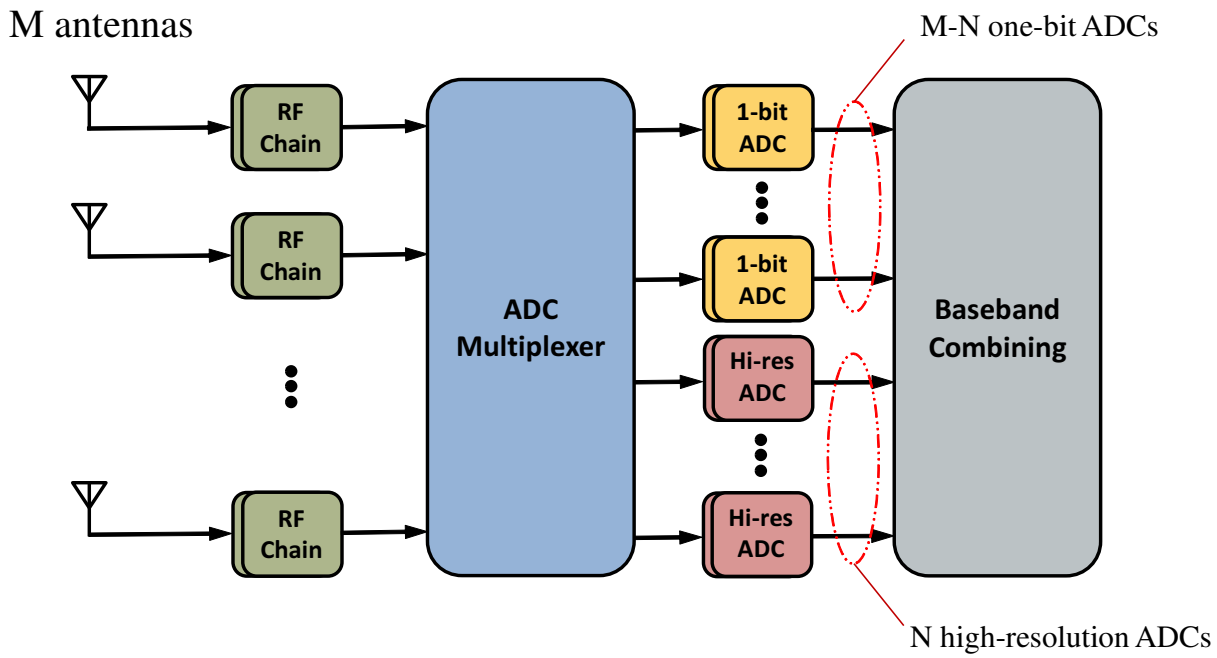

Figure 3.1: Mixed-ADC architecture.

terminals), which could be prohibitively large and may leave little room for data transmission. Hence, it is crucial to account for this fact in any SE analysis of mixed-ADC massive MIMO systems.

The round-robin training proposed in [24, 25, 27] based the channel estimate on only high-resolution observations, assuming that no data was collected from antennas during intervals when they were not connected to the high-resolution ADCs. In this chapter, we present an extension of the round-robin training approach that incorporates both high-resolution and one-bit measurements for the channel estimation. In contrast, our extension assumes that antennas collect one-bit observations and combine this data with the high-resolution samples to improve the channel estimation performance. We further illustrate the importance of using the Bussgang approach rather than the simpler additive quantization noise model in obtaining the most accurate characterization of the channel estimation performance for round-robin training. The analysis illustrates that the addition of the one-bit observations considerably improves performance at low SNR.

After estimating the channel, we study the SE performance of mixed-ADC architecture for different scenarios. We analytically derive the SE performance of the antenna selection algorithm for MRC and numerically study its performance for ZF detection, comparing against the simpler approach of assigning the high-resolution ADCs to an arbitrary fixed subset of the RF chains.

We compare the resulting SE with that achieved by mixed-ADC implementations that do not switch ADCs among the RF chains, and hence do not use round-robin training. We also compare against the SE for architectures that do not mix the ADC resolution across the array, but instead use uniform resolution with a fixed number of comparators for different array sizes. We show that, depending on the SNR, coherence interval, number of high-resolution ADCs, and the choice of the linear receiver, there are situations where each of the considered approaches shows superior performance. In particular, using uniform low-resolution ADCs is better than a mixed-ADC approach for an interference limited system. On the other hand, a mixed-ADC system, even one with round-robin training, is superior at higher SNRs when zero-forcing is used to reduce the interference.

In addition to the above contributions, we also discuss some of the issues related to implementing an ADC switch or multiplexer in hardware that allows different ADCs to be assigned to different antennas. We restrict our analysis and numerical examples to a single-carrier flat-fading scenario, although our methodology can be used in a straightforward way to extend the results to frequency-selective fading or multiple-carrier signals (*e.g.*, see our prior work in Section III.B of [15] for the SE analysis of an all-one-bit ADC system for OFDM and frequency selectivity). The reasons for focusing on the single-carrier flat-fading case are as follows: (1) the mixed-ADC assumption already makes the resulting analytical expressions quite complicated even for the simple flat-fading case, and it would be more difficult to gain insight into the problem if the expressions were further complicated; (2) the original round-robin training idea was proposed in [24] for the single-carrier flat-fading case, and thus we

analyze it under the same assumptions; (3) the main conclusions of the chapter are based on relative algorithm comparisons for the same set of assumptions, and we expect our general conclusions to remain unchanged if frequency rather than flat fading were considered; and (4) the flat fading case is still of interest in some applications, for example in a micro-cell setting with typical path-length differences of 50-100 m, the coherence bandwidth is between 3-6 MHz, which is not insignificant.

Further assumptions regarding the system model are outlined in the next section. Section 3.2 discusses channel estimation using round-robin training, and derives the LMMSE channel estimator that incorporates both the high-resolution and one-bit observations. A discussion of hardware and other practical considerations associated with using a mixed-ADC system with ADC/antenna switching is presented in Section 3.3. Section 3.4 then presents the analysis of the spectral efficiency for MRC and ZF receivers based on the imperfect channel state estimates, including an analytical performance characterization of antenna selection and architectures with uniform ADC resolution across the array. A number of numerical studies are then presented in Section 3.5 to illustrate the relative performance of the algorithms considered.

3.1 System Model

Consider the uplink of a single-cell multi-user MIMO system consisting of K single-antenna users that send their signals simultaneously to a BS equipped with M antennas. Assuming a single-carrier frequency flat channel and symbol-rate sampling , the $M \times 1$ signal received at the BS from the K users is given by

$$\boldsymbol{r} = \sum_{k=1}^{K} \sqrt{p_k} \boldsymbol{g}_k s_k + \boldsymbol{n}, \tag{3.1}$$

where p_k represents the average transmission power from the kth user, $\boldsymbol{g}_k = \sqrt{\beta_k}\boldsymbol{h}_k$ is the channel vector between the kth user and the BS where β_k models geometric attenuation and shadow fading, and $\boldsymbol{h}_k \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$ represents the fast fading and is assumed to be independent of other users' channel vectors. The symbol transmitted by the kth user is denoted by s_k where $\mathbb{E}\left\{|s_k|^2\right\} = 1$ and is drawn from a CSCG codebook independent of the other users. Finally, $\boldsymbol{n} \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I}_M\right)$ denotes additive CSCG receiver noise at the BS. The assumption of symbol-rate sampling means that the matched filter at the receiver must be implemented in the analog domain. Better performance (e.g., higher rates) could be achieved by oversampling the ADCs, particularly those with one-bit resolution.

We consider a block-fading model with coherence bandwidth W_c and coherence time T_c. In this model, each channel remains constant in a coherence interval of length $T = T_c W_c$ symbols and changes independently between different intervals. Note that T is a fixed system parameter chosen as the minimum coherence duration of all users. At the beginning of each coherence interval, the users send their η-tuple mutually orthogonal pilot sequences ($K \leq \eta \leq T$) to the BS for channel estimation. Denoting the length of the training phase as η_{eff}, the remaining $T - \eta_{\text{eff}}$ symbols are dedicated to uplink data transmission.

3.2 Training Phase

In this section, we investigate the linear minimum mean squared error (LMMSE) channel estimator for different ADC architectures at the BS. In all scenarios, the pilot sequences are drawn from an $\eta \times K$ matrix $\boldsymbol{\Phi}$, where the kth column of $\boldsymbol{\Phi}$, $\boldsymbol{\phi}_k$, is the kth user's pilot sequence and $\boldsymbol{\Phi}^H \boldsymbol{\Phi} = \boldsymbol{I}_K$. Therefore, the $M \times \eta$ received signal at the BS before quantization becomes

$$\boldsymbol{X} = \sum_{k=1}^{K} \sqrt{\eta p_k} \boldsymbol{g}_k \boldsymbol{\phi}_k^T + \boldsymbol{N}, \tag{3.2}$$

where $\boldsymbol{N}$ is an $M \times \eta$ matrix with i.i.d. $\mathcal{CN}(0, \sigma_n^2)$ elements. Since the rows of $\boldsymbol{X}$ are mutually independent due to the assumption of spatially uncorrelated Gaussian channels and noise, we can analyze them separately. As a result, we will focus on the mth row of $\boldsymbol{X}$ which is

$$\boldsymbol{x}_m^T = \sum_{k=1}^{K} \sqrt{\eta p_k} g_{mk} \boldsymbol{\phi}_k^T + \boldsymbol{n}_m^T, \tag{3.3}$$

where g_{mk} is the mth element of the kth user channel vector, $\boldsymbol{g}_k$, and $\boldsymbol{n}_m^T$ is the mth row of $\boldsymbol{N}$. Since the analysis is not dependent on m, hereafter we drop this subscript and denote the received signal at the mth antenna by $\boldsymbol{x}$.

3.2.1 Estimation Using One-Bit Quantized Observations

In this subsection, to have a benchmark for comparison purposes, we consider the case in which all antennas at the BS are connected to one-bit ADCs. The received signal $\boldsymbol{x}^T$ after quantization by one-bit ADCs can be written as

$$\boldsymbol{y}_t^T = Q\left(\boldsymbol{x}^T\right), \tag{3.4}$$

where the element-wise one-bit quantization operation $Q(\cdot)$ replaces each input entry with the quantized value $\frac{1}{\sqrt{2}}(\pm 1 \pm j)$, depending on the sign of the real and imaginary parts. According to the Bussgang decomposition [49], the following linear representation of the quantization can be employed [15]:

$$Q\left(\boldsymbol{x}^T\right) = \sqrt{\frac{2}{\pi}} \boldsymbol{x}^T \mathbf{D}_{\boldsymbol{x}}^{-\frac{1}{2}} + \boldsymbol{q}_t^T, \tag{3.5}$$

where $\mathbf{D}_{\boldsymbol{x}} = \text{diag}\{\mathbf{C}_{\boldsymbol{x}}\}$ and $\mathbf{C}_{\boldsymbol{x}}$ denotes autocorrelation matrix of $\boldsymbol{x}$, which can be calculated as

$$\mathbf{C}_{\boldsymbol{x}} = \sum_{k=1}^{K} \eta p_k \beta_k \boldsymbol{\phi}_k^* \boldsymbol{\phi}_k^T + \sigma_n^2 \mathbf{I}_\eta. \tag{3.6}$$

In addition, $\boldsymbol{q}_t$ represents quantization noise which is uncorrelated with $\boldsymbol{x}$ and its autocorrelation matrix can be derived based on the *arcsine law* as [52]

$$\mathbf{C}_{\boldsymbol{q}_t} = \frac{2}{\pi}\arcsin\{\mathbf{D}_{\boldsymbol{x}}^{-\frac{1}{2}}\mathbf{C}_{\boldsymbol{x}}\mathbf{D}_{\boldsymbol{x}}^{-\frac{1}{2}}\} - \frac{2}{\pi}\mathbf{D}_{\boldsymbol{x}}^{-\frac{1}{2}}\mathbf{C}_{\boldsymbol{x}}\mathbf{D}_{\boldsymbol{x}}^{-\frac{1}{2}}. \tag{3.7}$$

Much of the existing work on massive MIMO systems with low-resolution ADCs employs the simple additive quantization noise model (AQNM) for their analysis [21, 22, 23], [26, 27, 28, 29, 30, 31], [53] which is valid only for low SNRs and does not capture the correlation among the elements of $\boldsymbol{q}_t$, which turns out to be of crucial importance in our analysis. Hence, we consider the Bussgang decomposition instead and will show its effect on the system performance analysis. Stacking the rows of (3.5) into a matrix, the one-bit quantized observation at the BS becomes

$$\mathbf{Y} = \sqrt{\frac{2}{\pi}}\mathbf{X}\mathbf{D}_{\boldsymbol{x}}^{-\frac{1}{2}} + \mathbf{Q}, \tag{3.8}$$

where $\mathbf{Q}$ is an $M \times \eta$ matrix whose mth row is $\boldsymbol{q}_t^T$. The LMMSE estimate of the channel $\boldsymbol{G} = [\boldsymbol{g}_1, ..., \boldsymbol{g}_K]$ based on just one-bit quantized observations (3.8) is given in the following theorem.

Theorem 3.1. *The LMMSE estimate of the k-th user channel, $\boldsymbol{g}_k$, given the one-bit quantized observations $\mathbf{Y}$ is [15]*

$$\hat{\boldsymbol{g}}_k = \frac{\beta_k}{\beta_k + \sigma_{w_k}^2}\sqrt{\frac{1}{\eta p_k}}\mathbf{Y}\bar{\boldsymbol{\phi}}_k^*, \tag{3.9}$$

where

$$\bar{\boldsymbol{\phi}}_k \triangleq \sqrt{\frac{\pi}{2}} \mathbf{D}_{\boldsymbol{x}}^{\frac{1}{2}} \boldsymbol{\phi}_k \tag{3.10}$$

$$\sigma_{w_k}^2 = \frac{1}{\eta p_k} \left(\sigma_n^2 + \bar{\boldsymbol{\phi}}_k^T \mathbf{C}_{\boldsymbol{q}_t} \bar{\boldsymbol{\phi}}_k^* \right). \tag{3.11}$$

Define the channel estimation error $\boldsymbol{\varepsilon} \triangleq \hat{\boldsymbol{g}}_k - \boldsymbol{g}_k$. *Then we have*

$$\sigma_{\hat{g}_k}^2 = \frac{\beta_k^2}{\beta_k + \sigma_{w_k}^2} \quad \textit{and} \quad \sigma_{\varepsilon_k}^2 = \frac{\sigma_{w_k}^2 \beta_k}{\beta_k + \sigma_{w_k}^2}, \tag{3.12}$$

where $\sigma_{\hat{g}_k}^2$ *and* $\sigma_{\varepsilon_k}^2$ *are the variances of the independent zero-mean elements of* $\hat{\boldsymbol{g}}_k$ *and* $\boldsymbol{\varepsilon}$, *respectively.*

From Theorem 3.1, it is apparent that in the channel estimation analysis of massive MIMO systems with one-bit ADCs, the estimation error is directly affected not only by the inner product of the pilot sequences, but also by their outer product as well [15]. To get insight into the impact of the one-bit quantization on the channel estimation, in the next corollary we adopt the statistics-aware power control policy proposed in [51]. Apart from its practical advantages, this policy is especially suitable specially for one-bit ADCs since it avoids near-far blockage and hence strong interference. Moreover, this power control approach also leads to simple expressions and provides analytical convenience for our derivation in Section 3.4. Although not the focus of this work, we note that in general a massive MIMO system employing a mixed-ADC architecture will be more resilient than an all one-bit implementation to the near-far effect and jamming. This is an interesting topic for further study.

Corollary 3.1. *For the case in which power control is performed, i.e.,* $p_k = \frac{p}{\beta_k}$ *for some fixed value* p *and for* $k \in \mathcal{K} = \{1, \cdots, K\}$, *the number of users is equal to the length of pilot*

sequences, i.e., $\eta = K$, and the pilot matrix satisfies $\mathbf{\Phi}\mathbf{\Phi}^H = \mathbf{I}_K$, we have

$$\mathbf{C}_{\boldsymbol{x}} = \left(Kp + \sigma_n^2\right)\mathbf{I}_K = \mathbf{D}_{\boldsymbol{x}} \tag{3.13}$$

$$\mathbf{C}_{\boldsymbol{q}_t} = \left(1 - \frac{2}{\pi}\right)\mathbf{I}_K, \tag{3.14}$$

which yields

$$\sigma^2_{\hat{g}_k} = \frac{2}{\pi}\frac{\beta_k}{1 + \frac{\sigma_n^2}{Kp}} \tag{3.15}$$

$$\sigma^2_{\varepsilon_k} = \frac{\left(\frac{Kp}{\sigma_n^2}\left(1 - \frac{2}{\pi}\right) + 1\right)\beta_k}{1 + \frac{Kp}{\sigma_n^2}}. \tag{3.16}$$

Corrollary 3.1 states conditions under which $\mathbf{C}_{\boldsymbol{q}_t}$ is diagonal. In addition, it is evident that the channel estimation suffers from an error floor at high SNRs.

3.2.2 Channel Estimation with Few Full Resolution ADCs

Channel estimation with coarse observations suffers from large errors especially in the high SNR regime. On the other hand, while estimating all channels using high-resolution ADCs is desirable, the resulting power consumption burden makes this approach practically infeasible. This motivates the use of a mixed-ADC architecture for channel estimation to eliminate the large estimation error caused by one-bit quantization while keeping the power consumption penalty at an acceptable level. In the approach described in [24, 25, 27] , $N \ll M$ pairs of

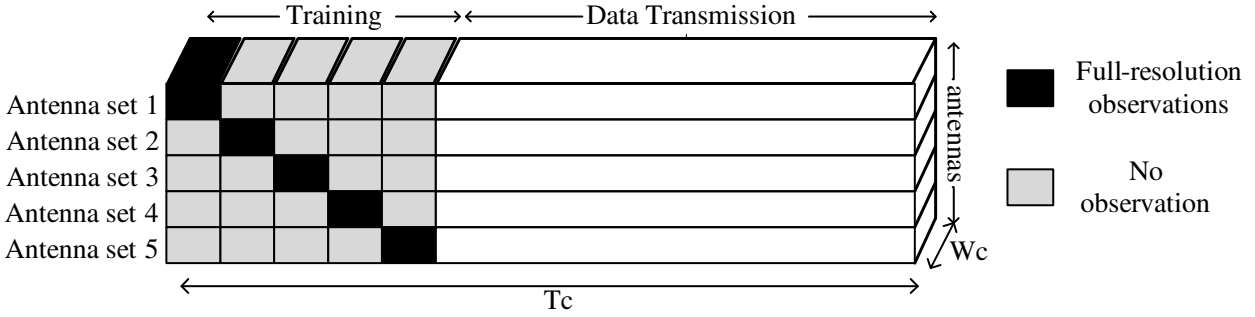

Figure 3.2: Transmission protocol for estimation using full-resolution observations.

high-resolution ADCs are deployed and switched between different antennas during different transmission intervals in an approach referred to as "round-robin" training. In this approach, the M BS antennas are grouped into M/N sets[1]. In the first training sub-interval, users send their mutually orthogonal pilots to the BS while the N high-resolution ADC pairs are connected to the first set of N antennas. After receiving the pilot symbols from all users in the η-symbol-length training sub-interval, the high-resolution ADCs are switched to the next set of antennas and so on. In this manner, after $(M/N)\eta$ pilot transmissions (M/N sub-intervals), we can estimate each channel based on observations with only high-resolution ADCs. This round-robin channel estimation protocol is illustrated in Fig. 3.2 for a mixed-ADC system with $M/N = 5$.

Stacking all $N \times \eta$ full-resolution observations into an $M \times \eta$ matrix, $\mathbf{X}$, the LMMSE estimate of the k-th user channel, $\boldsymbol{g}_k$, is [4]

$$\hat{\boldsymbol{g}}_k = \frac{1}{1 + \frac{\sigma_n^2}{\eta p_k \beta_k}} \frac{1}{\sqrt{\eta p_k}} \mathbf{X} \boldsymbol{\phi}_k^*, \tag{3.17}$$

and the resulting variances of the channel estimate and the error are given respectively by

$$\sigma_{\hat{g}_k}^2 = \frac{\beta_k}{1 + \frac{\sigma_n^2}{\eta p_k \beta_k}} \quad \text{and} \quad \sigma_{\varepsilon_k}^2 = \frac{\beta_k}{1 + \frac{\eta p_k \beta_k}{\sigma_n^2}}. \tag{3.18}$$

[1]We assume M/N is an integer throughout the chapter.

Eq. (3.18) states that by employing only N pairs of high-resolution ADCs and by expending a larger portion of the coherence interval for channel estimation, the channel can be estimated with the same precision as that achieved by conventional high-resolution ADC massive MIMO systems. However, this comes at the high cost of repeating the training data M/N times, which can significantly reduce the time available for data transmission. Indeed, we will see later that in some cases, a mixed-ADC implementation with round-robin training achieves a lower SE than a system with all one-bit ADCs because of the long training interval (even with the improvements we propose below for the round-robin method). However, we will also see that there are other situations for which the mixed-ADC round-robin method provides a large gain in SE. The primary goal of this chapter is to elucidate under what conditions these and other competing approaches provide the best performance.

Before analyzing the tradeoff between the gain (lower channel estimation error) and cost (longer training period) of the round-robin approach, in the next subsection we propose channel estimation based on the use of both full-resolution and one-bit data received by the BS in order to further improve the performance of the mixed-ADC architecture with round-robin channel estimation. To our knowledge, this approach has not been considered in prior work on mixed-ADC massive MIMO.

3.2.3 Estimation Using Joint Full-Resolution/One-Bit Observations

While channel estimation performance based on coarsely quantized observations suffers from large errors in the high SNR regime, it provides reasonable performance for low SNRs. Hence, in this subsection we consider joint channel estimation based on observations from both high-resolution and one-bit ADCs to further improve the channel estimation accuracy. Unlike the previous subsection in which the one-bit ADCs were not employed, here we incorporate their

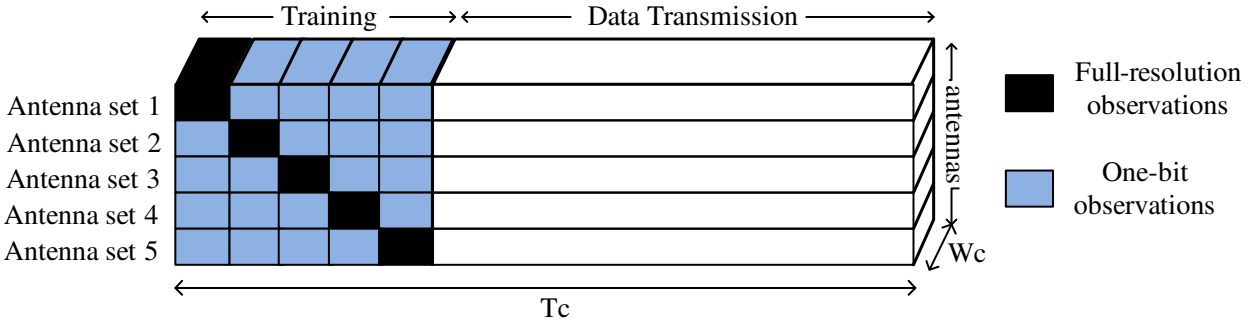

Figure 3.3: Transmission protocol for estimation using full-resolution/one-bit observations.

coarse observations into the channel estimation procedure. The protocol for this method is illustrated in Fig. 3.3 for a mixed-ADC system with $M/N = 5$. It can be seen that, in addition to one set of full-resolution observations for each antenna, there are $(M/N) - 1$ sets of one-bit observations which are also taken into account for channel estimation. The next theorem characterizes the performance of this approach.

Theorem 3.2. *Stacking all $N \times \eta$ full-resolution observations into an $M \times \eta$ matrix, $\mathbf{X}$, and all $(M/N) - 1$ $N \times \eta$ one-bit quantized observations into $M \times \eta$ matrices, $\mathbf{Y}_t$, $t \in \mathcal{T} = \{1, ..., M/N - 1\}$, the LMMSE estimate of the k-th user channel, $\boldsymbol{g}_k$, is*

$$\hat{\boldsymbol{g}}_k = \sqrt{\frac{1}{\eta p_k}} \left(w_{\infty_k} \mathbf{X} \boldsymbol{\phi}_k^* + w_{1_k} \sum_{t=1}^{\frac{M}{N}-1} \mathbf{Y}_t \bar{\boldsymbol{\phi}}_k^* \right), \tag{3.19}$$

where

$$w_{\infty_k} = \frac{\frac{\eta p_k}{\sigma_n^2}}{\frac{1}{\beta_k} + \frac{\eta p_k}{\sigma_n^2} + \varsigma_k(p_k)} \tag{3.20}$$

$$w_{1_k} = \frac{\left(\frac{M}{N} - 1\right)^{-1} \varsigma_k(p_k)}{\frac{1}{\beta_k} + \frac{\eta p_k}{\sigma_n^2} + \varsigma_k(p_k)} \tag{3.21}$$

$$\varsigma_k(p_k) = \frac{\left(\frac{M}{N} - 1\right)}{\sigma_{w_k}^2 + \left(\frac{M}{N} - 2\right) \varrho_k} \tag{3.22}$$

$$\sigma_{w_k}^2 = \frac{1}{\eta p_k} \left(\sigma_n^2 + \bar{\boldsymbol{\phi}}_k^T \mathbf{C}_{\boldsymbol{q}_t} \bar{\boldsymbol{\phi}}_k^* \right) \tag{3.23}$$

$$\varrho_k = \frac{1}{\eta p_k} \bar{\boldsymbol{\phi}}_k^T \bar{\mathbf{C}}_{\boldsymbol{q}_t} \bar{\boldsymbol{\phi}}_k^* \tag{3.24}$$

$$\bar{\mathbf{C}}_{\boldsymbol{q}_t} = \frac{2}{\pi} \arcsin\{\bar{\mathbf{D}}_{\boldsymbol{x}}^{-\frac{1}{2}} \bar{\mathbf{C}}_{\boldsymbol{x}} \bar{\mathbf{D}}_{\boldsymbol{x}}^{-\frac{1}{2}}\} - \frac{2}{\pi} \bar{\mathbf{D}}_{\boldsymbol{x}}^{-\frac{1}{2}} \bar{\mathbf{C}}_{\boldsymbol{x}} \bar{\mathbf{D}}_{\boldsymbol{x}}^{-\frac{1}{2}} \tag{3.25}$$

$$\bar{\mathbf{C}}_{\boldsymbol{x}} = \sum_{k=1}^{K} \eta p_k \beta_k \boldsymbol{\phi}_k^* \boldsymbol{\phi}_k^T \tag{3.26}$$

$$\bar{\mathbf{D}}_{\boldsymbol{x}} = \mathrm{diag}\{\bar{\mathbf{C}}_{\boldsymbol{x}}\}. \tag{3.27}$$

This approach yields the following variances for the channel estimate and the estimation error, respectively:

$$\sigma_{\hat{g}_k}^2 = \frac{\frac{\eta p_k}{\sigma_n^2} + \varsigma_k(p_k)}{\frac{1}{\beta_k} + \frac{\eta p_k}{\sigma_n^2} + \varsigma_k(p_k)} \beta_k \tag{3.28}$$

$$\sigma_{\varepsilon_k}^2 = \frac{1}{\frac{1}{\beta_k} + \frac{\eta p_k}{\sigma_n^2} + \varsigma_k(p_k)}. \tag{3.29}$$

Proof. From (3.2), the observations from the high-resolution ADCs can be written as

$$\boldsymbol{v}(0) = \sqrt{\frac{1}{\eta p_k}} \mathbf{X} \boldsymbol{\phi}_{\boldsymbol{k}}^{*} = \boldsymbol{g}_k + \tilde{\boldsymbol{n}}(0), \tag{3.30}$$

where $\tilde{\boldsymbol{n}}(0) \sim \mathcal{CN}(\mathbf{0}, \frac{\sigma_n^2}{\eta p_k} \boldsymbol{I}_M)$. In addition, from (3.8), the observations from the one-bit ADCs become

$$\boldsymbol{v}(t) = \sqrt{\frac{1}{\eta p_k}} \mathbf{Y}_t \bar{\boldsymbol{\phi}}_{\boldsymbol{k}}^{*} = \boldsymbol{g}_k + \tilde{\boldsymbol{n}}(t) + \tilde{\boldsymbol{q}}(t), \quad t \in \mathcal{T}, \tag{3.31}$$

where $\tilde{\boldsymbol{n}}(t) \sim \mathcal{CN}(\mathbf{0}, \frac{\sigma_n^2}{\eta p_k} \boldsymbol{I}_M)$ is independent of $\tilde{\boldsymbol{n}}(t')$ for $t \neq t'$, and $\tilde{\boldsymbol{q}}(t) = \sqrt{\frac{1}{\eta p_k}} \mathbf{Q}(t) \bar{\boldsymbol{\phi}}_{\boldsymbol{k}}^{*}$. Since the elements of $\boldsymbol{v}(t)$ are independent, we can estimate the mth channel g_{mk} separately. Therefore, stacking all the observations in a vector, we can write

$$\underbrace{\begin{bmatrix} v_m(0) \\ \vdots \\ v_m(t) \\ \vdots \\ v_m(\frac{M}{N}-1) \end{bmatrix}}_{\mathbf{v}} = \underbrace{\begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{1}_{\frac{M}{N}}} g_{mk} + \underbrace{\begin{bmatrix} \tilde{n}_m(0) \\ \vdots \\ \tilde{n}_m(t) + \tilde{q}_m(t) \\ \vdots \\ \tilde{n}_m(\frac{M}{N}-1) + \tilde{q}_m(\frac{M}{N}-1) \end{bmatrix}}_{\mathbf{u}}. \tag{3.32}$$

As a result, the LMMSE estimation of the mth channel coefficient for the kth user is [54]

$$\hat{g}_{mk} = \left(\frac{1}{\beta_k} + \mathbf{1}_{\frac{M}{N}}^T \mathbf{C}_{\mathbf{u}}^{-1} \mathbf{1}_{\frac{M}{N}} \right)^{-1} \mathbf{1}_{\frac{M}{N}}^T \mathbf{C}_{\mathbf{u}}^{-1} \mathbf{v}. \tag{3.33}$$

In Eq. (3.33), $\mathbf{C_u}$ denotes the covariance matrix of $\mathbf{u}$ which is a block diagonal matrix of the form

$$\mathbf{C_u} = \begin{bmatrix} \frac{\sigma_n^2}{\eta p_k} & 0 & \dots & 0 \\ 0 & \sigma_{w_k}^2 & \dots & \varrho_k \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \varrho_k & \dots & \sigma_{w_k}^2 \end{bmatrix} = \begin{bmatrix} \frac{\sigma_n^2}{\eta p_k} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{bmatrix}, \tag{3.34}$$

where

$$\varrho_k = \mathbb{E}\{(\tilde{n}_m(t) + \tilde{q}_m(t))\,(\tilde{n}_m(t') + \tilde{q}_m(t'))^*\}, \quad t \neq t', \tag{3.35}$$

can be easily calculated with the aid of the Bussgang decomposition and the arcsine law as in (3.24). Substituting (3.34) into (3.33), we have

$$\hat{g}_{mk} = \left(\frac{1}{\beta_k} + \frac{\eta p_k}{\sigma_n^2} + \mathbf{1}_{\frac{M}{N}-1}^T \mathbf{S}^{-1} \mathbf{1}_{\frac{M}{N}-1}\right)^{-1} \times \begin{bmatrix} \frac{\sigma_n^2}{\eta p_k} & \mathbf{1}_{\frac{M}{N}-1}^T \mathbf{S}^{-1} \end{bmatrix} \mathbf{v}. \tag{3.36}$$

To calculate the inverse of the matrix $\mathbf{S}$, we re-write it as

$$\mathbf{S} = \left(\sigma_{w_k}^2 - \varrho_k\right) \mathbf{I}_{\frac{M}{N}-1} + \varrho_k \mathbf{1}_{\frac{M}{N}-1} \mathbf{1}_{\frac{M}{N}-1}^T, \tag{3.37}$$

and use Woodbury's matrix identity:

$$\mathbf{S}^{-1} = \frac{1}{\sigma_{w_k}^2 - \varrho_k} \mathbf{I}_{\frac{M}{N}-1} - \frac{1}{\left(\sigma_{w_k}^2 - \varrho_k\right)^2} \left(\frac{1}{\varrho_k} + \frac{(\frac{M}{N} - 1)}{\sigma_{w_k}^2 - \varrho_k}\right)^{-1} \mathbf{1}_{\frac{M}{N}-1} \mathbf{1}_{\frac{M}{N}-1}^T, \tag{3.38}$$

which yields

$$\mathbf{1}_{\frac{M}{N}-1}^T \mathbf{S}^{-1} = \frac{1}{\sigma_{w_k}^2 + \left(\frac{M}{N} - 2\right) \varrho_k} \mathbf{1}_{\frac{M}{N}-1}^T, \tag{3.39}$$

$$\mathbf{1}_{\frac{M}{N}-1}^{T}\mathbf{S}^{-1}\mathbf{1}_{\frac{M}{N}-1} = \frac{\left(\frac{M}{N}-1\right)}{\sigma_{w_k}^2 + \left(\frac{M}{N}-2\right)\varrho_k}. \tag{3.40}$$

Substituting (3.39) and (3.40) into (3.36) completes the proof. ■

Theorem 3.2 demonstrates the optimal approach for combining the observations from high-resolution and one-bit ADCs. In addition, this highlights the importance of considering the correlation among the one-bit observations in the analysis of mixed-ADC channel estimation, something that could not be addressed by the widely-used AQNM approach. More precisely, it can be seen that the impact of joint high-resolution/one-bit channel estimation is manifested in the variance of the channel estimation error by the term $\varsigma_k(p_k)$. To see this, assume that the correlation among one-bit observations in different training sub-intervals is ignored (as would be the case with the AQNM approach). As shown in the appendix, this is equivalent to setting $\varrho_k = 0$ in (3.24). Under this assumption, $\varsigma_k(p_k)$ becomes

$$\varsigma_{k_0}(p_k) = \frac{\left(\frac{M}{N}-1\right)}{\sigma_{w_k}^2} > \varsigma_k(p_k), \tag{3.41}$$

and thus, $\sigma_{\varepsilon_k}^2 > \sigma_{\varepsilon_{k_0}}^2$ where $\sigma_{\varepsilon_{k_0}}^2$ denotes the estimation error for $\varrho_k = 0$. Consequently, the AQNM model yields an overly optimistic assessment of the channel estimation error compared with the more accurate Bussgang analysis. We will see below that the impact of the AQNM approximation is significant for mixed-ADC channel estimation.

The next corollary provides insight into the impact of the system parameters on the joint high-resolution/one-bit LMMSE estimation.

Corollary 3.2. *For the case in which power control is performed, i.e., $p_k = \frac{p}{\beta_k}$ for $k \in \mathcal{K}$, the number of users is equal to the length of pilot sequences, i.e., $\eta = K$, and the pilot matrix*

satisfies $\mathbf{\Phi}\mathbf{\Phi}^H = \mathbf{I}_K$*, we have*

$$\bar{\mathbf{C}}_{\boldsymbol{x}} = Kp\mathbf{I}_K = \bar{\mathbf{D}}_{\boldsymbol{x}}, \tag{3.42}$$

and

$$\bar{\mathbf{C}}_{\boldsymbol{q}_t} = \left(1 - \frac{2}{\pi}\right)\mathbf{I}_K, \tag{3.43}$$

which yields

$$\sigma^2_{\hat{g}_k} = \frac{\frac{Kp}{\sigma_n^2} + \varsigma(p)}{1 + \frac{Kp}{\sigma_n^2} + \varsigma(p)}\beta_k \quad \textit{and} \quad \sigma^2_{\varepsilon_k} = \frac{1}{1 + \frac{Kp}{\sigma_n^2} + \varsigma(p)}\beta_k, \tag{3.44}$$

where

$$\varsigma(p) = \frac{\left(\frac{M}{N} - 1\right)}{\frac{\pi}{2}\frac{\sigma_n^2}{Kp} + \left(\frac{\pi}{2} - 1\right)\left(\frac{M}{N} - 1\right)}. \tag{3.45}$$

In addition,

$$w_\infty = \frac{\frac{Kp}{\sigma_n^2}}{1 + \frac{Kp}{\sigma_n^2} + \varsigma(p)} \quad \textit{and} \quad w_1 = \frac{\left(\frac{M}{N} - 1\right)^{-1}\varsigma(p)}{1 + \frac{Kp}{\sigma_n^2} + \varsigma(p)}, \tag{3.46}$$

where w_∞ *and* w_1 *denote the weights of the high-resolution and one-bit observations in the LMMSE estimation, respectively.*

Corallary 3.2 states that in contrast to Theorem 1 where the correlation among one-bit observations within each training sub-interval can be eliminated by carefully selecting the system parameters as in Corollary 3.1, we cannot overcome the correlation among one-bit observations from different training sub-intervals. This phenomenon makes the addition of the one-bit observations less useful especially in the high SNR regime. For instance, in the

asymptotic case, as the SNR = $\frac{p}{\sigma_n^2}$ goes to infinity, we have

$$\varsigma \longrightarrow \frac{1}{\frac{\pi}{2} - 1}, \tag{3.47}$$

$$w_\infty \longrightarrow 1, \; w_1 \longrightarrow 0. \tag{3.48}$$

It is apparent from (3.47) that in the asymptotic regime ς tends to a finite value and also is independent of M/N. Moreover, (3.48) implies that the optimal approach for high SNRs is to estimate the channel based solely on the high-resolution observations.

The error for the three channel estimation approaches in Eqs. (3.12), (3.18), and (3.29) is depicted in Fig. 3.4 for a case with $M = 100$ antennas, $K = 10$ users, and various numbers of high-resolution ADCs, N and training lengths η. The label "Joint" refers to round-robin channel estimation that includes the one-bit observations as described in the previous section, "Full resolution" indicates the performance achieved using a full array of high-resolution ADCs, and "One-bit" refers to the performance of an all-one-bit architecture. We also plot the performance predicted for the Joint approach based on the AQNM analysis, which ignores the correlation among the one-bit observations. We see that the AQNM-based analysis yields an overly optimistic prediction for the channel estimation error. In particular, unlike AQNM, the more accurate Bussgang analysis shows that channel estimation with all an one-bit BS actually outperforms the Joint method for low SNRs, a critical observation in analyzing whether or not a mixed-ADC implementation makes sense. However, we see that the mixed-ADC architecture eventually overcomes the error floor of the all one-bit system for high SNRs and in such cases can reduce the estimation error dramatically. Fig. 3.4 focuses on channel estimation performance, but does not reflect the full impact of the round-robin training on the overall system spectral efficiency, since reducing N increases the amount

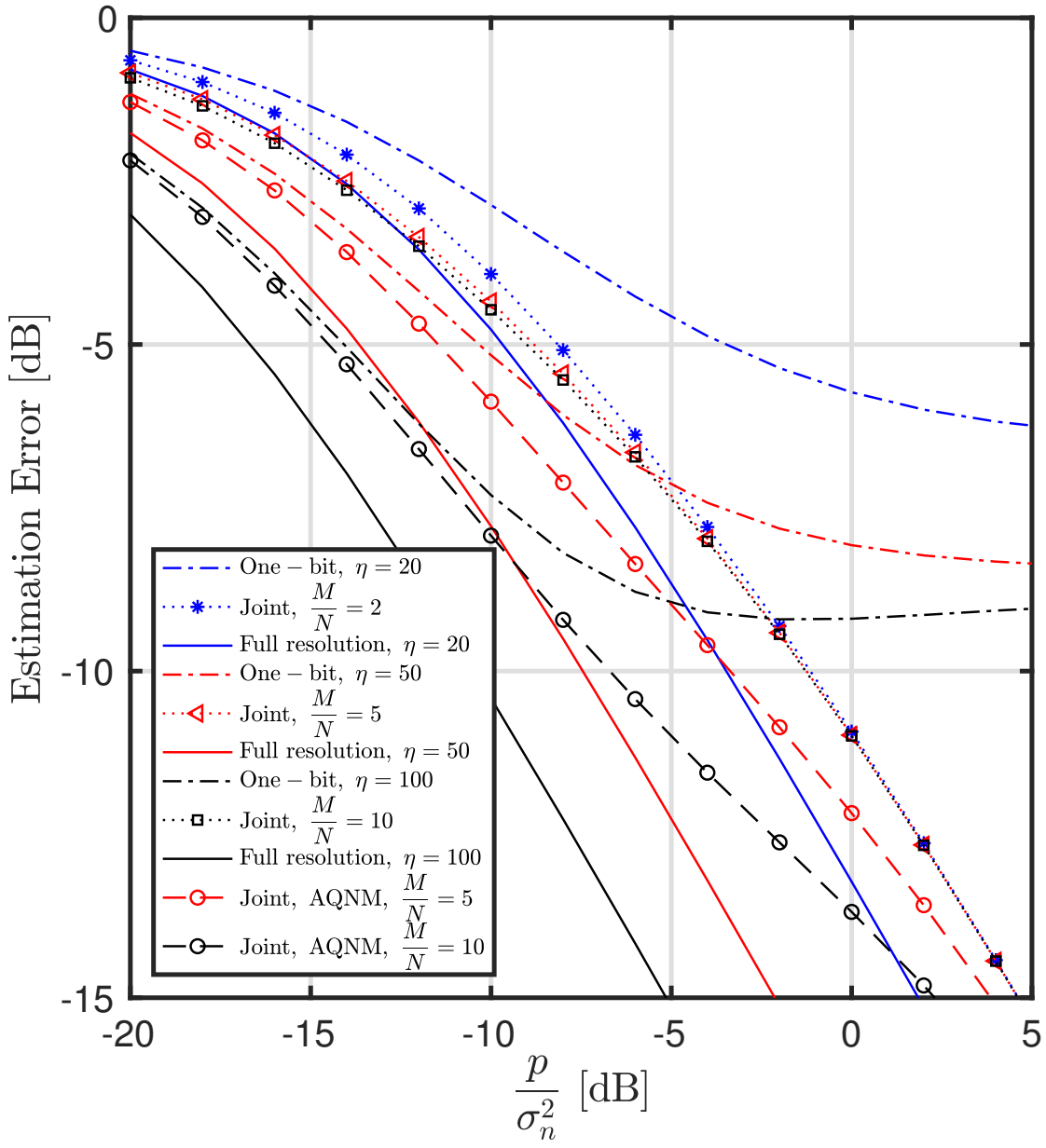

Figure 3.4: Channel estimation error $\sigma_{\varepsilon_k}^2/\beta_k$ versus p/σ_n^2.

of training required by the round-robin method. This will be taken into account when we analyze the SE in Section 3.4.

3.3 Practical Considerations

The improvement in channel estimation performance provided by the round-robin training clearly comes at the expense of a significantly increased training overhead. For example, consider a simple worst-case example with a 400 Hz Doppler spread in a narrowband channel of 400 kHz bandwidth; in this case, the coherence time is roughly 1000 symbols. For higher bandwidths or smaller cells with lower mobility, the coherence time can easily approach 10,000 symbols or more. A mixed-ADC array of 128 antennas with 16 high-resolution ADCs would require repeating the pilots 8 times, which for 20 users would amount to 160 symbols, or 16% of the coherence time when $T = 1000$ symbols. This is a relatively high price to pay, and as we will see later, in many instances the resulting loss in SE cannot be offset by the improved channel estimate. However, we will also see that on the other hand, there are other situations where the opposite is true, where the round-robin method leads to significant gains in SE even taking the training overhead into account.

Besides the extra training overhead, the round-robin method has the disadvantage of requiring extra RF switching or multiplexing hardware prior to the ADCs, as shown in Fig. 3.1. It is unlikely that a single large $M \times M$ multiplexer would be used for this purpose, since complete flexibility in assigning a given high-resolution ADC to any possible antenna is not needed. A more likely architecture would employ a bank of smaller multiplexers that allows one high-resolution ADC to be switched among a smaller subarray of antennas, ensuring that each RF chain has access to high-resolution training data during one of the round-robin intervals. Such an approach is similar to the simplified "subarray switching" schemes proposed for antenna selection in massive MIMO [55, 56, 57]. In an interesting earlier example,

a large 108×108 multiplexer chipset for a local area network application was developed in [58], composed of several 36×36 differential crosspoint ASIC switches that consume less than 100 mW each, with a bandwidth of 140 MHz and a 0 dB insertion loss.

In the 20 years since [58], RF switch technology has advanced considerably. For the example discussed above involving a 128-element array with 16 high-resolution ADCs and 112 one-bit ADCs, the multiplexing could be achieved using 16 8×8 analog switches arranged in parallel. Consider the Analog Devices ADV3228 8×8 crosspoint switch as an example of an off-the-shelf component for such an architecture[2]. The ADV3228 has a 750 MHz bandwidth, a switching time of 15 ns, and a power consumption of 500 mW, which is similar to that of an 8-bit ADC (for example, see Texas Instruments' ADC08B200 8-bit 200 MS/s ADC[3]). Since the switches can be implemented at a lower intermediate frequency prior to the I-Q demodulation, only one per subarray is required, and thus the total power consumption of the switches would be less than half that of the ADCs.

Note that for the vast majority of the coherence time, the switch is idle. To accommodate the round-robin training, the switches only need to be operated $\frac{M}{N} - 1$ times, once for every repetition of the training data. This reduces the actual power consumption to below the specification, and further reduces the impact of the additional training. Short guard intervals would need to be inserted between the training intervals to account for the switching transients, but these will typically not impact the SE. For the example discussed above with 128 antennas and 8 switches, 7 switching events are required for a total switching time of 105 ns, which is insignificant compared to the coherence time of 2.5 ms at a 400 Hz Doppler.

The insertion loss of the analog switches would also have to be taken into account in an actual implementation, since this will directly reduce the overall SNR of the received signals. Harmonic interference due to nonlinearities in the switch are likely not an issue; for example, the

[2]See http://www.analog.com/en/products/switches-multiplexers/buffered-analog-crosspoint-switches/adv3228.html#product-overview for product details.

[3]http://www.ti.com/product/ADC08B200/technicaldocuments.

specifications for a Texas Instruments switch (LMH6583) similar to the ADV3228 indicate that the power of the second and third harmonic distortions were -76 dBc. Furthermore, it has been shown that the use of signal combining with a massive antenna array provides significant robustness to such nonlinearities and other hardware imperfections [6, 11, 7, 8, 9, 10].

3.4 Spectral Efficiency

Although channel estimation with a mixed-ADC architecture using round-robin training can substantially improve the channel estimation accuracy, it requires a longer training interval and, therefore, leaves less room for data transmission in each coherence interval. More precisely, $(M/N)\eta$ symbol transmissions are required for round-robin channel estimation which could be large when the number of high-resolution ADCs, N, is small[4]. Despite losing a portion of the coherence interval for channel estimation due to the mixed-ADC architecture, the improvement in the signal-to-quantization-interference-and-noise ratio (SQINR) can be significant owing to more accurate channel estimation, and thus a higher rate would be expected during this shorter data transmission period. In this section, we study this system performance trade-off in terms of spectral efficiency for the three mentioned channel estimation approaches.

In the data transmission phase, all users simultaneously send their data symbols to the BS. To begin, assume the antennas are ordered so that the last N antennas are connected to high-resolution ADCs in this phase. A more thoughtful assignment of the high-resolution ADCs will be considered below. From equation (3.1), and based on the Bussgang decomposition,

[4]Note that in designing a mixed-ADC system with round-robin channel training, one should consider the ratio M/N in scaling the system instead of just increasing the number of antennas M. In particular, increasing the number of BS antennas requires increasing of the high-resolution ADCs, N, as well.

the received signal at the BS after one-bit quantization is

$$\boldsymbol{y}_d = \begin{bmatrix} \sqrt{\frac{2}{\pi}}\bar{\mathbf{D}}^{-\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix} \boldsymbol{r} + \underbrace{\begin{bmatrix} \bar{\boldsymbol{q}}_d \\ \mathbf{0} \end{bmatrix}}_{\boldsymbol{q}_d} \tag{3.49}$$

$$\bar{\mathbf{D}} = \text{diag}\{\mathbf{C}_{\boldsymbol{r}}\} \tag{3.50}$$

$$\mathbf{C}_{\boldsymbol{r}} = \sum_{k=1}^{K} p_k \bar{\boldsymbol{g}}_k \bar{\boldsymbol{g}}_k^H + \sigma_n^2 \mathbf{I}_{M-N}, \tag{3.51}$$

where $\bar{\boldsymbol{g}}_k$ denotes the $M-N$ elements of $\boldsymbol{g}_k$ corresponding to the $M-N$ one-bit ADCs and $\bar{\boldsymbol{q}}_d$ is the $(M-N)\times 1$ quantization noise in the data transmission phase. It is apparent that the covariance matrix in (3.51) is not diagonal which makes analytical tractability difficult. However, by adopting statistics-aware power control [51], i.e., $p_k = \frac{p}{\beta_k}$, and assuming that the number of users is relatively large (typical for massive MIMO systems), channel hardening occurs [15], and (3.51) can be approximated as

$$\mathbf{C}_{\boldsymbol{r}} \cong \left(Kp + \sigma_n^2\right) \mathbf{I}_{M-N} = \bar{\mathbf{D}}. \tag{3.52}$$

As a result, according to the arcsine law (see (3.7)), the covariance matrix of the quantization noise in the data transmission phase becomes $\mathbf{C}_{\bar{\boldsymbol{q}}_d} \cong (1 - 2/\pi)\mathbf{I}_{M-N}$ and (3.49) simplifies to

$$\boldsymbol{y}_d \cong \mathbf{A}\left(\sum_{k=1}^{K} \sqrt{p}\boldsymbol{h}_k s_k + \boldsymbol{n}\right) + \boldsymbol{q}_d \tag{3.53}$$

$$\mathbf{A} = \begin{bmatrix} \alpha \mathbf{I}_{M-N} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix},$$

where $\alpha \triangleq \sqrt{\frac{2}{\pi} \frac{1}{(Kp+\sigma_n^2)}}$.

For data detection, the BS selects a linear receiver $\boldsymbol{W} \in \mathbb{C}^{M \times K}$ as a function of the channel estimate. Note that the quantization model considered in (3.4) and (3.5) does not preserve the power of the input of the quantizer since the power of the output is forced to be 1. Thus we premultiply the received signal as follows to offset this effect:

$$\hat{\boldsymbol{y}}_d = \mathbf{A}^{-1} \boldsymbol{y}_d. \tag{3.54}$$

By employing the linear detector $\mathbf{W}$, the resulting signal at the BS is

$$\hat{\boldsymbol{s}} = \boldsymbol{W}^H \hat{\boldsymbol{y}}_d. \tag{3.55}$$

Thus, the kth element of $\hat{\boldsymbol{s}}$ is

$$\hat{s}_k = \sqrt{p} \boldsymbol{w}_k^H \boldsymbol{h}_k s_k + \sqrt{p} \sum_{i=1, i \neq k}^{K} \boldsymbol{w}_k^H \boldsymbol{h}_i s_i + \boldsymbol{w}_k^H \boldsymbol{n} + \boldsymbol{w}_k^H \mathbf{A}^{-1} \boldsymbol{q}_d, \tag{3.56}$$

where $\boldsymbol{w}_k$ is the kth column of $\boldsymbol{W}$. We assume the BS treats $\boldsymbol{w}_k^H \boldsymbol{h}_k$ as the gain of the desired signal and the other terms of (3.56) as Gaussian noise when decoding the signal[5]. Consequently, we can use the classical bounding technique of [51] to derive an approximation for the ergodic achievable SE at the kth user as

$$\mathcal{S}_k = \mathcal{R}\left(\mathrm{SQINR}_k\right), \tag{3.57}$$

[5]Note that in general, the quantization noise is not Gaussian. However, to derive a lower bound for the SE, we assume it is Gaussian with covariance $\mathbf{C}_{\boldsymbol{q}_d}$.

where the effective SQINR_k is defined by

$$\mathrm{SQINR}_k = \frac{p\left|\mathbb{E}\left\{\boldsymbol{w}_k^H \boldsymbol{h}_k\right\}\right|^2}{p\sum_{i=1}^{K}\mathbb{E}\left\{\left|\boldsymbol{w}_k^H \boldsymbol{h}_i\right|^2\right\} - p\left|\mathbb{E}\left\{\boldsymbol{w}_k^H \boldsymbol{h}_k\right\}\right|^2 + \sigma_n^2\mathbb{E}\left\{\|\boldsymbol{w}_k\|^2\right\} + \alpha^{-2}\mathbb{E}\left\{\boldsymbol{w}_k^H \mathbf{C}_{\boldsymbol{q}_d}\boldsymbol{w}_k\right\}}, \quad (3.58)$$

and $\mathcal{R}(\theta) \triangleq (1-\eta_{\mathrm{eff}}/T)\log_2(1+\theta)$ where η_{eff} represents the training duration which is η and $(M/N)\eta$ for the pure one-bit and mixed-ADC architectures, respectively.

3.4.1 MRC Detection

3.4.1.1 Random Mixed-ADC Detection

In this subsection, we consider the case in which the high-resolution ADCs are connected to an arbitrary set of N antennas. Denoting the estimate of the channel by $\hat{\mathbf{H}} = [\hat{\boldsymbol{h}}_1, ..., \hat{\boldsymbol{h}}_K]$, setting $\mathbf{W} = \hat{\mathbf{H}}$, and following the same reasoning as in [15], the SE of the mixed-ADC architecture with MRC detection can be derived as

$$\mathcal{S}_k^{\mathrm{MRC}} = \mathcal{R}\left(\frac{pM\sigma_{\hat{h}}^2}{pK + \sigma_n^2 + \frac{\left(1-\frac{2}{\pi}\right)}{\alpha^2}\left(1-\frac{N}{M}\right)}\right), \quad (3.59)$$

where the channel estimate variance $\sigma_{\hat{h}}^2 = \sigma_{\hat{g}_k}^2/\beta_k$ depends on the estimation approach as denoted in (3.12), (3.18), and (3.28).

From (3.59), it can be observed that the gain of exploiting the mixed-ADC architecture is manifested in the SE expressions by two factors, channel estimation improvement by a factor of $\sigma_{\hat{h}}^2$, and quantization noise reduction by a factor of $1 - N/M$.

3.4.1.2 Mixed-ADC Detection with Antenna Selection

Having an accurate channel estimate can help us to employ the N high-resolution ADCs in an intelligent manner to further improve the performance of the mixed-ADC architecture. A careful look at the SQINR expression in (3.58) reveals that the effect of one-bit quantization on the SE is manifested by the last term of the denominator. Hence, one can maximize the SE by minimizing this term through smart use of the N high-resolution ADCs. We refer to this approach as *Mixed-ADC with Antenna Selection*. We consider an antenna selection scheme suggested by the SQINR expression in (3.58). In this approach, the N high-resolution ADCs are connected to the antennas corresponding to rows of $\hat{\mathbf{H}}$ with the largest energy, i.e. $\sum_{k=1}^{K}\left|\hat{h}_{mk}\right|^2$. Besides numerical evaluation in Section 3.5, in Theorem 3.3 we derive a bound for the SE achieved by MRC detection with antenna selection.

Theorem 3.3. *The spectral efficiency of the mixed-ADC system with antenna selection and an MRC receiver is lower bounded by*

$$\bar{\mathcal{S}}_k^{\mathrm{MRC}} = \mathcal{R}\left(\frac{pM\sigma_{\hat{h}}^2}{pK+\sigma_n^2+\frac{\left(1-\frac{2}{\pi}\right)}{MK\alpha^2}\left(\sum_{m=1}^{M-N}\chi_m\right)}\right), \tag{3.60}$$

where χ_m is defined as

$$\chi_m = \frac{M!}{(m-1)!\,(M-m)!}\sum_{\ell=0}^{M-m}(-1)^{-\ell}\binom{M-m}{\ell}\left(\Gamma(K)\right)^{-m-\ell}K^{1-m-\ell}\Gamma\left(1+K\left(m+\ell\right)\right) \times \mathcal{F}_A^{(m+\ell-1)}\left(1+K\left(m+\ell\right);K,\cdots,K;K+1,\cdots,K+1;-1,\cdots,-1\right) \tag{3.61}$$

and $\mathcal{F}_A$ denotes the Lauricella function of type A [59].

Proof. Denote the energy of the mth row, $m \in \mathcal{M} = \{1, ..., M\}$, of $\hat{\mathbf{H}}$ by $\mathcal{E}_m$, i.e.,

$$\mathcal{E}_m \triangleq \sum_{k=1}^{K} \left| \hat{h}_{mk} \right|^2. \tag{3.62}$$

To do antenna selection, we must connect the N high-resolution ADCs to the antennas corresponding to the largest $\mathcal{E}_m$. Suppose that the indices of the N antennas to which the high-resolution ADCs are connected are contained in the set $\mathcal{N}$. Hence, we have

$$\sum_{k=1}^{K} \mathbb{E}\left\{ \hat{\boldsymbol{h}}_k^H \mathbf{C}_{\boldsymbol{q}_d} \hat{\boldsymbol{h}}_k \right\} = K\mathbb{E}\left\{ \hat{\boldsymbol{h}}_k^H \mathbf{C}_{\boldsymbol{q}_d} \hat{\boldsymbol{h}}_k \right\} = \left(1 - \frac{2}{\pi}\right) \sum_{\mathcal{M}\setminus\mathcal{N}} \mathbb{E}\{\mathcal{E}_m\}. \tag{3.63}$$

Eq. (3.63) provides a criterion for connecting the N high-resolution ADCs in the data transmission phase. In fact, it states that, for the MRC receiver, the expected value in (3.63) will be minimized if the high-resolution ADCs are connected to the antennas corresponding to the largest $\mathcal{E}_m$. Denote $\mathcal{E}_{(m)}$ as the mth smallest value of $\mathcal{E}_m$, i.e.,

$$\mathcal{E}_{(1)} \leq \mathcal{E}_{(2)} \leq \cdots \leq \mathcal{E}_{(M)}.$$

Hence, $\mathcal{E}_{(m)}$ is the mth order statistic, and assuming that the $\mathcal{E}_{(m)}$ are statistically independent and identically distributed, we have [60]

$$\mathbb{E}\{\mathcal{E}_{(m)}\} = M\binom{M-1}{m-1} \int_{-\infty}^{\infty} x \left[F(x)\right]^{m-1} \left[1 - F(x)\right]^{M-m} dF(x), \tag{3.64}$$

where x is the realization of $\mathcal{E}_{(m)}$ and $F(x)$ is the cumulative distribution function of $\mathcal{E}_m$. For the case that we have considered, where the channel coefficients are i.i.d. Rayleigh distributed, the $\mathcal{E}_m$ are independent Gamma random variables with

$$F(x) = \gamma\left(\frac{x}{\sigma_{\hat{h}}^2}, K\right), \tag{3.65}$$

where $\gamma(.,.)$ denotes the incomplete Gamma function. From [61], the integral (3.64) can be

calculated in closed form for Gamma random variables as

$$\mathbb{E}\{\mathcal{E}_{(m)}\} = \sigma^2_{\hat{h}} \chi_m. \tag{3.66}$$

This is in contrast to the unordered case where $\mathbb{E}\{\mathcal{E}_m\} = K\sigma^2_{\hat{h}}$. As a result

$$\min \left\{ \mathbb{E}\left\{ \hat{\boldsymbol{h}}_k^H \mathbf{C}_{\boldsymbol{q}_d} \hat{\boldsymbol{h}}_k \right\} \right\} = \left(1 - \frac{2}{\pi}\right) \frac{\sigma^2_{\hat{h}}}{K} \sum_{m=1}^{M-N} \chi_m. \tag{3.67}$$

The remaining terms in (3.58) can be calculated similar to the case where the high-resolution ADCs are connected to arbitrary antennas. Plugging these terms and (3.67) into (3.58) and some algebraic manipulation results in (3.60).

■

The lower bound (3.60) explicitly reflects the benefit of antenna selection in the data transmission phase. By comparing (3.60) with (3.59), it is evident that antenna selection has improved the SE by replacing $1 - N/M$ by $\frac{\sum_{m=1}^{M-N} \chi_m}{MK}$. In Section 3.5 we illustrate how antenna selection improves SE for different SNRs. Note that Theorem 3.3 assumes the ability to make an arbitrary assignment of the high-resolution ADCs to different RF chains, which may not be possible if the ADC multiplexing is implemented by a bank of subarray switches. In the numerical results presented later, we show that this does not lead to a significant degradation in performance.

3.4.2 ZF Detection

In this section, we study the SE of the mixed-ADC architecture with ZF detection. To design a mixed-ADC adapted ZF detector, we re-write the last two terms of the denominator of

(3.58) as follows:

$$\boldsymbol{w}_k^H \left(\sigma_n^2 \mathbf{I}_M + \alpha^{-2} \mathbf{C}_{\boldsymbol{q}_d}\right) \boldsymbol{w}_k = \left[\mathbf{W}^H \mathbf{C}_{n_{\text{eff}}} \mathbf{W}\right]_{kk}, \tag{3.68}$$

where $\mathbf{C}_{n_{\text{eff}}} = \sigma_n^2 \mathbf{I}_M + \alpha^{-2} \mathbf{C}_{\boldsymbol{q}_d}$. Accordingly, the ZF detector for the mixed-ADC architecture can be written as

$$\mathbf{W} = \mathbf{C}_{n_{\text{eff}}}^{-1} \hat{\mathbf{H}} \left(\hat{\mathbf{H}}^H \mathbf{C}_{n_{\text{eff}}}^{-1} \hat{\mathbf{H}}\right)^{-1}. \tag{3.69}$$

Plugging (3.69) into (3.58) yields

$$\text{SQINR}_k^{\text{ZF}} = \frac{p}{pK\left(1 - \sigma_{\hat{h}}^2\right) \mathbb{E}\left\{\left[\left(\hat{\mathbf{H}}^H \mathbf{C}_{n_{\text{eff}}}^{-1} \hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}^H \mathbf{C}_{n_{\text{eff}}}^{-2} \hat{\mathbf{H}} \left(\hat{\mathbf{H}}^H \mathbf{C}_{n_{\text{eff}}}^{-1} \hat{\mathbf{H}}\right)^{-1}\right]_{kk}\right\} + \mathbb{E}\left\{\left[\left(\hat{\mathbf{H}}^H \mathbf{C}_{n_{\text{eff}}}^{-1} \hat{\mathbf{H}}\right)^{-1}\right]_{kk}\right\}} \tag{3.70}$$

Similar to the MRC case, the SQINR in (3.70) suggests the same antenna selection approach for ZF detection. In general, calculating the expected values in (3.70) is not tractable neither for arbitrary-antenna mixed-ADC detection nor mixed-ADC with antenna selection. Hence, we numerically evaluate the performance of mixed-ADC with ZF detection in the next section.

3.4.3 Massive MIMO with Uniform ADC Resolution

Contrary to the mixed-ADC architecture where the ADC comparators are concentrated in a few antennas, uniformly spreading the comparators over the array is an alternative approach [20, 21, 22, 62, 63]. In this subsection, we provide the SE expressions for such systems. These expressions will be used in the next section for performance comparisons with the

mixed-ADC architecture.

The SE for the case of all one-bit ADCs was derived in [15] using the Bussgang decomposition. For ADC resolutions of 2 bits or higher, the AQNM model is sufficiently accurate. Using AQNM and following the same reasoning as in [22, 62, 63], the SE of a massive MIMO system with uniform resolution ADCs can be derived as

$$\tilde{\mathcal{S}}_k^{\mathrm{MRC}} = \mathcal{R}\left(\frac{pM\tilde{\sigma}_{\hat{h}}^2}{pK + \sigma_n^2 + \frac{(1-\alpha_0)}{\alpha_0^2}\left(p\left(\tilde{\sigma}_{\hat{h}}^2 + K\right) + \sigma_n^2\right)}\right) \tag{3.71}$$

$$\tilde{\mathcal{S}}_k^{\mathrm{ZF}} = \mathcal{R}\left(\frac{p\,(M-K)\,\tilde{\sigma}_{\hat{h}}^2}{pK\left(1-\tilde{\sigma}_{\hat{h}}^2\right) + \sigma_n^2 + \frac{(M-K)\tilde{\sigma}_{\hat{h}}^2}{\alpha^2}\mathbb{E}\left\{\boldsymbol{w}_k^H \mathbf{C}_0 \boldsymbol{w}_k\right\}}\right), \tag{3.72}$$

for MRC and ZF detection, respectively. In (3.71) and (3.72),

$$\tilde{\sigma}_{\hat{h}}^2 = \frac{\alpha_0^2 \eta p}{\alpha_0^2 \eta p + \alpha_0^2 \sigma_n^2 + \alpha_0\,(1-\alpha_0)\left(pK + \sigma_n^2\right)}, \tag{3.73}$$

α_0 is a scalar depending on the ADC resolution and can be found in Table I of [22], $\boldsymbol{w}_k$ is the kth column of $\mathbf{W} = \hat{\mathbf{H}}\left(\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right)^{-1}$, and $\mathbf{C}_0$ denotes the covariance matrix of the quantization noise based on the AQNM model [22]. The detailed calculation of $\mathbb{E}\left\{\boldsymbol{w}_k^H \mathbf{C}_0 \boldsymbol{w}_k\right\}$ in (3.72) is provided in [63] which we do not include here for the sake of brevity.

3.5 Numerical Results

By substituting from (3.12), (3.18), and (5.16) into (3.59), (3.60), and (3.70), we can evaluate the performance of mixed-ADC architectures for different system settings. For all of the following experiments, we consider a system with $M = 100$ antennas at the BS, and $K = 10$

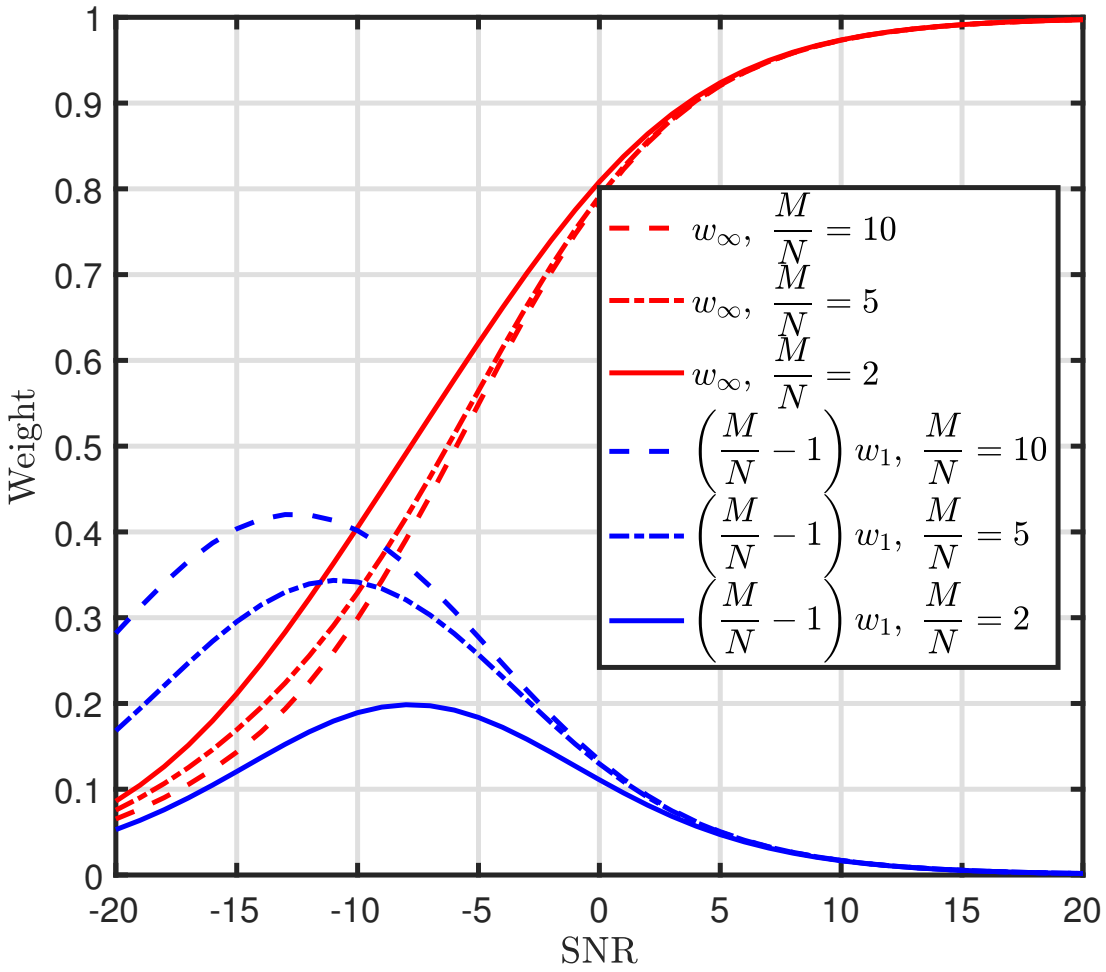

Figure 3.5: Weights used in the LMMSE channel estimator for high-resolution and one-bit observations.

users. Also, we assume the power control approach of [51] is used, so that $p_k\beta_k = p$ for all k. We also assume that an optimal resource allocation has been performed [62, 64] such that the training length, η_{eff}, transmission power during the training phase, p_t, and data transmission phase, p_d are optimized under a power constraint $\eta_{\text{eff}}p_t + (T - \eta_{\text{eff}})p_d = P_{\text{ave}}T$. In the following figures, the SNR is defined as SNR $\triangleq P_{\text{ave}}/\sigma_n^2$.

Fig. 3.5 illustrates the optimal weights for combining high-resolution and one-bit observations for the joint high-resolution/one-bit LMMSE channel estimation. Interestingly, it can be seen that when M/N is large, the one-bit observations are emphasized in the low SNR

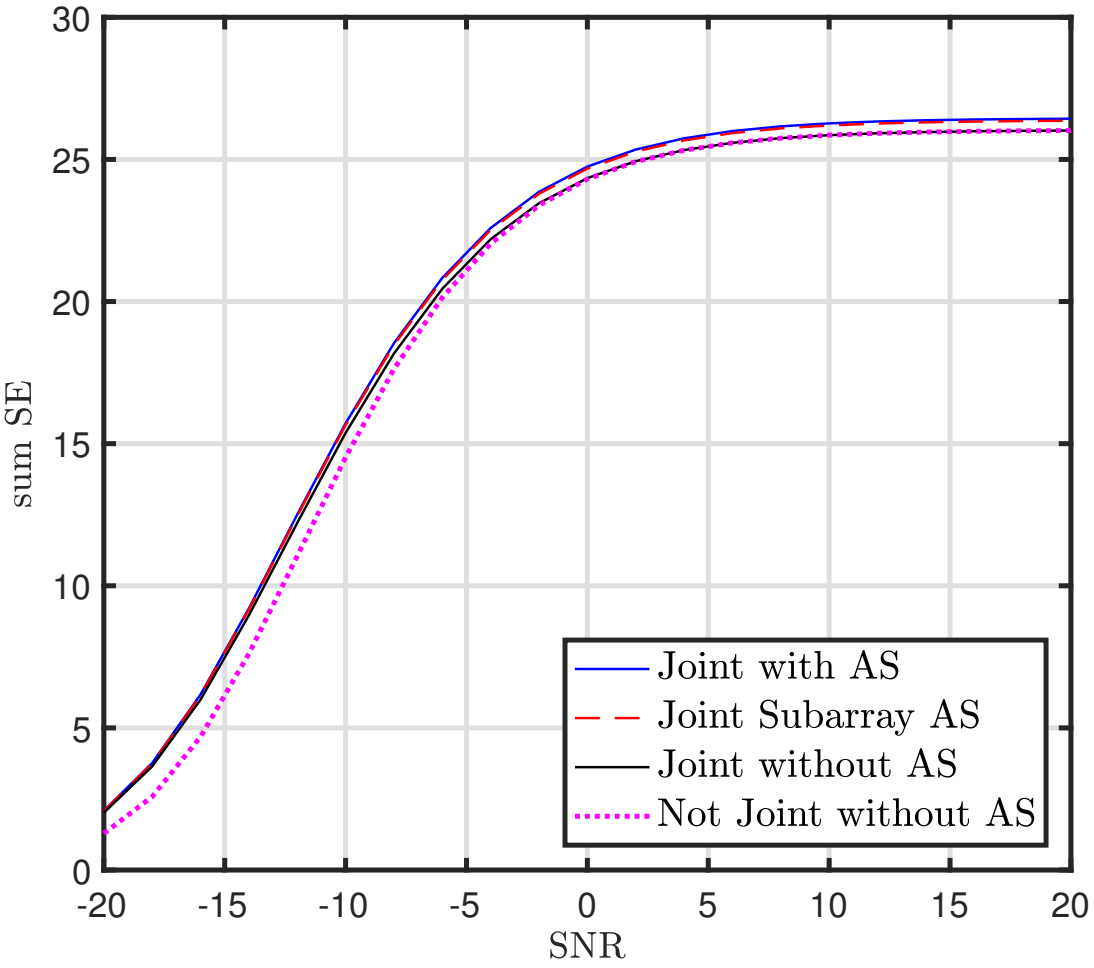

Figure 3.6: Sum SE for MRC detection versus SNR for $M = 100$, $N = 20$, and $T = 400$.

regime relative to the high-resolution observations. In addition, in contrast to the weights for the high-resolution observations, which rise monotonically with increasing SNR, the weight for the one-bit observations grows at first and then decreases to zero.

To study the performance improvement due to joint channel estimation and antenna selection in mixed-ADC massive MIMO, the sum SE for the MRC and ZF detectors for a system with coherence interval $T = 400$ symbols and $N = 20$ high-resolution ADCs is depicted in Fig. 3.6 and Fig. 3.7, respectively. In these and subsequent figures, "Joint with AS" indicates that the channel estimation was performed with both one-bit and high-resolution ADCs and

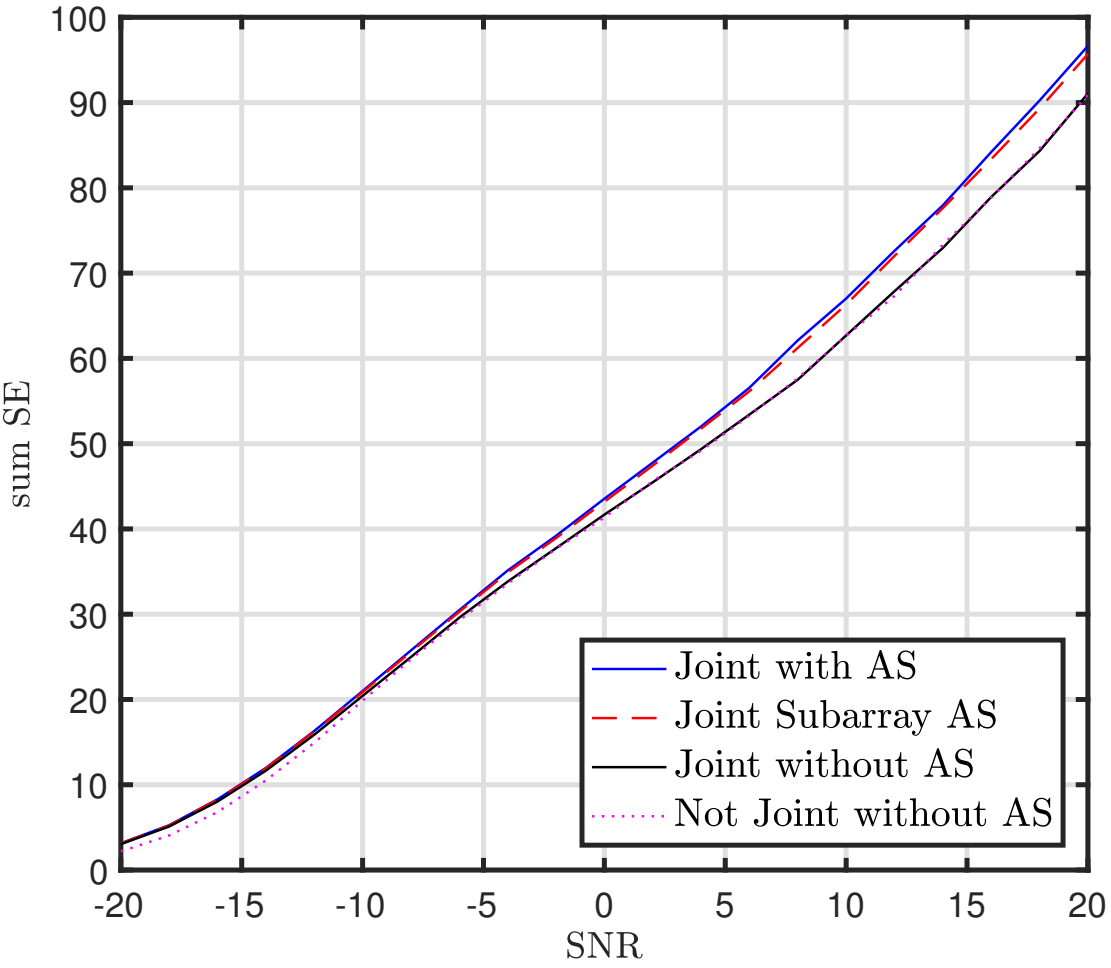

Figure 3.7: Sum SE for ZF detection versus SNR for $M = 100$, $N = 20$, and $T = 400$.

that antenna selection (AS) was used for data detection, "Joint without AS" represents the same case without antenna selection, "Joint Subarray AS" means that the antenna selection only occurred within each M/N-element subarray (one high-resolution ADC assigned to the strongest channel within each subarray), and "Not Joint without AS" represents the case in which channel is estimated based on only high-resolution observations and no antenna selection is employed.

For both MRC and ZF, it can be seen that antenna selection slightly improves the SE for high SNRs, where the channel estimation is most accurate. At low SNR, we see that

joint channel estimation provides a gain from the use of one-bit ADCs, which provide useful information at these SNRs. We also see that the constrained AS required when the switching is only performed within subarrays provides nearly identical performance to the case where arbitrary AS is allowed.

Note that the main reason for the small gain for antenna selection is due to the fact that, with multiple users, selecting a given antenna does not benefit all users simultaneously, and the strong users responsible for a given antenna being selected will in general be different for different antennas. Thus, the improvement due to increased signal-to-noise ratio for some users is somewhat offset by the fact that other users may experience a lower SNR on those same antennas. We would see a much larger benefit for antenna selection if only a single user were present.

Figs. 3.8 and 3.9 provide a comparison among a mixed-ADC massive MIMO system with joint channel estimation and antenna selection, an all-one-bit architecture ("One-bit"), and a mixed-ADC without round-robin training for which the high-resolution ADCs are connected to a fixed set of antennas without ADC switching or antenna selection ("Non-round-robin") [28]. Since mixed-ADC channel estimation improves the channel estimation accuracy by expending a larger portion of the coherence interval for training, its benefit is directly related to the length of the coherence interval. For MRC detection, when $T = 400$, the mixed-ADC architecture performs better than the all-one-bit architecture for $N = 20$, but when $N = 10$ the all-one-bit architecture is better due to the larger training overhead incurred when N is smaller. However, for $T = 1000$, mixed-ADC outperforms the all-one-bit architecture at high SNRs for both $N = 10,\ 20$, while the all-one-bit case is still better for $N = 10$ at low SNRs. Round-robin training provides better SE performance at high SNR when $N = 20$ compared to the case without antenna switching, especially for the larger coherence interval. However, for other cases, the round-robin training overhead significantly reduces the SE, especially for $N = 10$ and the shorter coherence interval.

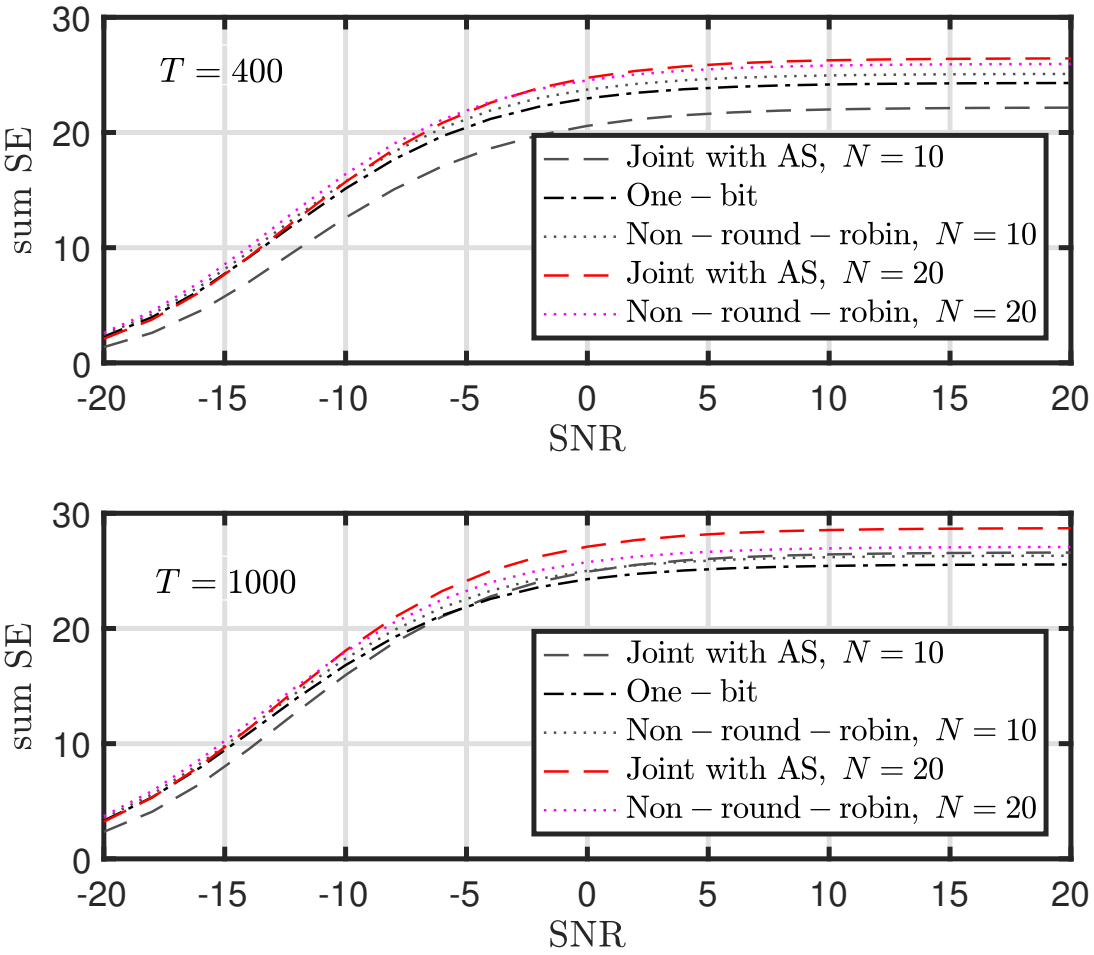

Figure 3.8: Sum SE for MRC detection versus SNR for $M = 100$, $N = 20$, 10, and $T = 400$, 1000.

For ZF detection, we see that the mixed-ADC architectures can provide very large gains in SE compared to the one-bit case at high SNRs, regardless of T. For low SNRs, there is little to no improvement. These cases still do not show a significant benefit for round-robin training compared with a fixed ADC assignment; only when $N = 20$ and $T = 1000$ do we see a slight improvement.

For $N = 20$, Figs. 3.10 and 3.11 show how the coherence interval T impacts the effectiveness of the mixed-ADC architecture for MRC and ZF detectors, respectively. For mixed-ADC MRC detection, it is apparent that the best choice among the three architectures (all one-bit,

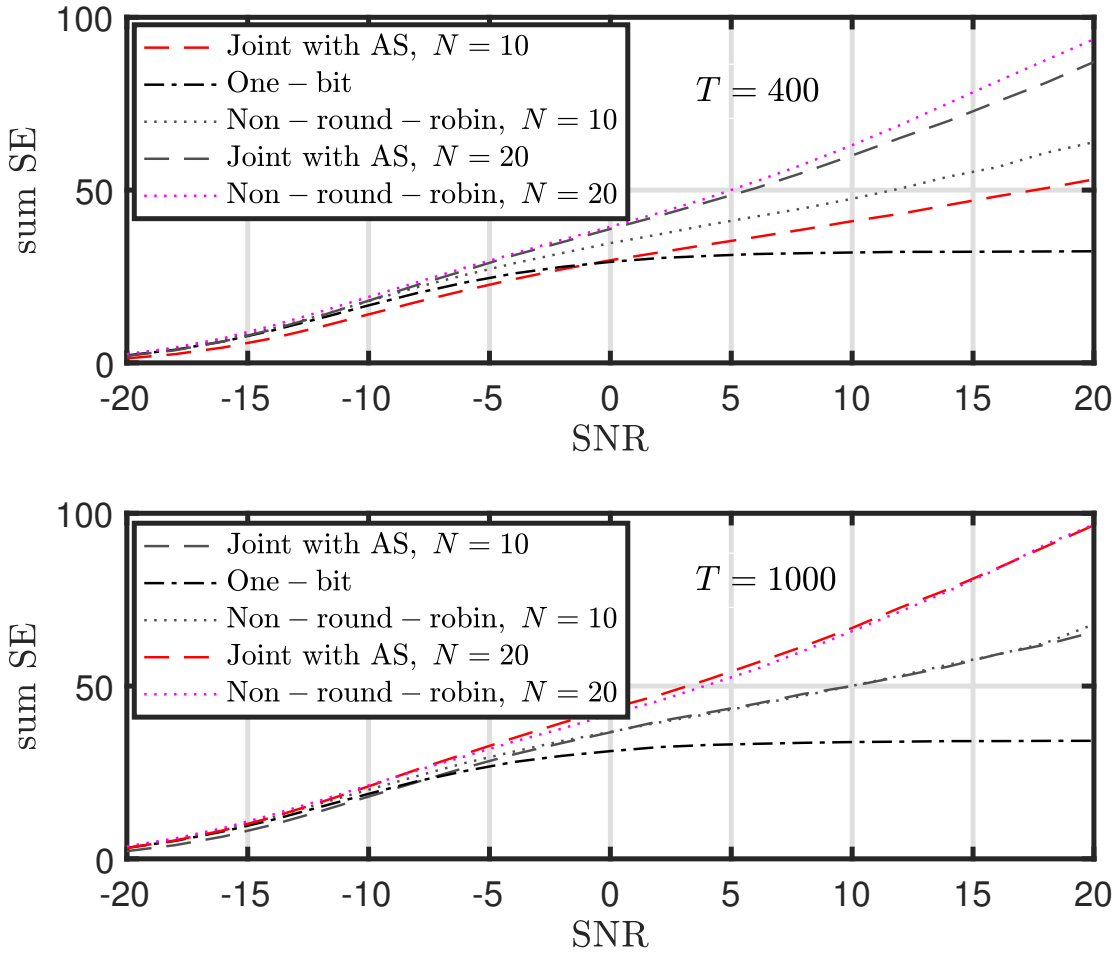

Figure 3.9: Sum SE for ZF detection versus SNR for $M = 100$, $N = 20,\ 10$, and $T = 400,\ 1000$.

mixed-ADC with and without round-robin training) depends on the SNR operating point and the length of the coherence interval. The advantage of round-robin training becomes apparent for long coherence intervals, where the increased training length has a smaller impact. The gain for round-robin training is greatest at higher SNRs. For shorter coherence intervals, mixed ADC with fixed antenna/ADC assignments provides the best SE, with the largest gains again coming at higher SNRs. For this value of N, the all-one-bit system generally has the lowest SE, although the difference is not large for MRC.

The next example investigates the impact of distributing the resolution (i.e., the comparators

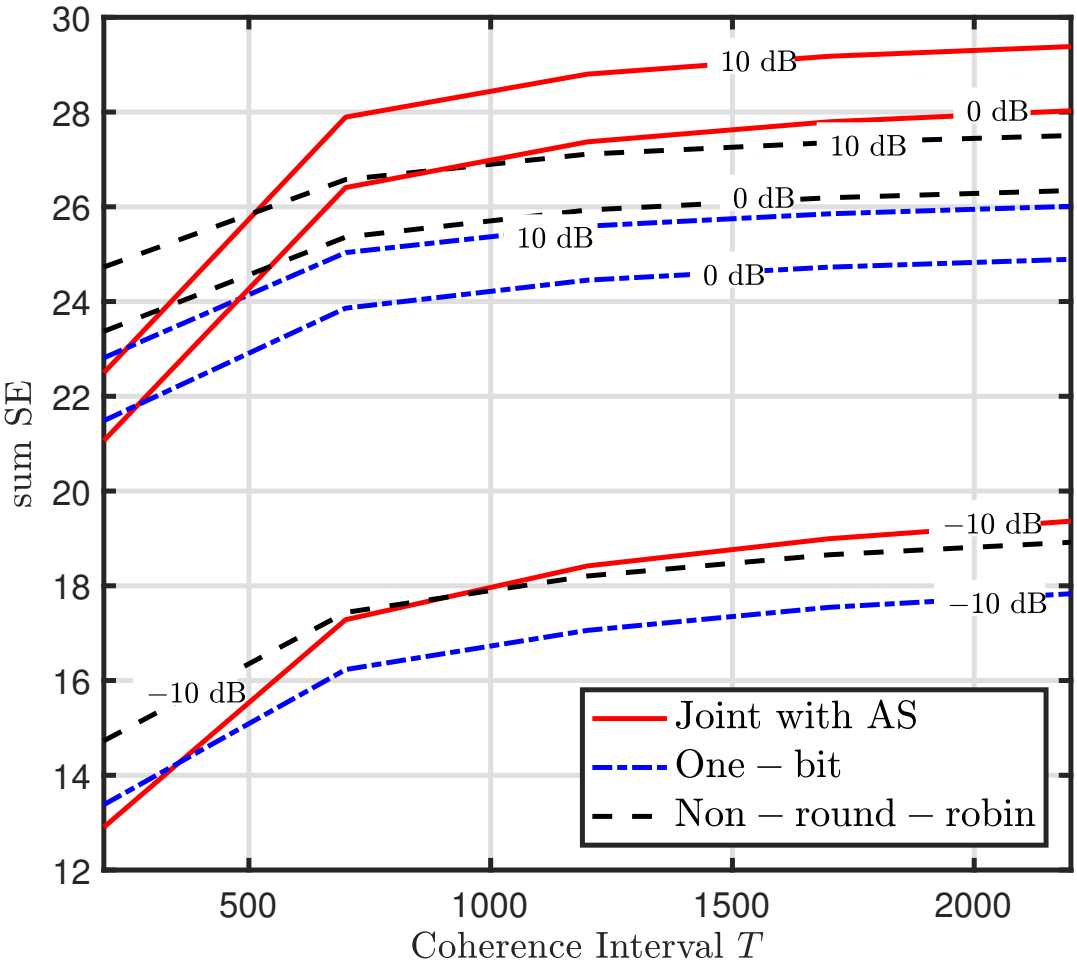

Figure 3.10: Sum SE for MRC detection versus T for $M = 100$, $N = 20$, and $SNR = -10, 0, 10$ dB.

of the ADCs) across the array with different numbers of antennas. If we assume that the "high-resolution" ADCs consist of 5 bits [14], a mixed-ADC architecture with $N = 20$ high-resolution and $M - N = 80$ one-bit ADCs will have 180 total comparators. Figs. 3.12 and 3.13 illustrate the SE achieved by distributing the 180 comparators across arrays of different length for MRC and ZF detection, respectively. In these figures, "Joint with AS" and "Non-round-robin" refer to mixed-ADC architectures with $N = 20$ 5-bit ADCs and $M - N = 80$ one-bit ADCs, "One-bit" corresponds to $M = 180$ antennas with one-bit ADCs, and "Multi-bit" indicates a system with either $M = 90$ 2-bit ADCs or $M = 60$ 3-bit ADCs. As we see

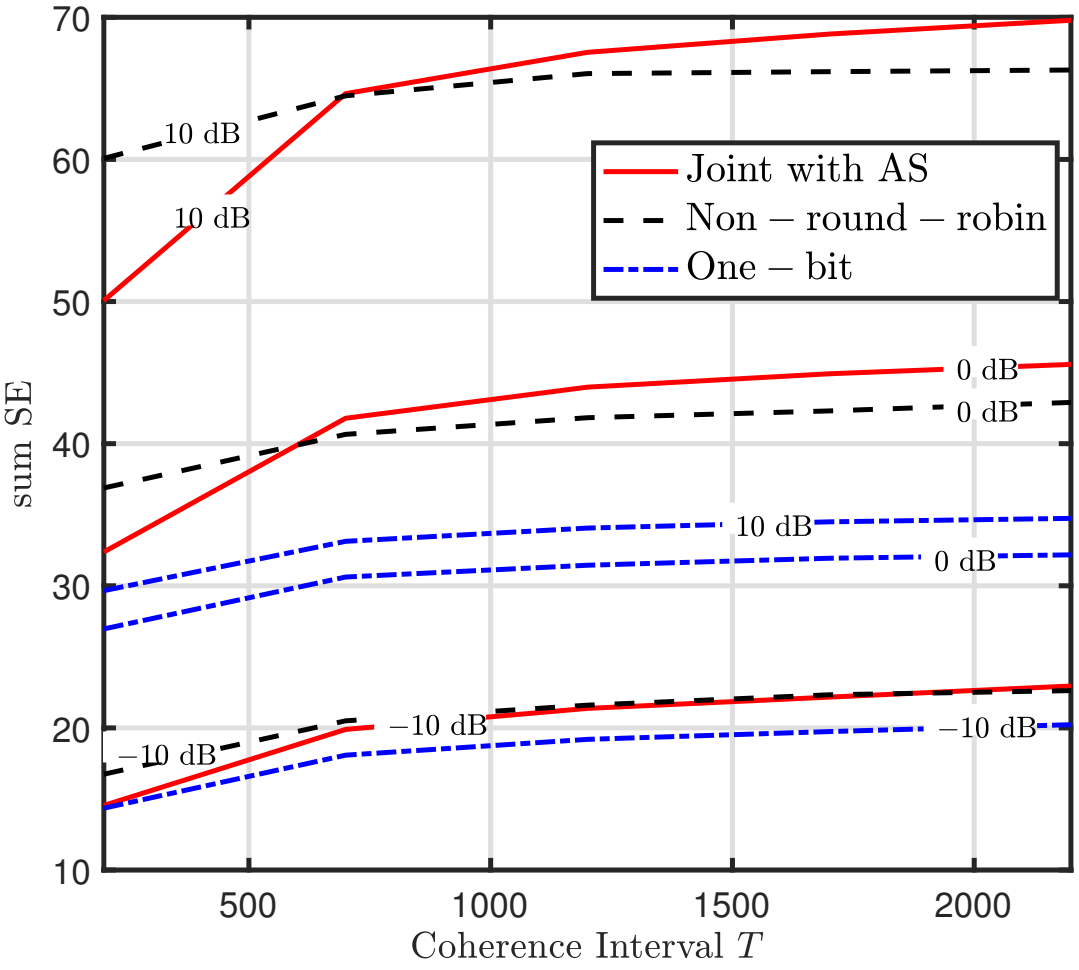

Figure 3.11: Sum SE for ZF detection versus T for $M = 100$, $N = 20$, and $SNR = -10, 0, 10$ dB.

in the figures, it can be inferred that for MRC detection, which is interference limited, it is better to have a larger number of antennas with lower-resolution ADCs instead of equipping the BS with fewer antennas and high resolution ADCs. This is consistent with the results of [31, 53], and is due to the fact that a larger number of antennas helps the system to more effectively cancel the interference. On the other hand, for ZF detection which is noise limited, the use of high-resolution ADCs avoids additional quantization noise imposed by the low-resolution ADCs, and is more beneficial than having a larger number of antennas with low-resolution ADCs at high SNR.

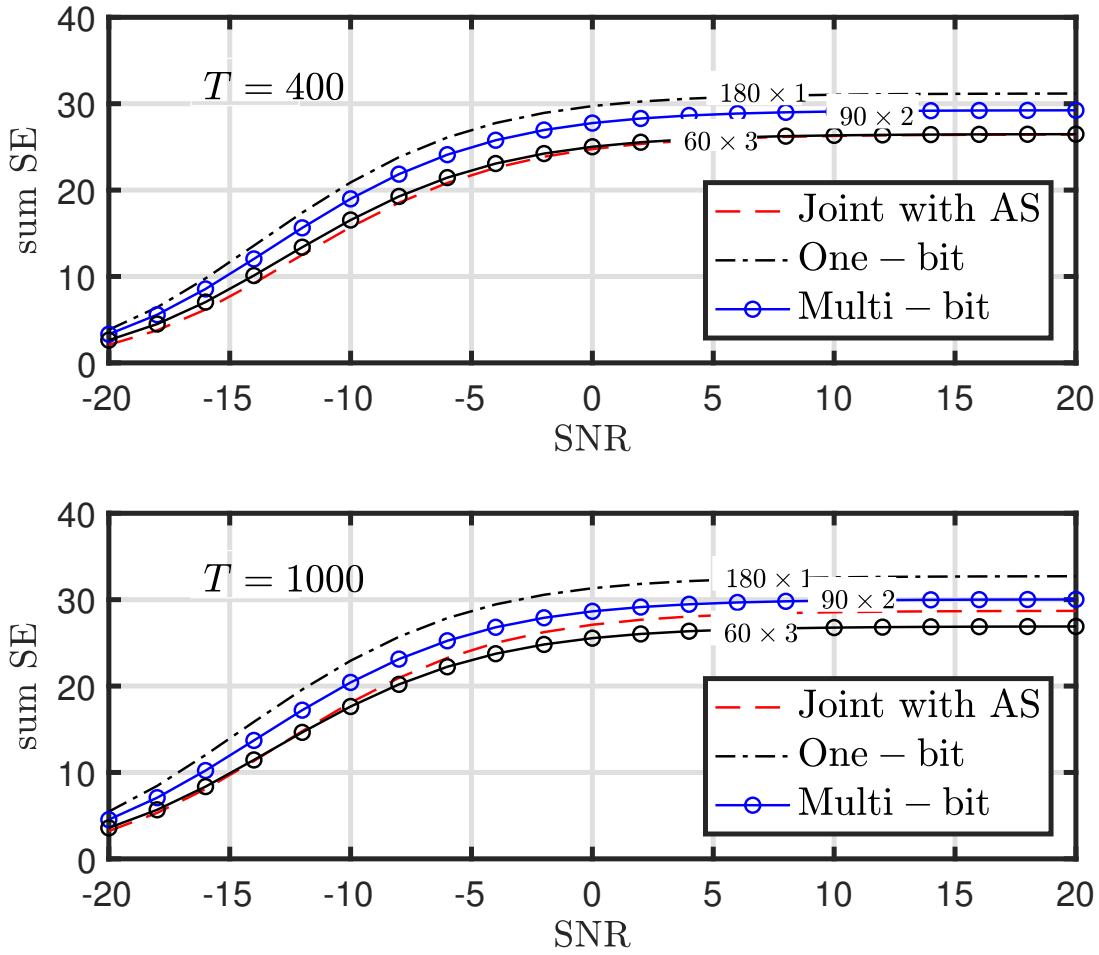

Figure 3.12: Sum SE for MRC detection versus SNR for 180 comparators and $T = 400,\ 1000$.

Finally, Figs. 3.14 and 3.15 show the impact of the number of high-resolution ADCs in a mixed-ADC system with $M = 100$ antennas, $K = 10$ users, and various numbers N of high-resolution ADCs, where $N = 100$ denotes the all-high-resolution system. It is apparent that with a large enough coherence interval and a sufficient number of high-resolution ADCs, the mixed-ADC implementation with joint round-robin channel estimation and antenna selection outperforms the all-one-bit architecture and mixed-ADC without round-robin training. The gains are greatest when ZF detection is used and the SNR is high, but such gains must be weighed against the increased power consumption and hardware complexity.

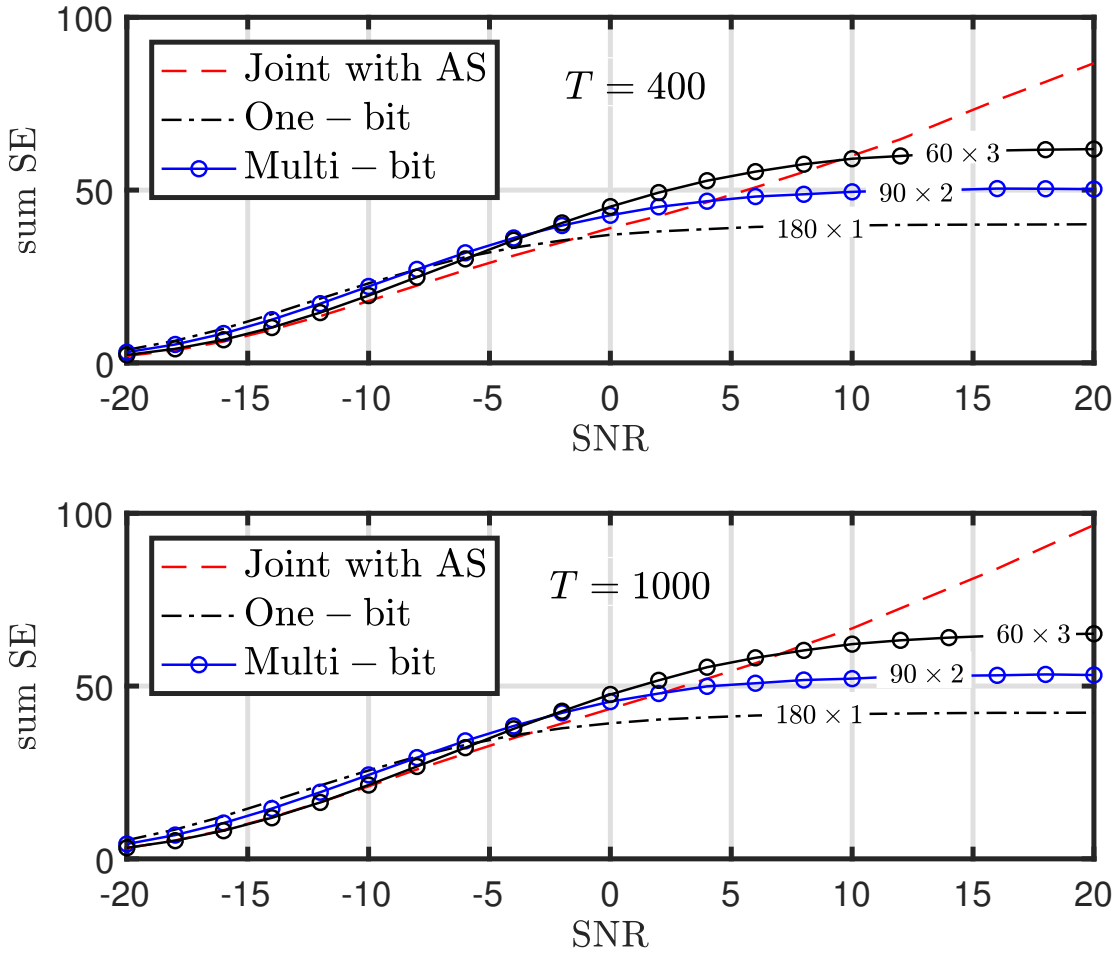

Figure 3.13: Sum SE for ZF detection versus SNR for 180 comparators and $T = 400,\ 1000$.

3.6 Summary

We studied the spectral efficiency of mixed-ADC massive MIMO systems with either MRC or ZF detection. We showed that properly accounting for the impact of the quantized receivers using the Bussgang decomposition is important for obtaining an accurate analysis of the SE. We introduced a joint channel estimation approach to leverage both high-resolution ADCs and one-bit ADCs and our analytical and numerical results confirmed the benefit of joint channel estimation for low SNRs.

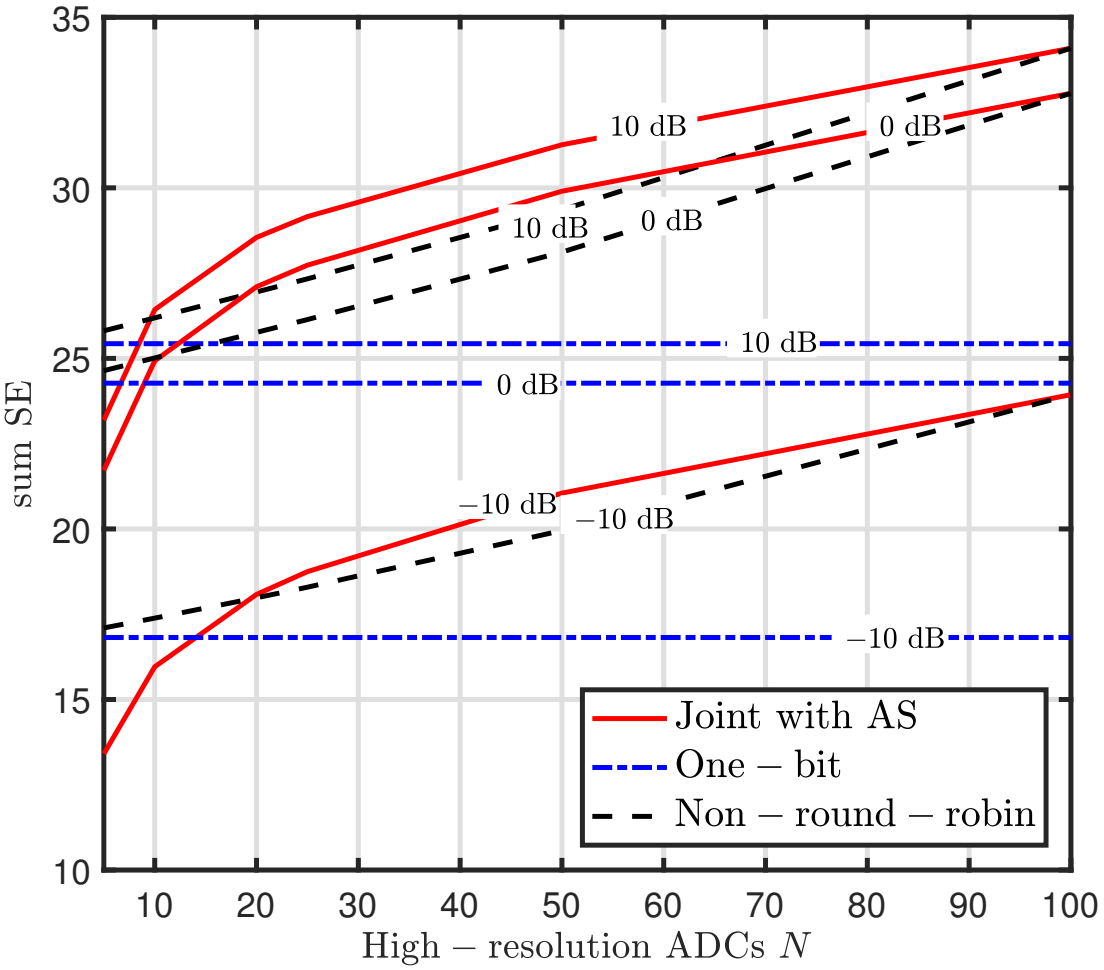

Figure 3.14: Sum SE for MRC detection versus N for $SNR = -10, 0, 10$ dB and $T = 1000$.

Mixed-ADC detection with MRC and ZF detectors and antenna selection were also studied. Analytical expressions were derived for MRC detection and a numerical performance analysis was performed for ZF detection. It was shown that antenna selection provides a slight advantage for high SNRs while this advantage tends to disappear for low SNRs.

We showed that the SNR, the number of high-resolution ADCs and the length of the coherence interval play a pivotal role in determining the performance of mixed-ADC systems. We showed that, in general, mixed-ADC architectures will have the greatest benefit compared to implementations with all low-resolution ADCs when ZF detection is used and the SNR

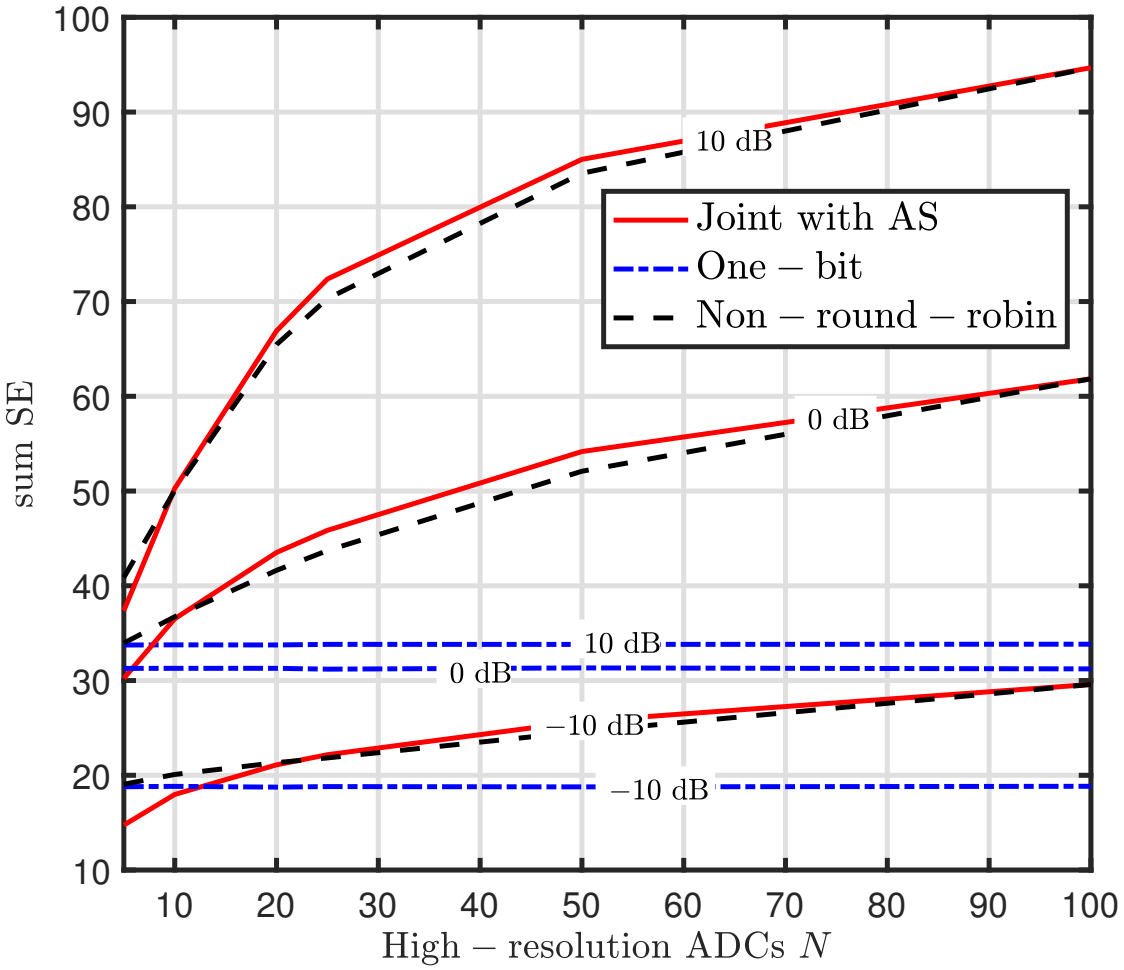

Figure 3.15: Sum SE for ZF detection versus N for $SNR = -10, 0, 10$ dB and $T = 1000$.

is relatively high. In such cases, the gain of the mixed-ADC approach can be substantial. Gains are also possible for MRC, but they are not as significant, and require larger numbers of high-resolution ADCs to see a benefit compared with the ZF case. The more complicated mixed-ADC approach based on ADC switching and round-robin training can achieve the best performance in some cases, particularly when the coherence interval is long and more high-resolution ADCs are available to reduce the number of training interval repetitions. Otherwise, a mixed-ADC implementation without ADC switching and extra training is preferred.

Chapter 4

Spectral Efficiency of One-Bit Sigma-Delta Massive MIMO

In this chapter, we study another approach in alleviating the performance loss due to coarse quantization, i.e., spatial $\Sigma\Delta$ architecture. In this method, using minimal additional hardware in the analog domain, the resulting spatial one-bit $\Sigma\Delta$ architecture can shape the quantization noise to angles of arrival away from those that correspond to the users of interest. Hence, performance close to that of systems with high-resolution ADCs can be achieved while reducing power consumption and complexity.

First, we elaborate the temporal $\Sigma\Delta$ operation. Then, we show how this idea can be extended to the spatial case. After introducing the spatial $\Sigma\Delta$ architecture, by proposing an appropriate linear model, we study the uplink spectral efficiency of a massive MIMO base station that employs one-bit spatial $\Sigma\Delta$ quantization, and compare it with the performance achievable by systems with infinite resolution and standard one-bit quantization for maximum ratio combining (MRC) and zero-forcing (ZF) receivers. Past work on quantifying the SE for standard one-bit quantization (*e.g.*, [15, 50]) has relied on a vectorized version of the

well-known Bussgang decomposition [49], which formulates an equivalent linear vector model for the array of non-linear quantizers assuming that the inputs to the quantizers are (at least approximately) jointly Gaussian. However, the vector Bussgang solution is not appropriate for the more complicated $\Sigma\Delta$ architecture, since it leads to a linear model that is inconsistent with the corresponding hardware implementation. Thus, we are led to derive an alternative linear model in which we apply a scalar version of the Bussgang approach to each quanitizer individually. This model is then used in turn to determine the overall sum SE.

The results of the analysis indicate the significant gain of the $\Sigma\Delta$ approach compared with standard one-bit quantization for users that lie in the angular sector where the shaped quantization error spectrum is low. For MRC, the one-bit $\Sigma\Delta$ array performs essentially the same for such users as a BS with infinite resolution ADCs. The angular sectorization of users in the spatial domain is not necessarily a drawback in cellular implementations, where cells are typically split into 120° regions using different arrays on the BS tower. In addition, there are many small-cell scenarios both indoors and outdoors where the targeted users are confined to relatively narrow angular sectors (auditoriums, plazas, arenas, etc.). Such situations will become even more prevalent as frequencies move to the millimeter wave band. However, the size of the sector of good performance for $\Sigma\Delta$ arrays depends on the amount of spatial oversampling. Unlike the temporal case, where oversampling factors of 10 or higher are not uncommon, the physical dimensions of the antenna and the loss due to increased mutual coupling for closely-spaced antennas places a limit on the amount of spatial oversampling that is possible in massive MIMO. Fortunately, our results indicate that spatial oversampling by factors of only 2-4 is sufficient to achieve good performance for angular sectors ranging from $80° - 150°$. Furthermore, the ability of the $\Sigma\Delta$ array to electronically steer the desired angular sector by means of the feedback phase shift provides desirable flexibility. For example, multiple sectors could be serviced in parallel with a single antenna array by deploying a bank of $\Sigma\Delta$ receivers tuned to different spatial frequencies, in order to cover a wider angular region.

At last, we introduce spatial feedback beamformer (FBB) $\Sigma\Delta$ architecture. This approach can be useful in mitigating the strong interference which can have detrimental impact on systems with one-bit quantizers.

In the next section we outline the basic system model, and provide some background on temporal $\Sigma\Delta$ modulation. In Section 4.2, we introduce the spatial $\Sigma\Delta$ architecture. We develop an equivalent linear model and characterize this architecture in Section 4.3. The model is then applied to analyze the spectral efficiency of the $\Sigma\Delta$ array in Section 4.4. FBB $\Sigma\Delta$ is discussed in section 4.5. While the analysis is conducted assuming that perfect channel state information (CSI) is available, we also discuss the impact of imperfect CSI in Section 4.6. Several simulation results are presented in Section 4.6, followed by the summary of this chapter.

4.1 System Model

Consider the uplink of a single-cell multi-user MIMO system consisting of K single-antenna users that send their signals simultaneously to a BS equipped with a uniform linear array (ULA) with M antennas. The $M \times 1$ signal received at the BS from the K users is given by

$$\boldsymbol{x} = \boldsymbol{G}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{s} + \boldsymbol{n}, \tag{4.1}$$

where $\boldsymbol{G} = [\boldsymbol{g}_1, \cdots, \boldsymbol{g}_K] \in \mathbb{C}^{M \times K}$ is the channel matrix between the users and the BS and $\boldsymbol{P}$ is a diagonal matrix whose kth diagonal element, p_k, represents the transmitted power of the kth user. The symbol vector transmitted by the users is denoted by $\boldsymbol{s} \in \mathbb{C}^{K \times 1}$ where $\mathbb{E}\left\{\boldsymbol{s}\boldsymbol{s}^H\right\} = \boldsymbol{I}_K$ and is drawn from a circularly symmetric complex Gaussian (CSCG) codebook independent of the other users, and, $\boldsymbol{n} \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I}_M\right)$ denotes additive CSCG receiver noise at the BS.

We consider a physical channel model described in the angular domain and comprised of L paths for each user with azimuth angular spread Θ [65]. In particular, for the kth user, the channel vector is modeled as

$$\boldsymbol{g}_k = \sqrt{\frac{\beta_k}{L}} \boldsymbol{A}_k \boldsymbol{h}_k, \tag{4.2}$$

where $\boldsymbol{A}_k$ is an $M \times L$ matrix whose ℓth column is the array steering vector corresponding to the direction of arrival (DoA) $\theta_{k\ell} \in \theta_0 + \left[-\frac{\Theta}{2}, \frac{\Theta}{2}\right]$, β_k models geometric attenuation and shadow fading from the kth user to the BS, and the elements of $\boldsymbol{h}_k \in \mathbb{C}^{L\times 1}$ are assumed to be distributed identically and independently as $\mathcal{CN}(0,1)$, and model the fast fading propagation. For a ULA, the steering vector for a signal with DoA $\theta_{k\ell}$ is expressed as

$$\boldsymbol{a}(u_{k\ell}) = \left[1, z_{k\ell}^{-1}, \cdots, z_{k\ell}^{-(M-1)}\right]^T, \tag{4.3}$$

where $u_{k\ell} = \sin(\theta_{kl})$, $z_{k\ell} = e^{j\omega_{s_{k\ell}}}$, and thus $\omega_{s_{k\ell}} = 2\pi\frac{d}{\lambda}u_{k\ell}$ represents the spatial frequency assuming antenna spacing d and wavelength λ.

In a standard implementation involving one-bit quantization, each antenna element at the BS is connected to a one-bit ADC. In such systems, the received baseband signal at the mth antenna becomes

$$y_m = Q_m(x_m) \ , \tag{4.4}$$

where $Q_m(.)$ denotes the one-bit quantization operation which is applied separately to the real and imaginary parts as

$$Q_m(x_m) = \alpha_{m,r}\text{sign}(\Re\mathfrak{e}(x_m)) + j\alpha_{m,i}\text{sign}(\Im\mathfrak{m}(x_m)) \ , \tag{4.5}$$

where $\alpha_{m,r}$ and $\alpha_{m,i}$ represent the output voltage levels of the one-bit quantizer. We will

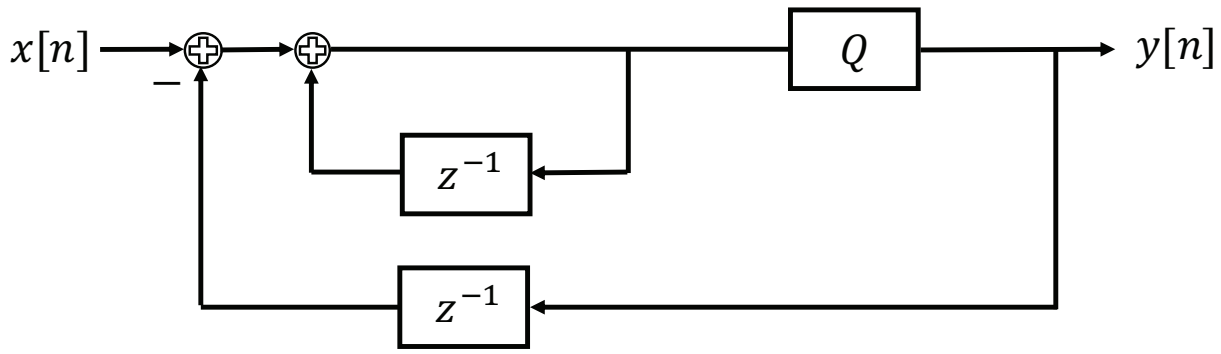

Figure 4.1: Block diagram for temporal ΣΔ modulator.

allow these levels to be a function of the antenna index m, unlike most prior work which assumes that the output levels are the same for all antennas. The necessity for this more general approach will become apparent later[1]. Finally, the received baseband signal at the BS is given by

$$\boldsymbol{y} = \mathcal{Q}(\boldsymbol{x}) = \left[\mathcal{Q}_1(x_1), \mathcal{Q}_2(x_2), \cdots, \mathcal{Q}_M(x_M)\right]^T. \tag{4.6}$$

4.2 ΣΔ Architecture

4.2.1 Temporal ΣΔ Modulation

In this subsection, we elaborate on temporal ΣΔ modulation to clarify the noise shaping characteristics of this technique.

Fig. 4.1 shows a block diagram representing the temporal ΣΔ modulator. To shape the quantization noise, the output signal is fed back and subtracted form the input (Δ-stage), and then this error is integrated (Σ-stage). To characterize the transfer function of this non-linear system, we substitute the one-bit quantizer with the equivalent linear model depicted

[1]While the one-bit ADC output levels will be optimized, this is a one-time optimization and the values do not change as a function of the user scenario or channel realization. Thus the ADCs are still truly "one-bit."

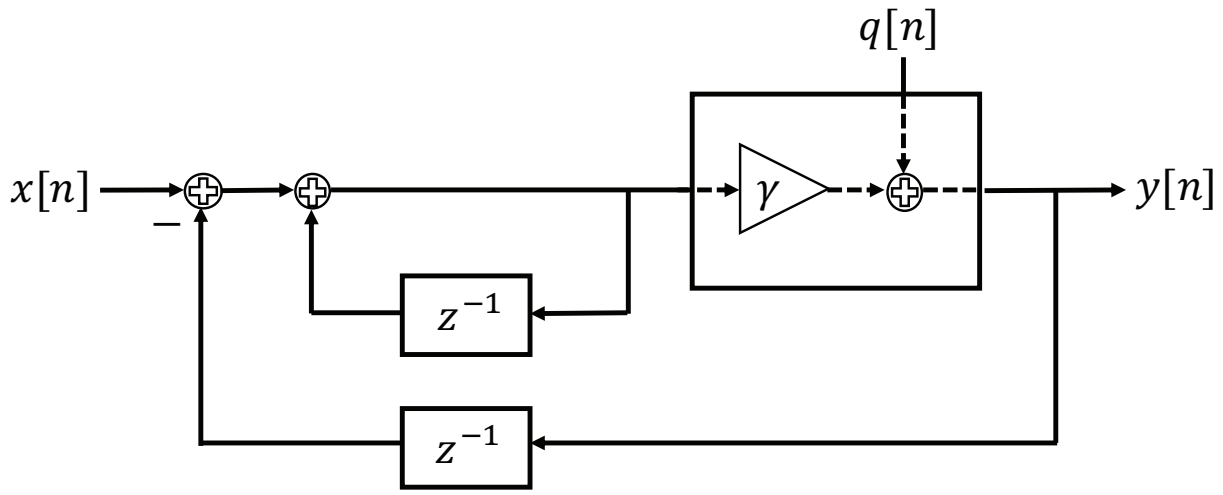

Figure 4.2: Block diagram for temporal ΣΔ modulator with equivalent linear model for quantization.

in Fig. 4.2 . The input-output relationship of the ΣΔ quantizer can then be written as

$$Y(z) = \frac{\gamma}{1-(1-\gamma)z^{-1}}X(z) + \frac{\left(1-z^{-1}\right)}{1-(1-\gamma)z^{-1}}Q(z), \tag{4.7}$$

where $X(z) = \sum_{n=0}^{\infty} x[n]\, z^{-n}$ denotes the z-transform. Simply stated, the objective of ΣΔ modulation is to pass the signal through an all-pass filter and the quantization noise through a high-pass filter. This objective can be realized by setting $\gamma \approx 1$. Since commercial quantizers are provided with a built-in automatic gain control (AGC), the $\gamma \approx 1$ condition is inherently satisfied in implementations of temporal ΣΔ modulators, and hence this issue is not generally discussed in the literature. However, as we show in the next subsection, the choice of the scaling factor is critical in the mathematical modeling of spatial ΣΔ architectures, and we derive a criterion for addressing this issue.

4.2.2 One-Bit Spatial ΣΔ Modulation

As mentioned earlier, the basic premise of temporal ΣΔ modulation can be adopted in the angle domain, in order to *spatially* shape the quantization noise in a desired way. Instead of

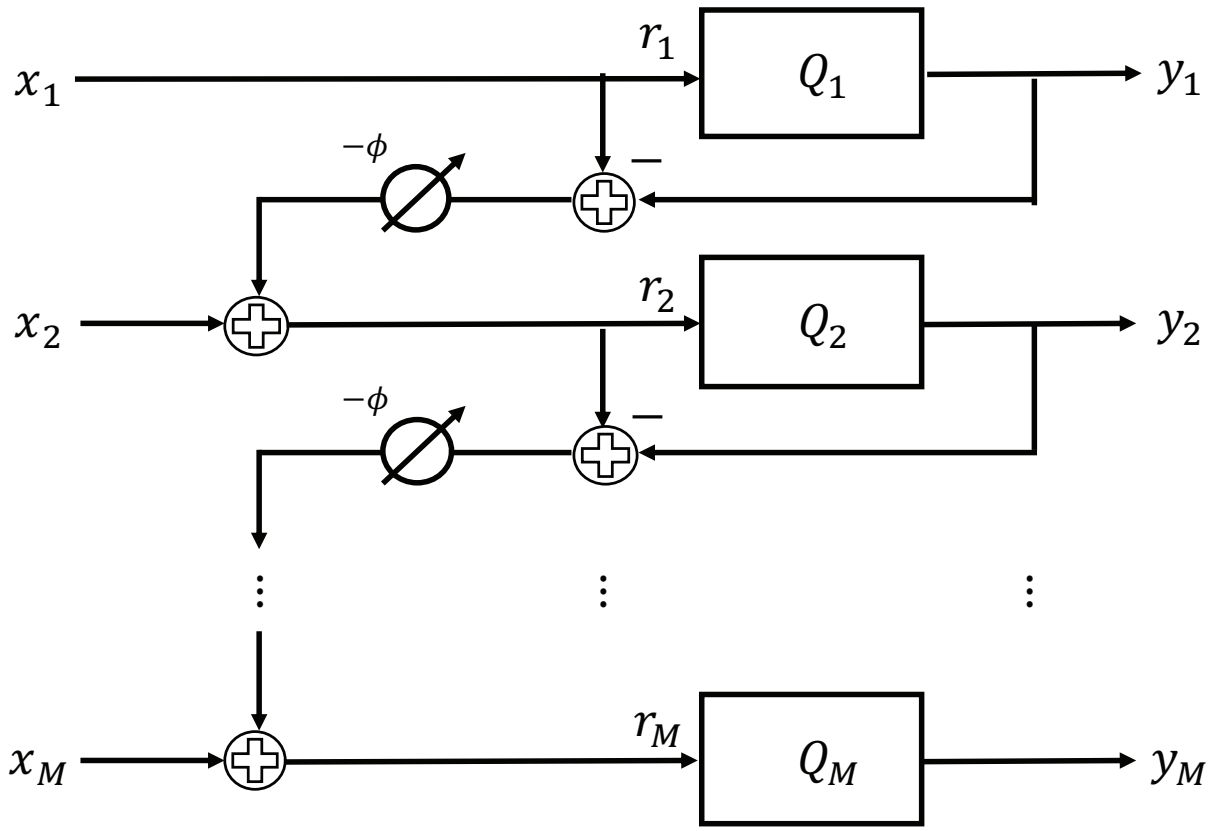

Figure 4.3: Spatial ΣΔ architecture.

forming the Δ component using a delayed sample of the quantized input as in the temporal case, we use the quantization error signal from an adjacent antenna. A direct transfer of the temporal ΣΔ idea to the angle domain as in [45, 46] pushes the quantization noise to higher spatial frequencies, which correspond to DoAs away from the array broadside ($|\theta| \gg 0°$), while the oversampling (reduced d/λ) pushes signals of interest near broadside closer to zero spatial frequency. However, by phase-shifting the quantization error in the feedback loop prior to the Δ stage, a ΣΔ frequency response can be obtained in which the quantization error is shaped away from a band of frequencies not centered at zero. This bandpass approach has been proposed for both the temporal (*e.g.,* see [37]) and spatial [48] versions of the ΣΔ architecture.

Fig. 4.3 shows the architecture of an angle-steered ΣΔ array. Using Fig. 4.3 and

$$y_m = \begin{cases} Q_1(x_1) & m = 1 \\ Q_m\left(x_m + e^{-j\phi}\left(x_{m-1} + e^{-j\phi}\left(\cdots\left(x_2 + e^{-j\phi}(x_1 - y_1) - y_2\right)\cdots\right) - y_{m-1}\right)\right) & m > 1 \end{cases} \tag{4.8}$$

we can formulate a compact input-output description of the spatial ΣΔ array by defining

$$\boldsymbol{U} = \begin{bmatrix} 1 & & & \\ e^{-j\phi} & 1 & & \\ \vdots & \ddots & \ddots & \\ e^{-j(M-1)\phi} & \cdots & e^{-j\phi} & 1 \end{bmatrix} \tag{4.9}$$

$$\boldsymbol{V} = \boldsymbol{U} - \boldsymbol{I}_M, \tag{4.10}$$

and expressing the input to the quantizers as

$$\boldsymbol{r} = \boldsymbol{U}\boldsymbol{x} - \boldsymbol{V}\boldsymbol{y}. \tag{4.11}$$

The output of the angle-steered one-bit ΣΔ array is then defined by

$$\boldsymbol{y} = \boldsymbol{Q}(\boldsymbol{r}). \tag{4.12}$$

4.3 Characterizing the Spatial ΣΔ Architecture

4.3.1 Linear Model

To analyze the performance of spatial ΣΔ processing, analogous to temporal ΣΔ, we will represent the one-bit quantization operation in (4.12) with an equivalent linear model as follows:

$$\boldsymbol{y} = Q(\boldsymbol{r}) = \boldsymbol{\Gamma}\boldsymbol{r} + \boldsymbol{q}, \tag{4.13}$$

where $\boldsymbol{\Gamma}$ is an $M \times M$ matrix and $\boldsymbol{q}$ denotes the effective quantization noise. The value of $\boldsymbol{\Gamma}$ that makes the equivalent quantization noise, $\boldsymbol{q}$, uncorrelated with $\boldsymbol{r}$ is $\boldsymbol{\Gamma}_0 = \boldsymbol{R}_{\boldsymbol{ry}}^H \boldsymbol{R}_{\boldsymbol{r}}^{-1}$. For the case where the elements of $\boldsymbol{r}$ are all jointly Gaussian, the computation of $\boldsymbol{R}_{\boldsymbol{ry}}$ is possible by resorting to the Bussgang theorem[2] [49]. This was the approach used in [15, 50] for a massive MIMO implementation with standard one-bit quantization, and the resulting $\boldsymbol{\Gamma}_0$ was a diagonal matrix.

For the case of the ΣΔ architecture, even if the matrix $\boldsymbol{\Gamma}_0$ could be computed, this decomposition would not be of interest, for at least two reasons. First, the equivalent quantization noise $\boldsymbol{q}$ that results from setting $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_0$ in (4.13) bears no connection to the quantization error fed from one antenna to the next as shown in Fig. 4.3. Setting $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_0$ would produce a model in which r_m and q_{m-1} are uncorrelated, but it is clear from Fig. 4.3 that r_m for the ΣΔ array directly depends on the quantization error from the $(m-1)$-th stage. Second, $\boldsymbol{\Gamma}_0$ cannot be a diagonal matrix[3], unlike the standard one-bit quantization case considered in [15]. The presence of off-diagonal elements in $\boldsymbol{\Gamma}_0$ implies that the model in (4.13) represents

[2]The result can also be extended to cases where $\boldsymbol{r}$ belongs to a limited class of distributions, see [66] for details

[3] If $\boldsymbol{\Gamma}_0$ were diagonal, it could be made equal to the identity matrix by a proper scaling of each y_m. However, $\boldsymbol{\Gamma}_0$ can never be the identity matrix because this implies that $r_m = x_m - e^{-j\phi} q_{m-1}$, while simultaneously r_m is uncorrelated with q_{m-1}, which is impossible.

the output of each quantizer as a linear combination of the inputs to that quantizer as well as other quantizers in the array. Such a model does not have an apparent connection with the scheme in Fig. 4.3, where each quantizer produces its output depending only on its input alone. These inconsistencies between the mathematical model based on $\mathbf{\Gamma} = \mathbf{\Gamma}_0$ and the physical block diagram of the $\Sigma\Delta$ array in Fig. 4.3 are the result of attempting to force $\boldsymbol{r}$ and $\boldsymbol{q}$ to be uncorrelated, when the architecture is actually propagating the quantization error from one stage to the next.

Consequently, in order to derive an appropriate model for the analysis of the $\Sigma\Delta$ architecture, we propose to apply the Bussgang decomposition to each quantizer individually. In particular, we formulate the model in (4.13) using a matrix $\mathbf{\Gamma} = \text{diag}\,(\gamma_1, \ldots, \gamma_M)$ that is forced to be diagonal. This is equivalent to imposing a model in which $\boldsymbol{r}$ and $\boldsymbol{q}$ are uncorrelated component-wise: $\mathbb{E}\left[r_m q_m^*\right] = 0$. The elements of $\mathbf{\Gamma}$ are given by

$$\gamma_m = \frac{\mathbb{E}\left[r_m y_m^*\right]}{\mathbb{E}\left[|r_m|^2\right]} = \alpha_m \frac{\mathbb{E}\left[|\mathfrak{Re}\left[r_m\right]| + |\mathfrak{Im}\left[r_m\right]|\right]}{\mathbb{E}\left[|r_m|^2\right]}, \tag{4.14}$$

where in the last equality and from now on, we assume that r_m is circularly symmetric. This assumption implies that the quantizer output levels are identical for the real and imaginary parts, and thus we use α_m to represent both $\alpha_{m,r}$ and $\alpha_{m,i}$. The expression in (4.14) has been defined in the literature as the equivalent gain of a non-linear device [66, 67].

As we will see later on, since the elements of $\mathbf{\Gamma}$ depend only on the signals at one stage of the $\Sigma\Delta$ architecture, they are much easier to compute than the elements of $\mathbf{\Gamma}_0$. Moreover, the resulting decomposition is consistent with Fig. 4.3. Given that no precondition is imposed on the correlation $\mathbb{E}\left[r_m q_l^*\right]$ for $m \neq l$, the model is compatible with the fact that the quantization noise of one stage appears in subsequent stages.

Plugging (4.13) into (4.11) and using some algebraic manipulations, we obtain the following

mathematical model for the ΣΔ architecture:

$$\boldsymbol{y} = \left(\boldsymbol{I} + \boldsymbol{\Gamma V}\right)^{-1}\boldsymbol{\Gamma U x} + \left(\boldsymbol{I} + \boldsymbol{\Gamma V}\right)^{-1}\boldsymbol{q}. \tag{4.15}$$

Equation (4.15) is the spatial ΣΔ equivalent to the temporal domain ΣΔ description in (4.7). Similar to the temporal case, (4.80) indicates that $\boldsymbol{\Gamma} = \boldsymbol{I}$ should hold for the spatial ΣΔ array to work as desired, that is, to pass $\boldsymbol{x}$ and $\boldsymbol{q}$ through spatial all-pass and high-pass filters, respectively. If $\boldsymbol{\Gamma} = \boldsymbol{I}$, then (4.80) becomes

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{U}^{-1}\boldsymbol{q}\ , \tag{4.16}$$

and the m-th element of $\boldsymbol{y}$ is expressed as

$$y_m = x_m + \left(q_m - e^{-j\phi}q_{m-1}\right)\ , \tag{4.17}$$

which explicitly shows the quantization noise-shaping characteristic of the spatial ΣΔ architecture. The only task remaining to complete our proposed linear model is to calculate the power of the equivalent quantization noise. The condition $\boldsymbol{\Gamma} = \boldsymbol{I}$ for the adequate operation of the ΣΔ scheme determines the quantization levels that have to be set. Setting (4.14) equal to 1, we obtain the optimum value of α_m:

$$\alpha_m^{\star} = \frac{\mathbb{E}\left[|r_m|^2\right]}{\mathbb{E}\left[|\Re\mathrm{e}\left[r_m\right]| + |\Im\mathrm{m}\left[r_m\right]|\right]} = \frac{\mathbb{E}\left[|\Re\mathrm{e}\left[r_m\right]|^2\right]}{\mathbb{E}\left[|\Re\mathrm{e}\left[r_m\right]|\right]}. \tag{4.18}$$

It is worth noting that (4.18) is different from

$$\alpha_m = \mathbb{E}\left[|\Re\mathrm{e}\left[r_m\right]|\right], \tag{4.19}$$

which leads to the Lloyd-Max one-bit quantizer that minimizes the mean-squared-error

(MSE) between the input and the output of the quantizer. However, the Lloyd-Max approach makes the quantization error uncorrelated with the quantizer output, but not with the input.

While the expression derived in (4.18) is useful, it is difficult to analytically evaluate the expectations in closed form, and it is not clear how the output level could be tuned using analog processing in the RF chain (e.g., via an AGC or some other type of calibration). To address this issue, we use the assumption that r_m is Gaussian inherent in the Bussgang decomposition to find an approximation for $\alpha_m^\star$ that is easier to deal with, both for the subsequent mathematical analysis and from the viewpoint of a hardware implementation. The validity of the approximation will be apparent in the numerical examples presented later. If r_m is Gaussian, we can write

$$\alpha_m^\star = \frac{\sqrt{\pi \mathbb{E}\left[|r_m|^2\right]}}{2}. \tag{4.20}$$

In the discussion below, we show how to express (4.20) in terms of the statistics of the array output $\boldsymbol{x}$, which provides an analytical solution and clarifies how the quantizer output levels could be set in a practical setting.

4.3.2 Quantization Noise Power

In this section, we calculate the power of the effective quantization noise and the power of the quantizers' inputs, which is needed to properly set the output levels using (4.20). With $\boldsymbol{\Gamma} = \boldsymbol{I}$, (4.13) becomes

$$\boldsymbol{y} = \boldsymbol{r} + \boldsymbol{q}. \tag{4.21}$$

Since r_m and q_m are uncorrelated, and using (4.20), we obtain

$$\mathbb{E}\left[|q_m|^2\right] = \mathbb{E}\left[|y_m|^2\right] - \mathbb{E}\left[|r_m|^2\right] = \left(\frac{\pi}{2} - 1\right)\mathbb{E}\left[|r_m|^2\right]. \tag{4.22}$$

To determine $\mathbb{E}\left[|r_m|^2\right]$, we substitute (4.21) into (4.11), so that

$$\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{U}^{-1}\boldsymbol{V}\boldsymbol{q}. \tag{4.23}$$

It can be shown that

$$\boldsymbol{U}^{-1}\boldsymbol{V} = e^{-j\phi}\boldsymbol{Z}_{-1}\,, \tag{4.24}$$

where[4]

$$\boldsymbol{Z}_{-1} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ \vdots & \ddots & \ddots & \\ 0 & \ddots & 1 & 0 \end{bmatrix}. \tag{4.25}$$

Moreover, following the same reasoning as in Appendix A of [15], it can be shown that $\mathbb{E}\left[x_{m'}q_m^*\right] \approx 0,\ \forall m, m' \in \mathcal{M} = \{1, \cdots, M\}$. This results in $\boldsymbol{R}_{\boldsymbol{qx}} \approx \boldsymbol{0}$. Therefore,

$$\boldsymbol{R}_{\boldsymbol{r}} = \boldsymbol{R}_{\boldsymbol{x}} + \boldsymbol{Z}_{-1}\boldsymbol{R}_{\boldsymbol{q}}\boldsymbol{Z}_{-1}^{H}\,. \tag{4.26}$$

Eq. (4.26) implies that

$$\mathbb{E}\left[|r_m|^2\right] = \begin{cases} \mathbb{E}\left[|x_m|^2\right] & m = 1 \\ \mathbb{E}\left[|x_m|^2\right] + \mathbb{E}\left[|q_{m-1}|^2\right] & m > 1 \end{cases} \tag{4.27}$$

[4]Note that $\boldsymbol{Z}_{-1}$ is the spatial domain equivalent of the delay operator z^{-1} for the z-transform in the time domain.

Substituting (4.22) into (4.27) and noting that $\mathbb{E}\left[|r_1|^2\right] = \mathbb{E}\left[|x_1|^2\right]$, we obtain the following recursive equality to calculate $\mathbb{E}\left[|r_m|^2\right]$ for $m > 1$:

$$\mathbb{E}\left[|r_m|^2\right] = \mathbb{E}\left[|x_m|^2\right] + \left(\frac{\pi}{2} - 1\right)\mathbb{E}\left[|r_{m-1}|^2\right]. \tag{4.28}$$

Let

$$\boldsymbol{p}_{\chi} = \left[\mathbb{E}\left[|\chi_1|^2\right], \mathbb{E}\left[|\chi_2|^2\right], \cdots, \mathbb{E}\left[|\chi_M|^2\right]\right]^T, \tag{4.29}$$

where $\boldsymbol{\chi}$ can be any element of the set $\boldsymbol{\chi} \in \{\boldsymbol{r}, \boldsymbol{x}, \boldsymbol{q}\}$. Then, using (4.22) and (4.28), we have

$$\boldsymbol{p}_{\boldsymbol{r}} = \boldsymbol{\Pi}\boldsymbol{p}_{\boldsymbol{x}} \tag{4.30}$$

$$\boldsymbol{p}_{\boldsymbol{q}} = \left(\frac{\pi}{2} - 1\right)\boldsymbol{\Pi}\boldsymbol{p}_{\boldsymbol{x}}, \tag{4.31}$$

where

$$\boldsymbol{\Pi} = \begin{bmatrix} 1 & & & & & \mathbf{0} \\ \left(\frac{\pi}{2}-1\right) & 1 & & & & \\ \vdots & \ddots & 1 & & & \\ \left(\frac{\pi}{2}-1\right)^m & \ddots & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ \left(\frac{\pi}{2}-1\right)^{M-1} & \cdots & \left(\frac{\pi}{2}-1\right)^m & \cdots & \left(\frac{\pi}{2}-1\right) & 1 \end{bmatrix}. \tag{4.32}$$

Equation (4.30) shows that the calculation of $\mathbb{E}\left[|r_m|^2\right]$ needed in (4.20) can be formulated in terms of the power of the antenna outputs $\mathbb{E}\left[|x_m|^2\right]$, for which simple expressions exist from (4.1). This further implies that control of $\mathbb{E}\left[|x_m|^2\right]$ via an AGC would allow the quan-

tizer output levels to be set without feedback from the digital baseband. In the following remark, we show that, using the optimal quantizer output settings, the power of the quantization noise does not grow with m despite the fact that it is propagated from one antenna to the next.

Remark 4.1. *Eq. (4.31) implies that, by appropriately selecting the quantizers' output levels, the quantization noise power does not increase without bound. In particular, consider the case where the power of the received signal is constant over the array elements, i.e.,* $\boldsymbol{p_x} = p_x \mathbf{1}$. *Then,*

$$\mathbb{E}\left[|q_m|^2\right] = \left(\frac{\pi}{2} - 1\right) \frac{1 - \left(\frac{\pi}{2} - 1\right)^m}{1 - \left(\frac{\pi}{2} - 1\right)} p_x \xrightarrow[m\to\infty]{} \frac{\frac{\pi}{2} - 1}{2 - \frac{\pi}{2}} p_x \,, \tag{4.33}$$

which shows that, in the limit of a large number of antenna elements, the quantization noise power converges to a constant value of approximately 1.33 times the input power.

4.3.3 Quantization Noise Power Density

In the time domain, it is well-known that sampling a band-limited signal by a rate N times larger than the Nyquist rate and down-sampling after quantization can reduce the in-band quantization noise power by a factor of $1/N$ and $1/N^3$ for standard one-bit and $\Sigma\Delta$ modulation, respectively [68]. In this subsection, we look for a similar behaviour for quantization across an array in space. More precisely, we want to quantify how spatial oversampling, i.e., decreasing the antenna spacing, d/λ, (or equivalently, increasing the number of antennas for space-constrained arrays) can reduce the quantization noise power for the in-band angular spectrum. To do so, we define the quantization noise power density as

$$\rho_q(u) \triangleq \frac{1}{M} \boldsymbol{a}(u)^H \boldsymbol{R}\ \boldsymbol{a}(u), \tag{4.34}$$

where $\boldsymbol{R}$ is the covariance matrix of the quantization noise. To differentiate the two cases, we denote the covariance matrix of the quantization noise for standard one-bit quantization as $\boldsymbol{R}_{\boldsymbol{q}_1}$, and the covariance of the $\Sigma\Delta$ quantization noise as $\boldsymbol{R}_{\boldsymbol{q}_{\Sigma\Delta}}$. Expressions for these covariance matrices will be derived later in this subsection. Hence, the normalized received quantization noise power over some angular region, Θ, is given by[5]

$$\mathcal{P}_q = \frac{1}{2\delta} \int_{-\delta}^{\delta} \rho_q(u)\, du, \tag{4.35}$$

where $\delta = \sin\left(\frac{\Theta}{2}\right)$. Next we find $\mathcal{P}_q$ for standard one-bit and $\Sigma\Delta$ quantization.

4.3.3.1 One-bit Quantization

Unlike [15], for standard one-bit quantization, we choose the quantizer output levels as $\alpha_m = \sqrt{\pi \mathbb{E}\left[|x_m|^2\right]}/2$ so that $y_m = Q(x_m) = x_m + q_m$. This causes no loss of generality for standard one-bit quantization, since the value of the quantizer output has no impact on the performance of the resulting system. Therefore, the covariance matrix of the quantization noise can be written as

$$\boldsymbol{R}_{\boldsymbol{q}_1} = \boldsymbol{R}_{\boldsymbol{y}} - \boldsymbol{R}_{\boldsymbol{x}}\,, \tag{4.36}$$

where the arc-sine law [52, 69] is used to obtain

$$\boldsymbol{R}_{\boldsymbol{y}} = \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{x}})^{\frac{1}{2}} \sin^{-1}(\boldsymbol{\Lambda})\ \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{x}})^{\frac{1}{2}}, \tag{4.37}$$

and

$$\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{x}})^{-\frac{1}{2}}\, \Re\mathfrak{e}(\boldsymbol{R}_{\boldsymbol{x}})\, \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{x}})^{-\frac{1}{2}} + j \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{x}})^{-\frac{1}{2}}\, \Im\mathfrak{m}(\boldsymbol{R}_{\boldsymbol{x}})\, \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{x}})^{-\frac{1}{2}}\,. \tag{4.38}$$

[5]To simplify the calculation of the quantization noise power, we assume without loss of generality that the $\Sigma\Delta$ array is steered to broadside ($\theta = 0$).

Note that the arc-sine in (4.37) is applied separately to each element of the matrix argument, and also separately to the real and imaginary parts of the matrix elements.

From [15], we have that $\operatorname{diag}\left(\boldsymbol{R}_{\boldsymbol{y}}\right) = \frac{\pi}{2}\operatorname{diag}\left(\boldsymbol{R}_{\boldsymbol{x}}\right)$. Since the off-diagonal elements of $\boldsymbol{\Lambda}$ are small, we use the approximation $\sin^{-1}(x) \approx \zeta x$, where $\zeta > 1$, to obtain

$$\boldsymbol{R}_{\boldsymbol{q}_1} \approx (\zeta - 1)\boldsymbol{R}_{\boldsymbol{x}} + \left(\frac{\pi}{2} - \zeta\right)\operatorname{diag}\left(\boldsymbol{R}_{\boldsymbol{x}}\right). \tag{4.39}$$

Moreover, from (4.1), $\boldsymbol{R}_{\boldsymbol{x}}$ becomes

$$\boldsymbol{R}_{\boldsymbol{x}} = \sum_{k=1}^{K} p_k \beta_k \frac{1}{L} \sum_{\ell=1}^{L} \boldsymbol{a}\left(u_{k\ell}\right) \boldsymbol{a}\left(u_{k\ell}\right)^H + \sigma_n^2 \boldsymbol{I}, \tag{4.40}$$

where for $L \gg 1$, $u_{k\ell}$ can be taken as a random variable uniformly distributed in $[-\delta, \delta]$. That is,

$$\frac{1}{L} \sum_{\ell=1}^{L} \boldsymbol{a}\left(u_{k\ell}\right) \boldsymbol{a}\left(u_{k\ell}\right)^H \approx \mathbb{E}\left[\boldsymbol{a}(u)\, \boldsymbol{a}(u)^H\right] = \frac{1}{2\delta} \int_{-\delta}^{\delta} \boldsymbol{a}(u)\, \boldsymbol{a}(u)^H \, du. \tag{4.41}$$

Therefore,

$$\boldsymbol{R}_{\boldsymbol{x}} = \sum_{k=1}^{K} p_k \beta_k \frac{1}{2\delta} \int_{-\delta}^{\delta} \boldsymbol{a}(u)\, \boldsymbol{a}(u)^H \, du + \sigma_n^2 \boldsymbol{I}. \tag{4.42}$$

Now we are ready to calculate the standard one-bit quantization noise power, $\mathcal{P}_{q_1}$.

Proposition 4.1. *The normalized quantization noise power for standard one-bit quantization is*

$$\mathcal{P}_{q_1} = (\zeta - 1) \times \left[\sigma_n^2 + \frac{1}{M} \sum_{k=1}^{K} p_k \beta_k \sum_{n=0}^{M-1} \sum_{m=0}^{M-1} \operatorname{sinc}^2\left(2\pi \frac{d}{\lambda}(m-n)\,\delta\right)\right] + \frac{\frac{\pi}{2} - \zeta}{M} \operatorname{Tr}\left[\boldsymbol{R}_{\boldsymbol{x}}\right], \tag{4.43}$$

where $\operatorname{sinc}(x) \triangleq \frac{\sin(x)}{x}$.

Proof. Plugging (4.42) into (4.39) results in

$$\mathcal{P}_{q_1} = (\zeta - 1) \times \left[\sigma_n^2 + \frac{1}{4\delta^2 M} \sum_{k=1}^{K} p_k \beta_k \iint_{-\delta}^{\delta} \left| \boldsymbol{a}(v)^H \boldsymbol{a}(u) \right|^2 du dv \right] + \frac{\frac{\pi}{2} - \zeta}{M} \operatorname{Tr}[\boldsymbol{R}_{\boldsymbol{x}}]. \quad (4.44)$$

Using Eq. (10) in [70] yields

$$\frac{1}{4\delta^2} \iint_{-\delta}^{\delta} \left| \boldsymbol{a}(v)^H \boldsymbol{a}(u) \right|^2 du dv = \mathbb{E}\left[\left| \boldsymbol{a}(v)^H \boldsymbol{a}(u) \right|^2 \right] = \sum_{n=0}^{M-1} \sum_{m=0}^{M-1} \operatorname{sinc}^2 \left(2\pi \frac{d}{\lambda} (m-n) \delta \right), \quad (4.45)$$

which completes the proof. ∎

Remark 4.2. *Consider the case that* $M \gg 1$. *Then, from (4.43)*

$$\mathcal{P}_{q_1} \overset{(a)}{\approx} (\zeta - 1) \sigma_n^2 + (\zeta - 1) \left[\frac{1}{2\delta} \left(\frac{d}{\lambda} \right)^{-1} - \frac{1}{4\pi^2 \delta^2} \left(\frac{d}{\lambda} \right)^{-2} f\left(\frac{d}{\lambda} \right) \right] \sum_{k=1}^{K} p_k \beta_k + \left(\frac{\pi}{2} - \zeta \right) \sum_{k=1}^{K} p_k \beta_k, \quad (4.46)$$

where $f(x) \triangleq \frac{2}{M} \sum_{n=1}^{M-1} \frac{\sin^2(2\pi x \delta n)}{n}$ *and in (a) we have used Eq. (14) of [70]. Equation (4.46) states that, for standard one-bit quantization, increasing the spatial oversampling in a large antenna array* $(d/\lambda \to 0)$ *increases the quantization noise power proportional to* $(d/\lambda)^{-1}$.

Remark 4.3. *Consider the fixed-aperture case where* $d_0 = M \frac{d}{\lambda}$ *is a constant* (i.e., *the antenna spacing decreases proportionally to the increase in the number of antennas). Then, from*

(4.43)

$$\mathcal{P}_{q_1} \xrightarrow[M\to\infty]{} (\zeta - 1)\left[\sigma_n^2 + M\sum_{k=1}^{K} p_k\beta_k\right] + \left(\frac{\pi}{2} - \zeta\right)\sum_{k=1}^{K} p_k\beta_k. \tag{4.47}$$

Equation (4.47) states that, for standard one-bit quantization, increasing the number of antennas for an array with a fixed aperture, d_0, increases the quantization noise power linearly with M.

4.3.3.2 ΣΔ Quantization

From (4.16), the covariance of the quantization noise for the ΣΔ architecture is $\boldsymbol{R}_{\boldsymbol{q}_{\Sigma\Delta}} = \boldsymbol{U}^{-1}\boldsymbol{R}_{\boldsymbol{q}}\boldsymbol{U}^{-H}$. We derive an expression for the normalized quantization noise power of the ΣΔ array, $\mathcal{P}_{q_{\Sigma\Delta}}$, in the next proposition.

Proposition 4.2. *The quantization noise power for spatial ΣΔ quantization is*

$$\mathcal{P}_{q_{\Sigma\Delta}} = \frac{2}{M}\left(\mathrm{Tr}\left[\boldsymbol{R}_{\boldsymbol{q}}\right] - \sigma_{q_M}^2\right)\left[1 - \mathrm{sinc}\left(2\pi\frac{d}{\lambda}\delta\right)\right] + \frac{\sigma_{q_M}^2}{M}, \tag{4.48}$$

where $\sigma_{q_M}^2 = \mathbb{E}\left[|q_M|^2\right]$.

Proof. Substituting $\boldsymbol{R}_{\boldsymbol{q}_{\Sigma\Delta}} = \mathbb{E}\left[\boldsymbol{U}^{-1}\boldsymbol{q}\boldsymbol{q}^H\boldsymbol{U}^{-H}\right]$ into (4.34) leads to

$$\mathcal{P}_{q_{\Sigma\Delta}} = \frac{1}{M}\frac{1}{2\delta}\int_{-\delta}^{\delta}\mathbb{E}\left[\left|\boldsymbol{a}\left(u\right)^H\boldsymbol{U}^{-1}\boldsymbol{q}\right|^2\right]du. \tag{4.49}$$

We set $\phi = 0$ due to the assumption of $u \in [-\delta, \delta]$ in the definition of the quantization noise power, and we note that

$$\boldsymbol{U}^{-1} = \boldsymbol{I}_M - \boldsymbol{Z}_{-1}\,. \tag{4.50}$$

Then

$$\boldsymbol{U}^{-1}\boldsymbol{q} = (\boldsymbol{I}_M - \boldsymbol{Z}_{-1})\,\boldsymbol{q} = \begin{bmatrix} q_1 \\ q_2 - q_1 \\ \vdots \\ q_M - q_{M-1} \end{bmatrix}. \tag{4.51}$$

In addition, from (4.23), and the fact that $\boldsymbol{R}_{\boldsymbol{qx}} \approx \boldsymbol{0}$, it can be readily shown that $\mathbb{E}\left[q_m q^*_{m\pm1}\right] \approx 0$. Hence, for the sake of analysis, we approximate $\mathbb{E}\left[q_m q^*_{m'}\right] \approx 0,\ \forall m \neq m' \in \mathcal{M}$, and therefore $\boldsymbol{R}_{\boldsymbol{q}} = \mathrm{diag}\left(\boldsymbol{p}_{\boldsymbol{q}}\right)$. Consequently,

$$\mathbb{E}\left[\left|\boldsymbol{a}\,(u)^H\,\boldsymbol{U}^{-1}\boldsymbol{q}\right|^2\right] = \left|1 - e^{j2\pi\frac{d}{\lambda}u}\right|^2 \sum_{m=1}^{M-1}\mathbb{E}\left[|q_m|^2\right] + \mathbb{E}\left[|q_M|^2\right] = 4\left(\mathrm{Tr}\left[\boldsymbol{R}_{\boldsymbol{q}}\right] - \sigma^2_{q_M}\right)\sin^2\left(\pi\frac{d}{\lambda}u\right) + \sigma^2_{q_M}. \tag{4.52}$$

By integrating (4.52) and using some algebraic manipulation, we arrive at (4.48). ■

Remark 4.4. *Consider the case that $M \gg 1$. Then, from (4.48)*

$$\mathcal{P}_{q_{\Sigma\Delta}} \overset{(a)}{\approx} \frac{4}{3}\frac{\frac{\pi}{2}-1}{2-\frac{\pi}{2}}\pi^2\delta^2\left(\frac{d}{\lambda}\right)^2 p_x\,, \tag{4.53}$$

where in (a) we have used $\mathrm{sinc}\,(x) \approx 1 - \frac{x^2}{6}$ *and*

$$\frac{1}{M}\left(\mathrm{Tr}\left[\boldsymbol{R}_{\boldsymbol{q}}\right] - \sigma^2_{q_M}\right) \approx \frac{\frac{\pi}{2}-1}{2-\frac{\pi}{2}}p_x \tag{4.54}$$

for $M \gg 1$ and assuming $\boldsymbol{p}_{\boldsymbol{x}} = p_x\boldsymbol{1}$. Equation (4.53) states that, by increasing the spatial oversampling ($d/\lambda \to 0$), the quantization noise power for the $\Sigma\Delta$ array tends to zero proportional to $(d/\lambda)^2$. This result is in contrast to that for the standard one-bit quantization power, which was shown earlier to increase proportional to $(d/\lambda)^{-1}$. Hence, the spatial $\Sigma\Delta$ architecture brings about an oversampling gain of $(d/\lambda)^3$ compared to the standard one-bit

architecture. While this is a promising result, as mentioned earlier the practical limitations of placing antenna elements close to each other prevent us from achieving a high degree of spatial oversampling.

Remark 4.5. *Consider the case that* $d_0 = M\frac{d}{\lambda}$ *is a constant. Then, from (4.48)*

$$M^2\mathcal{P}_{q_{\Sigma\Delta}} \xrightarrow[M\to\infty]{} \frac{4}{3}\frac{\frac{\pi}{2}-1}{2-\frac{\pi}{2}}\pi^2\delta^2 d_0^2 p_x. \tag{4.55}$$

Equation (4.55) states that, for spatial $\Sigma\Delta$ *quantization, increasing the number of antennas for an array with a fixed aperture,* d_0*, decreases the quantization noise power proportional to* $1/M^2$*. Hence, the spatial* $\Sigma\Delta$ *architecture brings about an oversampling gain of* M^3 *compared to the standard one-bit architecture.*

In the next section, we study the spectral efficiency of a massive MIMO system with spatial $\Sigma\Delta$ processing and discuss the impact of the spatial $\Sigma\Delta$ architecture on the system performance.

4.4 Spectral Efficiency

In this section, we study the SE of a massive MIMO system with spatial $\Sigma\Delta$ processing. We consider maximum ratio combining (MRC) and zero-forcing (ZF) receivers. We derive here an approximation for the SE of the system with an MRC receiver, and evaluate the SE for the ZF receiver in the next section, numerically. We first present the case where perfect channel state information (CSI) is assumed to be available at the BS, and then we discuss the impact of imperfect CSI on the system performance at the end of the section.

From (4.1) and (4.16), the received signal at a BS with a $\Sigma\Delta$ architecture can be modeled as

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{s} + \boldsymbol{n} + \boldsymbol{U}^{-1}\boldsymbol{q}. \tag{4.56}$$

Denoting the linear receiver by $\boldsymbol{W}$, we have

$$\hat{\boldsymbol{s}} = \boldsymbol{W}^{H}\boldsymbol{G}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{s} + \boldsymbol{W}^{H}\boldsymbol{n} + \boldsymbol{W}^{\boldsymbol{H}}\boldsymbol{U}^{-1}\boldsymbol{q}\,, \tag{4.57}$$

and the kth element of $\hat{\boldsymbol{s}}$ is given by

$$\hat{s}_k = \sqrt{p_k}\boldsymbol{w}_k^{H}\boldsymbol{g}_k s_k + \sum_{i=1, i\neq k}^{K} \sqrt{p_k}\boldsymbol{w}_k^{H}\boldsymbol{g}_i s_i + \boldsymbol{w}_k^{H}\boldsymbol{n} + \boldsymbol{w}_k^{H}\boldsymbol{U}^{-1}\boldsymbol{q}, \tag{4.58}$$

where $\boldsymbol{w}_k$ is the kth column of $\boldsymbol{W}$. We assume the BS treats $\boldsymbol{w}_k^{H}\boldsymbol{g}_k$ as the desired channel and the other terms of (4.58) as worst-case Gaussian noise when decoding the signal. Consequently, a lower bound for the ergodic achievable SE at the kth user can be written as [71]

$$\mathcal{S}_k = \mathbb{E}\left[\log_2\left(1 + \frac{p_k\left|\boldsymbol{w}_k^{H}\boldsymbol{g}_k\right|^2}{\Omega}\right)\right], \tag{4.59}$$

where

$$\Omega = \sum_{i=1, i\neq k}^{K} p_k\left|\boldsymbol{w}_k^{H}\boldsymbol{g}_i\right|^2 + \sigma_n^2\|\boldsymbol{w}_k\|^2 + \boldsymbol{w}_k^{H}\boldsymbol{U}^{-1}\boldsymbol{R}_{\boldsymbol{q}}\boldsymbol{U}^{-H}\boldsymbol{w}_k. \tag{4.60}$$

4.4.1 MRC Receiver

For the case of an MRC receiver, $\boldsymbol{W} = \boldsymbol{G}$. The following proposition presents an approximation for the achievable SE of a massive MIMO system with spatial $\Sigma\Delta$ processing and an MRC receiver.

Proposition 4.3. *For a massive MIMO system employing a spatial* $\Sigma\Delta$ *architecture and an MRC receiver, the SE of the* k*th user assuming perfect CSI is given by*

$$\mathcal{S}_k \approx \log_2\left(1 + \frac{\mathcal{N}}{\mathcal{D}}\right) \tag{4.61}$$

where

$$\mathcal{N} = p_k\beta_k\left(\left|\mathrm{Tr}\left[\boldsymbol{\Sigma}_{kk}\right]\right|^2 + \mathrm{Tr}\left[\boldsymbol{\Sigma}_{kk}^2\right]\right),$$

$$\mathcal{D} = \sum_{i=1,i\neq k}^{K} p_i\beta_i\mathrm{Tr}\left[\boldsymbol{\Sigma}_{ik}\boldsymbol{\Sigma}_{ik}^H\right] + \sigma_n^2\mathrm{Tr}\left[\boldsymbol{\Sigma}_{kk}\right] + \frac{4}{L}\left(\mathrm{Tr}\left[\boldsymbol{R}_{\boldsymbol{q}}\right] - \sigma_{q_M}^2\right)\sum_{\ell=1}^{L}\sin^2\left(\frac{\phi - 2\pi\frac{d}{\lambda}\sin\left(\theta_{\mathrm{k}\ell}\right)}{2}\right) + \sigma_{q_M}^2,$$

and $\boldsymbol{\Sigma}_{ik} \triangleq \frac{1}{L}\boldsymbol{A}_i^H\boldsymbol{A}_k$.

Proof. From [71], an approximation for (4.59) can be calculated as

$$\mathcal{S}_k \approx \log_2\left(1 + \frac{p_k\mathbb{E}\left[\left|\boldsymbol{w}_k^H\boldsymbol{g}_k\right|^2\right]}{\mathbb{E}\left[\boldsymbol{\Omega}\right]}\right). \tag{4.62}$$

By setting $\boldsymbol{w}_k = \boldsymbol{g}_k$ and using Lemma 2 of [11] and Lemma 1 of [72], the expected values of the desired signal, interference, and thermal noise can be readily calculated. For the quantization noise term, note that

$$\boldsymbol{U}^{-1} = \boldsymbol{I}_M - e^{-j\phi}\boldsymbol{Z}_{-1}. \tag{4.63}$$

Therefore,

$$\boldsymbol{U}^{-1}\boldsymbol{q} = \left(\boldsymbol{I}_M - e^{-j\phi}\boldsymbol{Z}_{-1}\right)\boldsymbol{q} = \begin{bmatrix} q_1 \\ q_2 - e^{-j\phi}q_1 \\ \vdots \\ q_M - e^{-j\phi}q_{M-1} \end{bmatrix}. \tag{4.64}$$

In addition, the kth user channel vector can be written as

$$\boldsymbol{g}_k = \sqrt{\frac{\beta_k}{L}} \sum_{l=1}^{L} h_{kl} \boldsymbol{a}\left(\theta_{kl}\right), \tag{4.65}$$

where h_{kl} is the lth element of $\boldsymbol{h}_k$. Hence,

$$\mathbb{E}\left[\left|\boldsymbol{g}_k^H \boldsymbol{U}^{-1} \boldsymbol{q}\right|^2\right] = \frac{\beta_k}{L} \mathbb{E}\left[\left|\sum_{\ell=1}^{L} h_{kl}\left(1 - e^{-j\phi} z_{kl}\right) \sum_{m=1}^{M-1} q_m z_{kl}^{m-1} + q_M z_{kl}^{M-1}\right|^2\right], \tag{4.66}$$

which, after some algebraic manipulation, leads to (4.61) and the proof is complete. ■

Remark 4.6. *The noise shaping characteristic of the spatial* $\Sigma\Delta$ *architecture is explicitly manifested in (4.61). A similar characteristic is observed in [48] for* $\Sigma\Delta$ *precoding. It shows the importance of the design parameter* ϕ *which should be chosen to minimize* $\mathcal{G} = \frac{1}{L}\sum_{\ell=1}^{L} \sin^2\left(\frac{\phi - 2\pi\frac{d}{\lambda}\sin(\theta_{k\ell})}{2}\right)$ *for all users. By writing the steering angle as* $\phi = 2\pi\frac{d}{\lambda}\sin\left(\theta\right)$*, we have*

$$\mathcal{G} = \frac{1}{L} \sum_{\ell=1}^{L} \sin^2\left(\pi \frac{d}{\lambda}\Big(\sin\left(\theta\right) - \sin\left(\theta_{k\ell}\right)\Big)\right). \tag{4.67}$$

Eq. (4.67) indicates that $\mathcal{G}$ *could be made arbitrarily small by decreasing the relative antenna spacing* d/λ *(the spatial oversampling gain) or* $\sin\left(\theta\right) - \sin\left(\theta_{k\ell}\right)$ *(the angle steering gain). However, physical constraints on the antenna spacing and larger angular spreads,* Θ*, limit the lower bound on* $\mathcal{G}$*. For the case that* $L \gg 1$*,* $\sin\left(\theta_{k\ell}\right) = u_{k\ell}$ *can be taken as a random variable uniformly distributed in* $[\delta_1, \delta_2]$ *where* $\delta_1 = \sin\left(\theta_0 - \frac{\Theta}{2}\right)$ *and* $\delta_2 = \sin\left(\theta_0 + \frac{\Theta}{2}\right)$*. Hence,*

$$\mathcal{G} \approx \frac{1}{\delta_2 - \delta_1} \int_{\delta_1}^{\delta_2} \sin^2\left(\frac{\phi - 2\pi\frac{d}{\lambda}u}{2}\right) du = \frac{1}{2} + \frac{1}{4\pi}\left(\frac{d}{\lambda}\right)^{-1} \frac{1}{\delta_2 - \delta_1}\left(b_0 \sin\left(\phi\right) - b_1 \cos\left(\phi\right)\right), \tag{4.68}$$

where

$$b_0 = \cos\left(2\pi\frac{d}{\lambda}\delta_2\right) - \cos\left(2\pi\frac{d}{\lambda}\delta_1\right)$$

$$b_1 = \sin\left(2\pi\frac{d}{\lambda}\delta_2\right) - \sin\left(2\pi\frac{d}{\lambda}\delta_1\right).$$

In this case, the optimal value of the steering angle that minimizes $\mathcal{G}$ can be simply derived as

$$\phi^\star = \begin{cases} 0 & \delta_2 = -\delta_1 \\ -\tan^{-1}\left(\frac{b_0}{b_1}\right) & otherwise \end{cases} \tag{4.69}$$

which indicates that the optimal steering angle is dependent on δ_1, δ_2, and the relative antenna spacing d/λ.

4.4.2 ZF receiver

For the ZF receiver, $\boldsymbol{W} = \boldsymbol{G}\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}$. After substituting this for $\boldsymbol{W}$ in (4.62), the SE achieved for the kth user with the $\Sigma\Delta$ architecture and ZF receiver can be written as

$$\mathcal{S}_k = \mathbb{E}\left[\log_2\left(1 + \frac{p_k}{\|\boldsymbol{w}_k\|^2\sigma_n^2 + \left[\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}\boldsymbol{G}^H\boldsymbol{U}^{-1}\boldsymbol{R_q}\boldsymbol{U}^{-H}\boldsymbol{G}\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}\right]_{kk}}\right)\right]. \tag{4.70}$$

Although (4.70) does not provide direct insight into the effect of the shaped quantization noise on the SE, in Section 4.6 we numerically evaluate this expression and show the superior performance of the $\Sigma\Delta$ architecture compared with standard one-bit quantization.

4.5 FBB $\Sigma\Delta$ Architecture

One drawback of using one-bit ADCs is its susceptibility to strong interference, which can occur in MIMO systems when a jammer is present. Since a one-bit ADC has zero dynamic

range, a moderately strong jammer can effectively swamp the relatively weak signals of interest and significantly degrade performance. With one-bit ADCs, jammer mitigation must occur in the RF domain prior to sampling, in order for the coarsely quantized ADCs to reveal the dynamics of the signals of interest. Simulations show that while the ordinary spatial $\Sigma\Delta$ architecture is not effective in adequately alleviating the impact of the jammer, our proposed spatial FBB $\Sigma\Delta$ quantizer can provide performance that is close to that of an unquantized system.

In this section, we show that the feedback employed by the spatial $\Sigma\Delta$ architecture can be generalized for this purpose. The genesis of the idea comes from work described in [42], which is based on the use of temporally oversampled $\Sigma\Delta$ ADCs. Instead of simply using feedback of the delayed ADC output as in a standard temporal $\Sigma\Delta$ architecture, the approach of [42] employs an analog feedback beamformer (FBB) designed to temporally null the interference. Unlike [42], in the method presented here we take a different approach that does not employ temporal oversampling, but instead uses the *spatial* $\Sigma\Delta$ architecture. In particular, the feedback between adjacent antennas is generalized to include an FBB signal that also serves to spatially null the interference. We generalize the approach discussed in last sections to develop an equivalent linear model for the $\Sigma\Delta$ array that includes the FBB signal and sets the output level of the one-bit quantizers.

To analyze the FBB $\Sigma\Delta$ architecture, consider the uplink of a massive MIMO system consisting of a legitimate, single-antenna user that send its signal to a base station (BS) equipped with a uniform linear array of $\boldsymbol{M}$ antennas. In addition, a jammer is present that aims to impair the performance of the legitimate user. Accordingly, the $\boldsymbol{M} \times 1$ received signal at the BS is

$$\boldsymbol{x} = \sqrt{p}\boldsymbol{g}_U s_U + \boldsymbol{n} + \sqrt{q}\boldsymbol{g}_J s_J, \tag{4.71}$$

where p represents the average transmission power from the user, $\boldsymbol{g}_U = 1/\sqrt{L}\sum_{\ell=1}^{L} h_\ell \boldsymbol{a}(\theta_\ell)$ is the user's channel vector where L denotes the number of signal paths, $h_\ell \sim \mathcal{CN}(0,1)$ is the complex channel gain for the ℓ-th path, and

$$\boldsymbol{a}(\theta_\ell) = [1, e^{-j2\pi\frac{d}{\lambda}\sin(\theta_\ell)}, \cdots, e^{-j(M-1)2\pi\frac{d}{\lambda}\sin(\theta_\ell)}]^T \tag{4.72}$$

denotes the array response vector for angle of arrival θ_ℓ, where d and λ represent the antenna spacing and the carrier wavelength, respectively. The symbol $s_U \in \mathbb{C}$ transmitted from the user satisfies $\mathbb{E}\left[|s_U|^2\right] = 1$, and $\boldsymbol{n} \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I}\right)$ denotes additive CSCG receiver noise at the BS. In addition, q represents the jammer's average power, $\boldsymbol{g}_J = 1/\sqrt{L}\sum_{j=1}^{L} h_j \boldsymbol{a}\left(\theta_j\right)$ is the channel vector between the jammer and the BS, and s_J denotes the jammer's signal, where $\mathbb{E}\left[|s_J|^2\right] = 1$.

Since each antenna element at the BS is connected to a one-bit ADC, the received baseband signal at the BS becomes (4.6).

4.5.1 Temporal FBB ΣΔ Modulation

In this subsection, we elaborate on the temporal FBB ΣΔ modulation approach of [42] to clarify the noise shaping characteristics of this technique.

Fig. 4.4 shows a block diagram representing the Nth-order temporal FBB ΣΔ modulator with feedback weights $\boldsymbol{w} = [w_0, \cdots, w_{N-1}]^T$. To shape the quantization noise, the weighted output signals are fed back and subtracted from the input (Δ-stage), and then this error is integrated (Σ-stage).

To characterize the transfer function of this non-linear system, we substitute the one-bit quantizer with the equivalent linear model depicted in Fig. 4.5. The input-output relation-

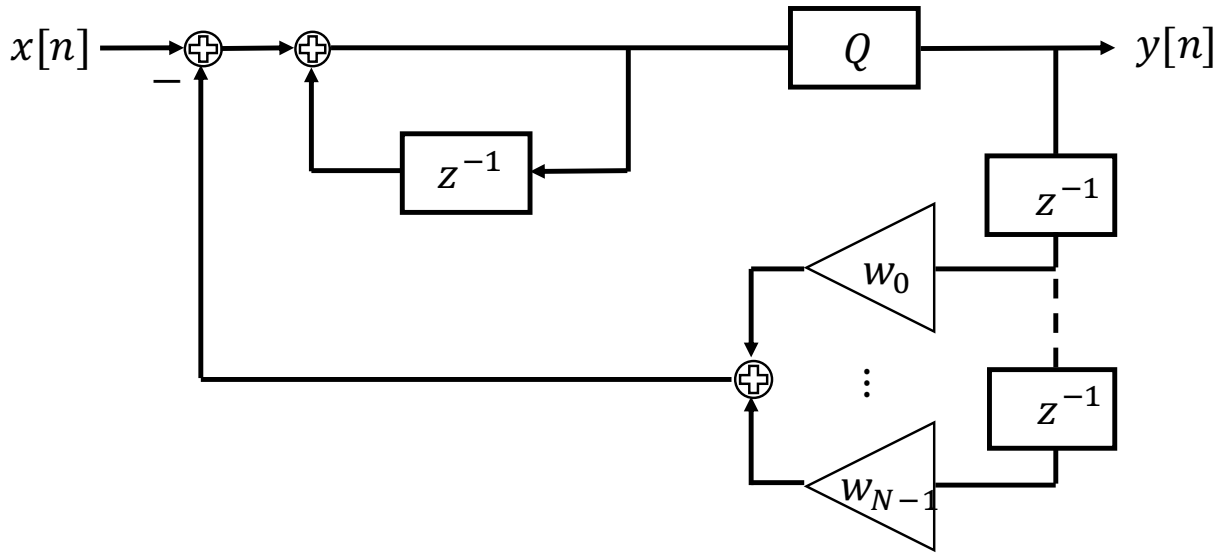

Figure 4.4: Block diagram for the Nth-order temporal FBB $\Sigma\Delta$ modulator.

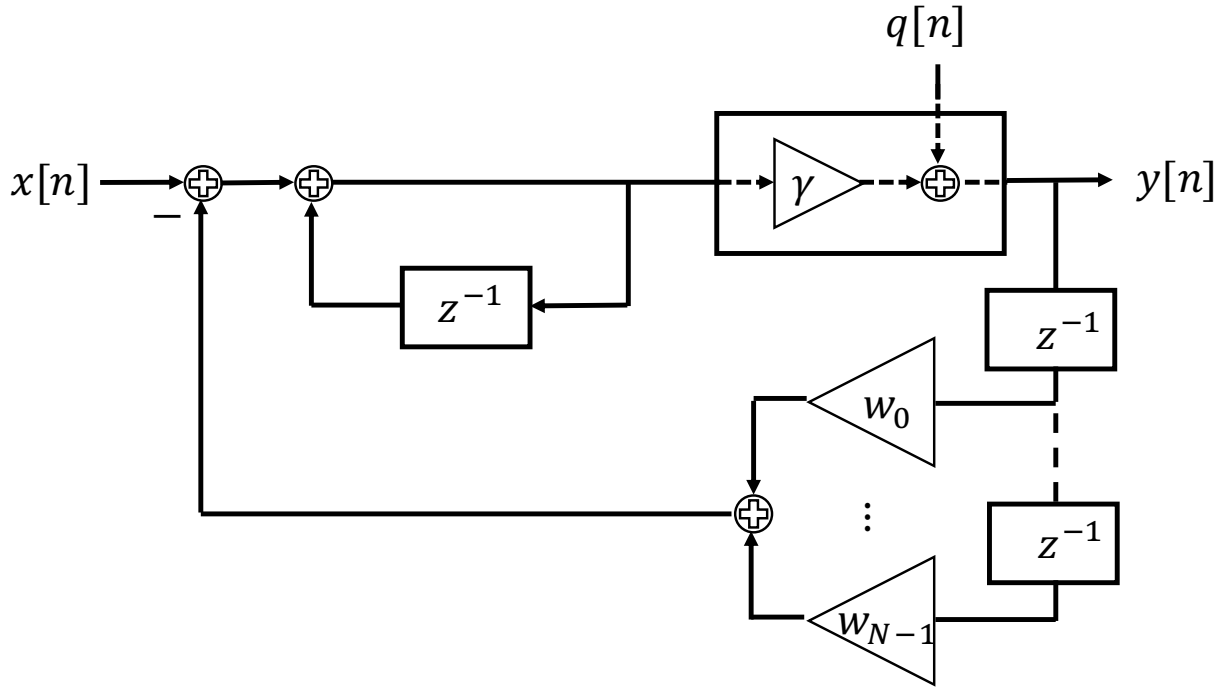

Figure 4.5: Equivalent linear model for Nth-order temporal FBB $\Sigma\Delta$.

ship of the FBB ΣΔ quantizer can then be written as

$$Y(z) = A(z)X(z) + B(z)Q(z), \tag{4.73}$$

where $X(z) = \sum_{n=0}^{\infty} x[n] z^{-n}$ denotes the z-transform and

$$A(z) = \frac{\gamma}{1 + (\gamma w_0 - 1) z^{-1} + \gamma w_1 z^{-2} + \cdots + \gamma w_{N-1} z^{-N}} \tag{4.74}$$

$$B(z) = \frac{1 - z^{-1}}{1 + (\gamma w_0 - 1) z^{-1} + \gamma w_1 z^{-2} + \cdots + \gamma w_{N-1} z^{-N}}. \tag{4.75}$$

Unlike ordinary ΣΔ modulation that passes the signal through an all-pass filter and the quantization noise through a high-pass filter, in (4.73) we see that $X(z)$ is passed through $A(z)$ and $Q(z)$ through $B(z)$ for FBB ΣΔ modulation. Hence, this approach not only provides a tool for shaping the quantization noise, but proper design of the feedback weights $\boldsymbol{w}$ allows for temporal filtering that passes the desired signal while eliminating undesirable contributions from other sources such as a jammer.

4.5.2 One-Bit Spatial FBB ΣΔ Modulation

The idea underlying temporal FBB ΣΔ modulation can be adapted to the angle domain, in order to *spatially* shape the quantization noise in a desired way and remove interference. It is worthwhile to note that although ordinary ΣΔ modulation provides a noise shaping characteristic, FBB ΣΔ not only shapes the quantization noise, but also suppresses the extra quantization noise caused by the jammer with appropriate feedback beamforming.

Fig. 4.6 shows the architecture of the angle-steered FBB ΣΔ array. Using Fig. 4.6, we can

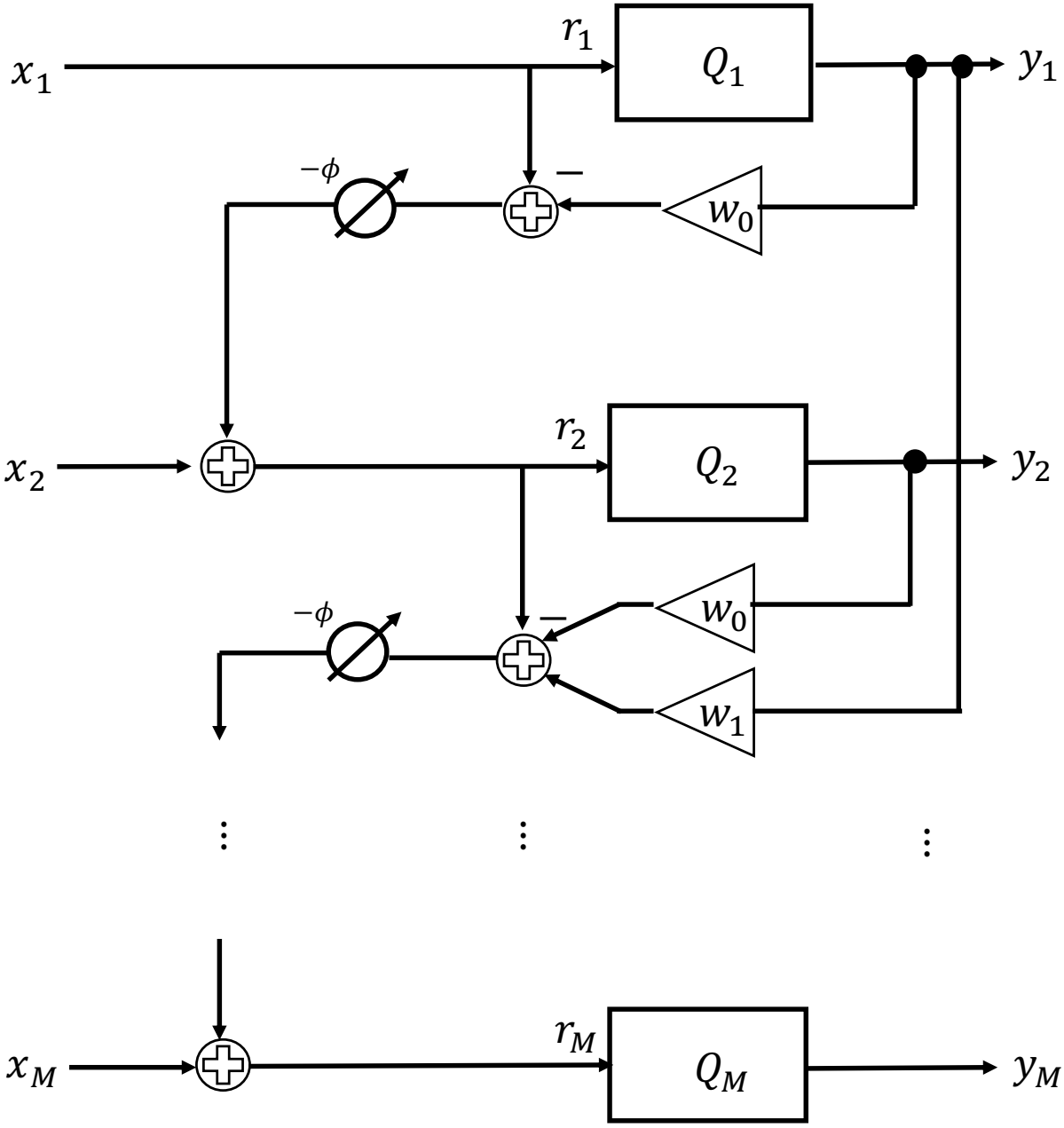

Figure 4.6: Spatial FBB ΣΔ architecture.

formulate a compact input-output description of the spatial FBB ΣΔ array by defining

$$\boldsymbol{V} = e^{-j\phi}\boldsymbol{Z}_{-1}\boldsymbol{U}\boldsymbol{W}, \tag{4.76}$$

where

$$\boldsymbol{W} = \begin{bmatrix} w_0 & & & & & \mathbf{0} \\ w_1 & w_0 & & & & \\ \vdots & \ddots & w_0 & & & \\ w_{N-1} & \ddots & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ \mathbf{0} & \cdots & w_{N-1} & \cdots & w_1 & w_0 \end{bmatrix}, \tag{4.77}$$

and expressing the input to the quantizers as

$$\boldsymbol{r} = \boldsymbol{U}\boldsymbol{x} - \boldsymbol{V}\boldsymbol{y}. \tag{4.78}$$

The output of the angle-steered one-bit FBB ΣΔ array is then defined by

$$\boldsymbol{y} = \mathcal{Q}(\boldsymbol{r}). \tag{4.79}$$

By using the linear model introduced in section 4.3.1, we obtain the following mathematical model for the FBB ΣΔ architecture:

$$\boldsymbol{y} = \boldsymbol{B}^{-1}\boldsymbol{U}\boldsymbol{x} + \boldsymbol{B}^{-1}\boldsymbol{q}, \tag{4.80}$$

where $\boldsymbol{B} = \boldsymbol{I} + \boldsymbol{V}$. Equation (4.80) is the spatial equivalent of the temporal FBB ΣΔ descrip-

tion in (4.73), with the following equivalences:

$$\boldsymbol{B}^{-1}\boldsymbol{U} \longleftrightarrow A(z) \qquad\qquad \boldsymbol{B}^{-1} \longleftrightarrow B(z). \tag{4.81}$$

Recall from section 4.3.1 that the condition $\boldsymbol{\Gamma} = \boldsymbol{I}$ and assuming that r_m is approximately Gaussian leads to the following choice for the output levels:

$$\alpha_m^{\star} = \frac{\sqrt{\pi \mathbb{E}\left[|r_m|^2\right]}}{2}. \tag{4.82}$$

Next, we calculate the power of the equivalent quantization noise $\boldsymbol{q}$, which is needed both to analytically assess the system performance and to compute $\alpha_m^{\star}$. In the discussion below, we show how to express (4.82) in terms of the statistics of the array input $\boldsymbol{x}$, which illustrates how the quantizer output levels can be analytically chosen in a practical setting. Moreover, we will show how spatial FBB $\Sigma\Delta$ impacts the power of the quantization noise and elaborate on how it differs from the ordinary spatial $\Sigma\Delta$ approach.

With $\boldsymbol{\Gamma} = \boldsymbol{I}$, we have

$$\boldsymbol{y} = \boldsymbol{r} + \boldsymbol{q}. \tag{4.83}$$

Since r_m and q_m are uncorrelated, and using (4.82), we obtain

$$\mathbb{E}\left[|q_m|^2\right] = \mathbb{E}\left[|y_m|^2\right] - \mathbb{E}\left[|r_m|^2\right] = \left(\frac{\pi}{2} - 1\right)\mathbb{E}\left[|r_m|^2\right]. \tag{4.84}$$

To determine $\mathbb{E}\left[|r_m|^2\right]$, we substitute (4.83) into (4.78), so that

$$\boldsymbol{r} = \boldsymbol{B}^{-1}\boldsymbol{U}\boldsymbol{x} - \boldsymbol{B}^{-1}\boldsymbol{V}\boldsymbol{q}. \tag{4.85}$$

Let us denote $\boldsymbol{\Psi} = \boldsymbol{B}^{-1}\boldsymbol{U}$ and $\boldsymbol{\Lambda} = \boldsymbol{B}^{-1}\boldsymbol{V}$. It is clear that $\boldsymbol{\Psi}$ is a lower triangular matrix and

$\boldsymbol{\Lambda}$ is a lower triangular matrix with zeros along the main diagonal. In addition, following the same reasoning as in Appendix A of [15], it can be shown that $\mathbb{E}\left[x_{m'}q_m^*\right] \approx 0,\ \forall m, m' \in \mathcal{M} = \{1, \cdots, M\}$. This results in $\boldsymbol{R_{qx}} \approx \mathbf{0}$. Therefore,

$$\boldsymbol{R_r} \approx \boldsymbol{\Psi R_x \Psi}^H + \boldsymbol{\Lambda R_q \Lambda}^H \ . \tag{4.86}$$

Eq. (4.86) implies that

$$\mathbb{E}\left[|r_m|^2\right] \approx \begin{array}{ll} \left[\boldsymbol{\Psi R_x \Psi}^H\right]_{mm} & m = 1 \\ \left[\boldsymbol{\Psi R_x \Psi}^H\right]_{mm} + \left[\boldsymbol{\Lambda R_q \Lambda}^H\right]_{mm} & m > 1 \end{array} \tag{4.87}$$

To approximate $\mathbb{E}\left[|r_m|^2\right]$, and for the sake of analysis, we assume $\boldsymbol{R_q}$ is diagonal. Since $\boldsymbol{\Lambda}$ is a lower triangular matrix with zeros along the main diagonal, we only need the first $m-1$ diagonal elements of $\boldsymbol{R_q}$ to specify $\mathbb{E}\left[|r_m|^2\right]$. Hence, we can recursively calculate $\mathbb{E}\left[|r_m|^2\right]$ for $m > 1$ using (4.84) and (4.87). In the next section, we show that the diagonal elements of $\boldsymbol{R_q}$ are much smaller than those for the ordinary $\Sigma\Delta$ architecture. This is because of the appropriate design of the feedback weights that lead to the elimination of strong interference before the one-bit quantization.

4.6 Numerical Results

4.6.1 $\Sigma\Delta$ architecture

In this section, we numerically evaluate the SE performance of $\Sigma\Delta$ massive MIMO systems in various scenarios. We assume static-aware power control in the network [51] so that $p_k = p_0/\beta_k$. In all of the cases considered, unless otherwise noted, we assume $M = 100$ antennas, $K = 10$ users, and an angular spread of $\Theta = 40°$ centered at $\theta_0 = 30°$. We assume

the same DoAs for all users, i.e., $\boldsymbol{A}_k = \boldsymbol{A},\ \forall k$, drawn uniformly from the interval $[10°, 50°]$, which corresponds to $u = \sin(\theta) \in [0.17, 0.77]$, and the steering angle of the $\Sigma\Delta$ array is set to $\phi = 2\pi\frac{d}{\lambda}\sin(\theta_0)$. The SNR is defined to be $\text{SNR} \triangleq \frac{p_0}{\sigma_n^2}$. We further assume CSCG symbols and 10^4 Monte Carlo trials for the simulations.

Fig. 4.7 shows the simulated and analytically derived quantization noise power density, i.e., $\rho_q(u),\ u \in [-1, 1]$, for $\Sigma\Delta$ and standard one-bit quantization when the relative antenna spacing is $d = \lambda/4$. We see that the quantization noise power for the $\Sigma\Delta$ array is substantially lower over the angles where the users are present, while the effect is the opposite for standard one-bit quantization – the quantization noise is higher for angles where the amplitude of the received signals is larger. We also observe that there is excellent agreement between the simulations and our theoretically derived expressions for both cases. Note that the careful design of the quantizer output levels is a critical aspect for achieving the desired $\Sigma\Delta$ noise shaping characteristic shown here.

The impact of spatial oversampling on the shape of the quantization noise spectrum is illustrated in Fig. 4.8. We see from the figure that, as discussed in Remarks 4.2, the quantization noise power for the standard one-bit ADC architecture grows as d/λ decreases. Analogously to temporal $\Sigma\Delta$ modulation where increasing the sampling rate helps to push the quantization noise to higher frequencies and widen the quantization-noise-free band, we can reduce the quantization noise power over wider angular regions by placing the antenna elements of the array closer together. For example, when $d = \lambda/2$, the $\Sigma\Delta$ quantization noise power is below that of the standard one-bit quantizer over a beamwidth of 40°. This beamwidth increases to about 80°, 150°, and 180° for $d = \lambda/4$, $d = \lambda/8$, and $d = \lambda/16$, respectively. Mutual coupling will impact these results as d decreases, but both the standard one-bit and $\Sigma\Delta$ approaches would be expected to degrade.

In Fig. 4.9, we compare the SE performance of $\Sigma\Delta$ and standard one-bit quantization for the case of an MRC receiver. It is clear that the derived theoretical SE expression in

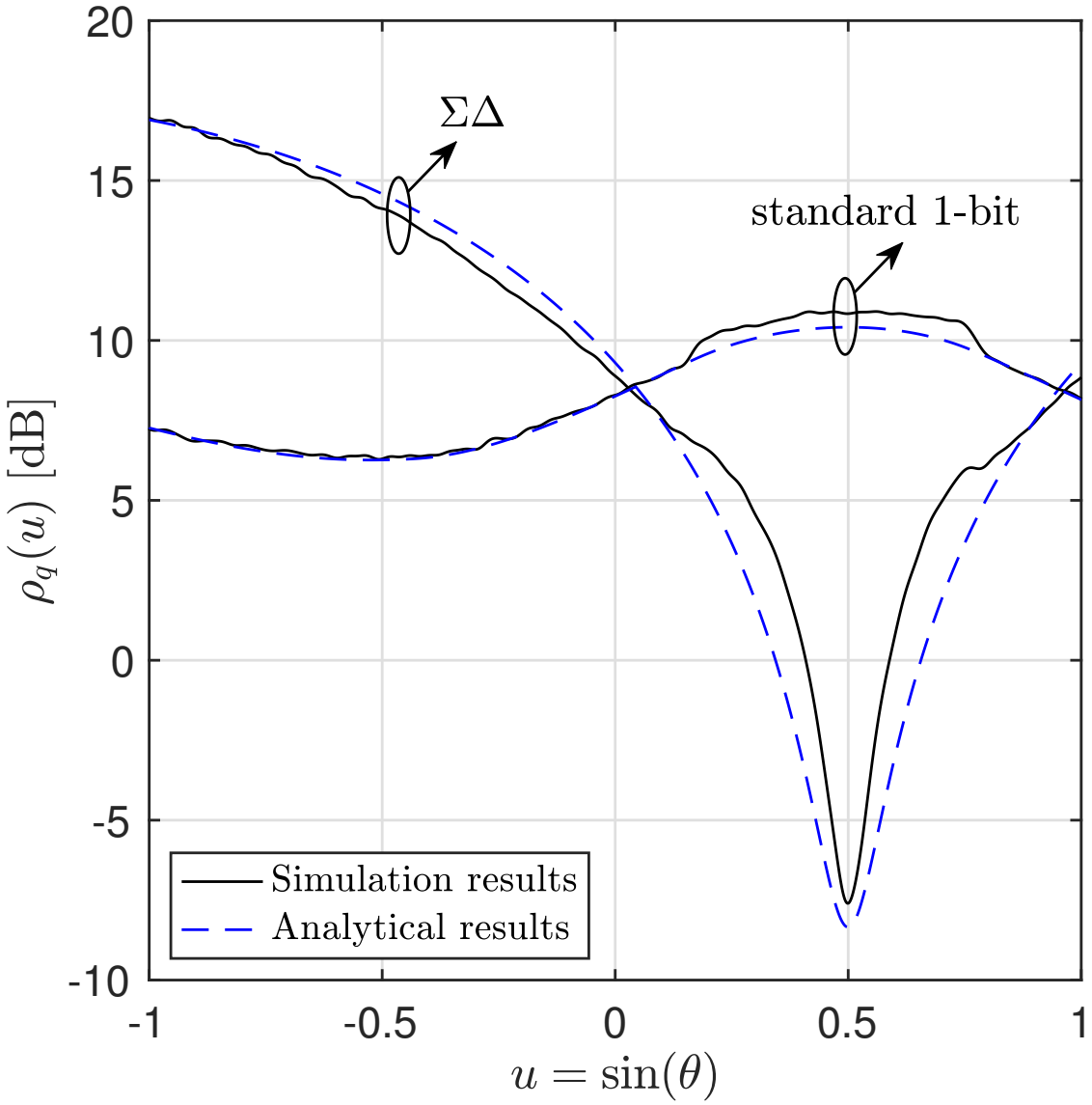

Figure 4.7: Spatial spectrum of the quantization noise for the ΣΔ and standard one-bit architectures when $L = 50$, $d = \lambda/4$, and SNR = 0 dB.

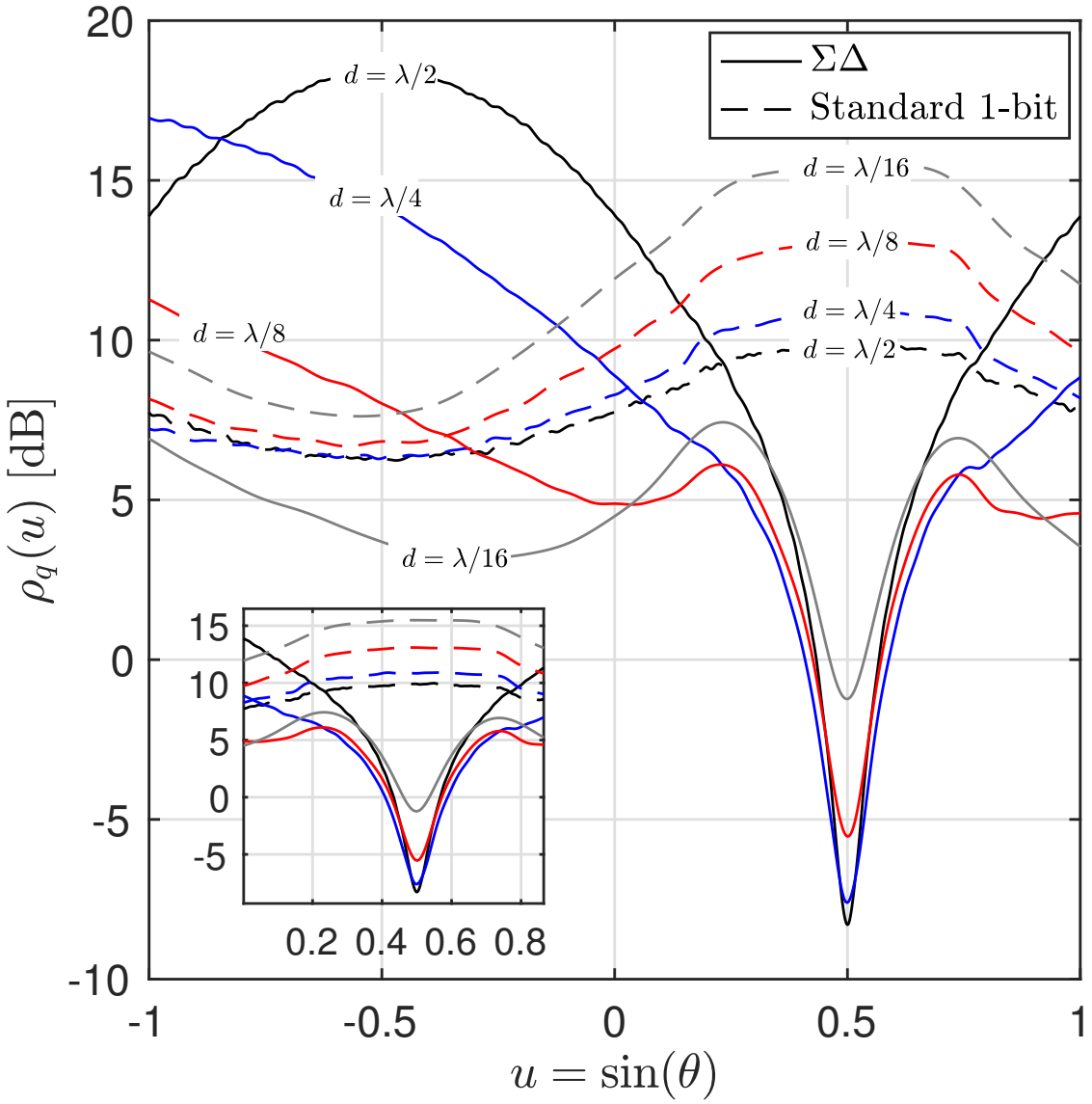

Figure 4.8: Spatial spectrum of the quantization noise for the ΣΔ and standard one-bit architectures for different antenna spacings when $L = 50$ and SNR = 0 dB.

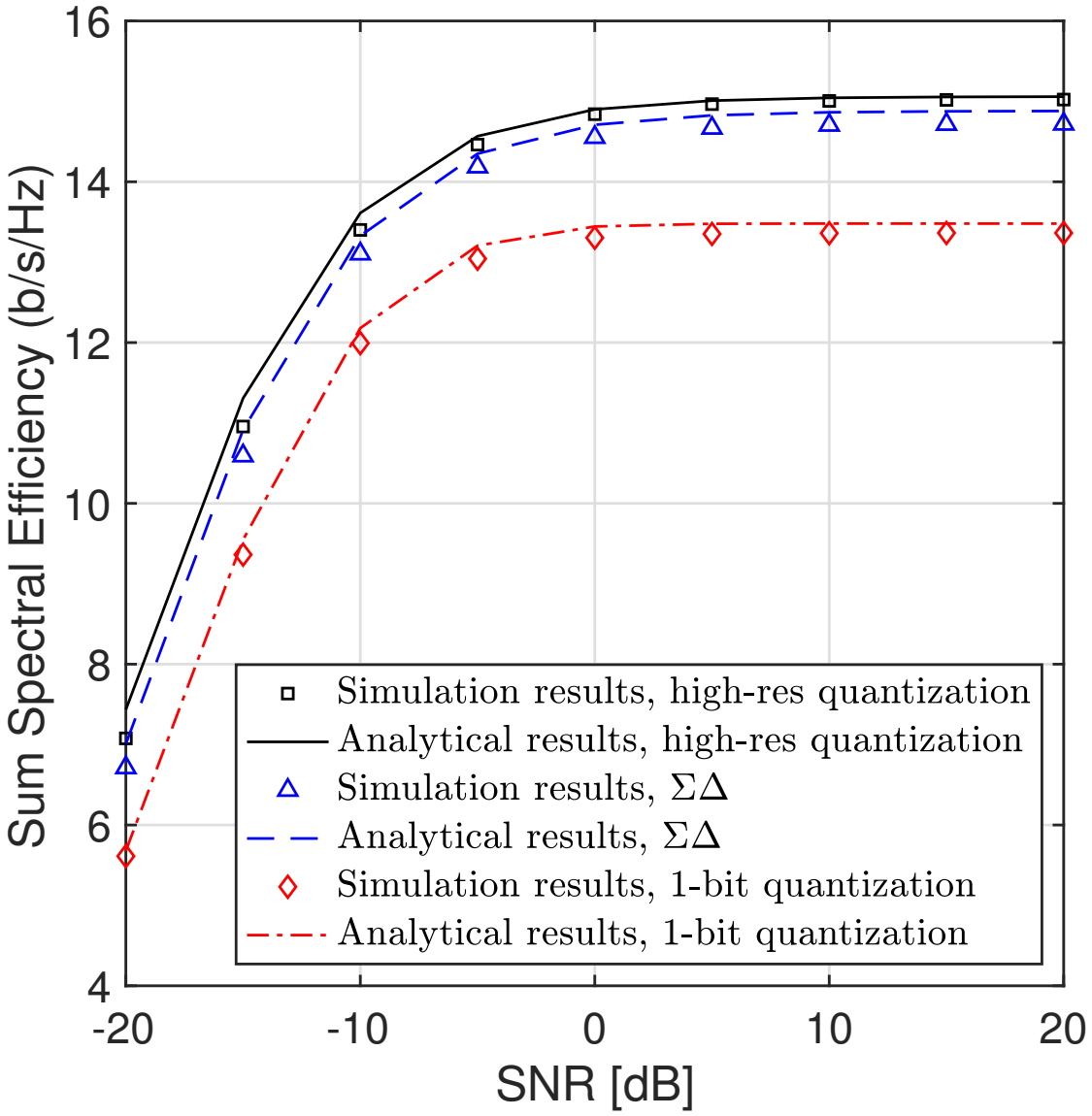

Figure 4.9: SE versus SNR for MRC receiver with perfect CSI, $L = 50$, and $d = \lambda/4$.

(4.61) very closely matches the simulated value of the expression in (4.59). The one-bit $\Sigma\Delta$ implementation achieves a significantly increased SE compared with standard one-bit quantization, and performs nearly identically to an MRC receiver with infinite resolution ADCs. It should be emphasized that this performance gain is achieved without paying a significant penalty in terms of power consumption (as with mixed-ADC architectures) or complicated processing (as required by non-linear receivers).

In Fig. 4.10 and 4.11, we numerically evaluate the SE when the ZF receiver is employed, using Eq. (4.70). The SE improvement of $\Sigma\Delta$ processing is much greater than for the case of MRC. For example, at SNR = 0 dB, about a 50% improvement in SE can be achieved by the spatial $\Sigma\Delta$ architecture compared with standard one-bit quantization, which confirms its ability to provide high SE with a simple architecture and low power consumption.

The effect of channel estimation error on the performance of the algorithms is also shown in Fig. 4.10 for the ZF receiver. For these results, we used a least squares (LS) channel estimator for each of the algorithms. In this approach, the channel estimate, $\hat{\boldsymbol{G}}$, becomes

$$\hat{\boldsymbol{G}} = \frac{1}{\eta\sqrt{p_0}}\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{Y}\boldsymbol{\Phi}^{*}, \tag{4.88}$$

where η is the training length, $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{A}^{\dagger}$ is the orthogonal projection onto $\boldsymbol{A}$, $\boldsymbol{Y} \in \mathbb{C}^{M\times\eta}$ is the received data during the channel estimation phase, and $\boldsymbol{\Phi} \in \mathbb{C}^{\eta\times K}$ is the orthogonal pilot matrix satisfying $\boldsymbol{\Phi}^{H}\boldsymbol{\Phi} = \eta\boldsymbol{I}$. We set $\eta = K$ and choose $\boldsymbol{\Phi}$ from among the columns of the discrete Fourier transform (DFT) matrix. Note that for the case of high-resolution quantization, $\boldsymbol{Y} = \sqrt{p_0}\boldsymbol{G}\boldsymbol{\Phi}^{T} + \boldsymbol{N}$, where the elements of $\boldsymbol{N}$ are independent $\mathcal{CN}(0,1)$ random variables. For standard one-bit and $\Sigma\Delta$ quantization, we pass $\boldsymbol{Y}$ through the corresponding quantization, and plug the output into (4.88) for channel estimation. Fig. 4.11 shows the performance of the ZF receiver with and without perfect CSI versus the number of antennas. The presence of imperfect CSI obviously degrades all of the algorithms, but we see that

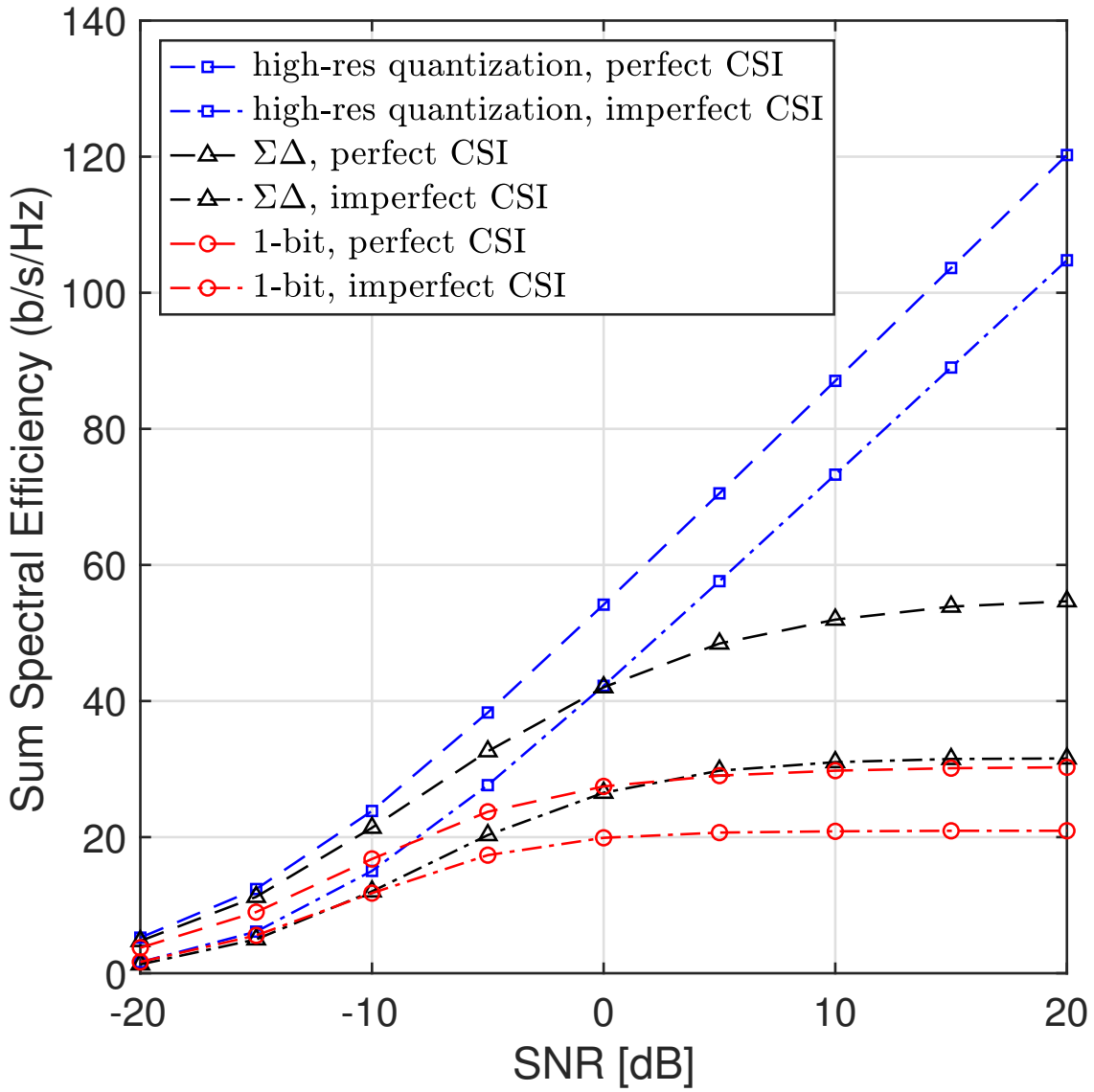

Figure 4.10: SE versus SNR for ZF receiver with and without channel estimation error. $L = 20$, $d = \lambda/4$.

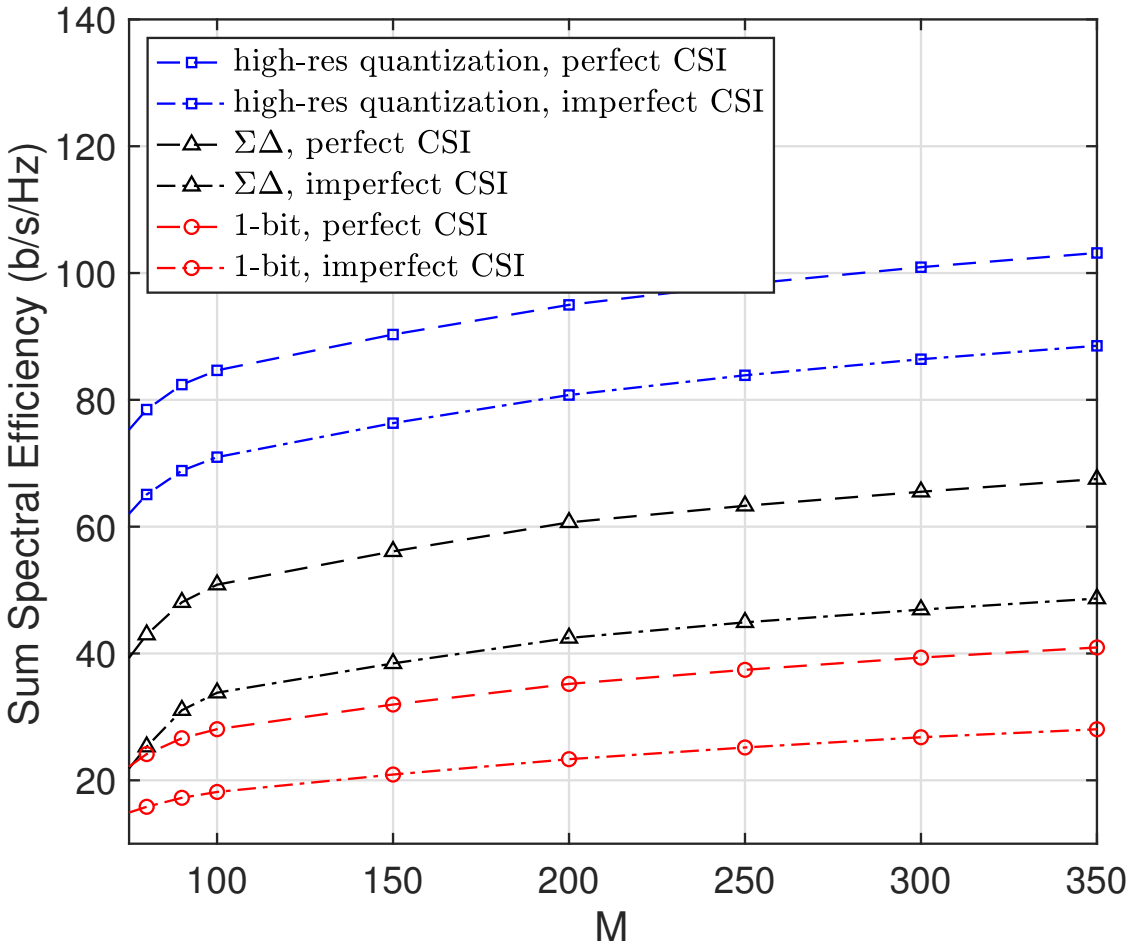

Figure 4.11: SE versus M for ZF receiver with and without channel estimation error. $L = 15$, $d = \lambda/4$, SNR = 10 dB.

the ΣΔ architecture provides a way to successfully bridge the performance gap between standard one-bit and high-resolution quantization with only a minimal increase in hardware complexity.

4.6.2 FBB ΣΔ architecture

This section describes the results of several Monte Carlo simulations in order to illustrate the performance of the FBB ΣΔ quantizer. In our simulations, we assume $L = 20$ multipath arrivals for both the legitimate user and the jammer with angles of arrival randomly uniformly

distributed in $\theta_l \in [\theta_0 - \delta, \theta_0 + \delta]$, where the center angle θ_0 is different for the user and the jammer. We set $\theta_{0_u} = -20°$ and $\theta_{0_j} = 60°$ for the desired user and jammer, respectively, with $\delta = 5°$. We further assume $d/\lambda = 1/4$, 8-PSK symbols, and 10^5 trials. The steering angle of the FBB $\Sigma\Delta$ array is set to $\phi = 2\pi\frac{d}{\lambda}\sin\left(\theta_{0_u}\right)$. We let $\sigma_n^2 = 1$, so that p and q denote the SNR of the user and the jammer, respectively. In all simulations, we consider a strong interference setting with $q = 20$ dB. We also assume that θ_{0_u} and θ_{0_j} are known at the BS[6]. Hence, following the same reasoning as in [42], the feedback weights are estimated as

$$\boldsymbol{w} = \left(e^{-j\phi}\boldsymbol{Z}_{-1}\boldsymbol{U}\bar{\boldsymbol{Y}}\right)^{\dagger}(\boldsymbol{U}\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}), \tag{4.89}$$

where

$$\bar{\boldsymbol{Y}} = \begin{bmatrix} \bar{y}_0 & & \boldsymbol{0} \\ \bar{y}_1 & \ddots & \\ \vdots & \ddots & \bar{y}_0 \\ \vdots & \vdots & \vdots \\ \bar{y}_{M-1} & \cdots & \bar{y}_{M-N} \end{bmatrix}, \ \bar{\boldsymbol{y}} = \tau\boldsymbol{a}\left(\theta_{0_u}\right) \tag{4.90}$$

$$\bar{\boldsymbol{x}} = \tau\boldsymbol{a}\left(\theta_{0_u}\right) + \boldsymbol{a}\left(\theta_{0_j}\right). \tag{4.91}$$

and τ is a constant. In the simulations that follow, we selected $\tau = \sqrt{10}$. Then, the solution in (4.89) is followed by iterative refinement (see Section III-C in [42]) to find the desired feedback weights.

Fig. 4.12 shows the simulated and analytically derived quantization noise power density

[6]Note that, although here we consider a single jammer in a known location, the feedback weights can in general be designed to reduce the impact of signals arriving from multiple sectors in which the jammers are known to lie, without precise knowledge of the actual jammer locations.

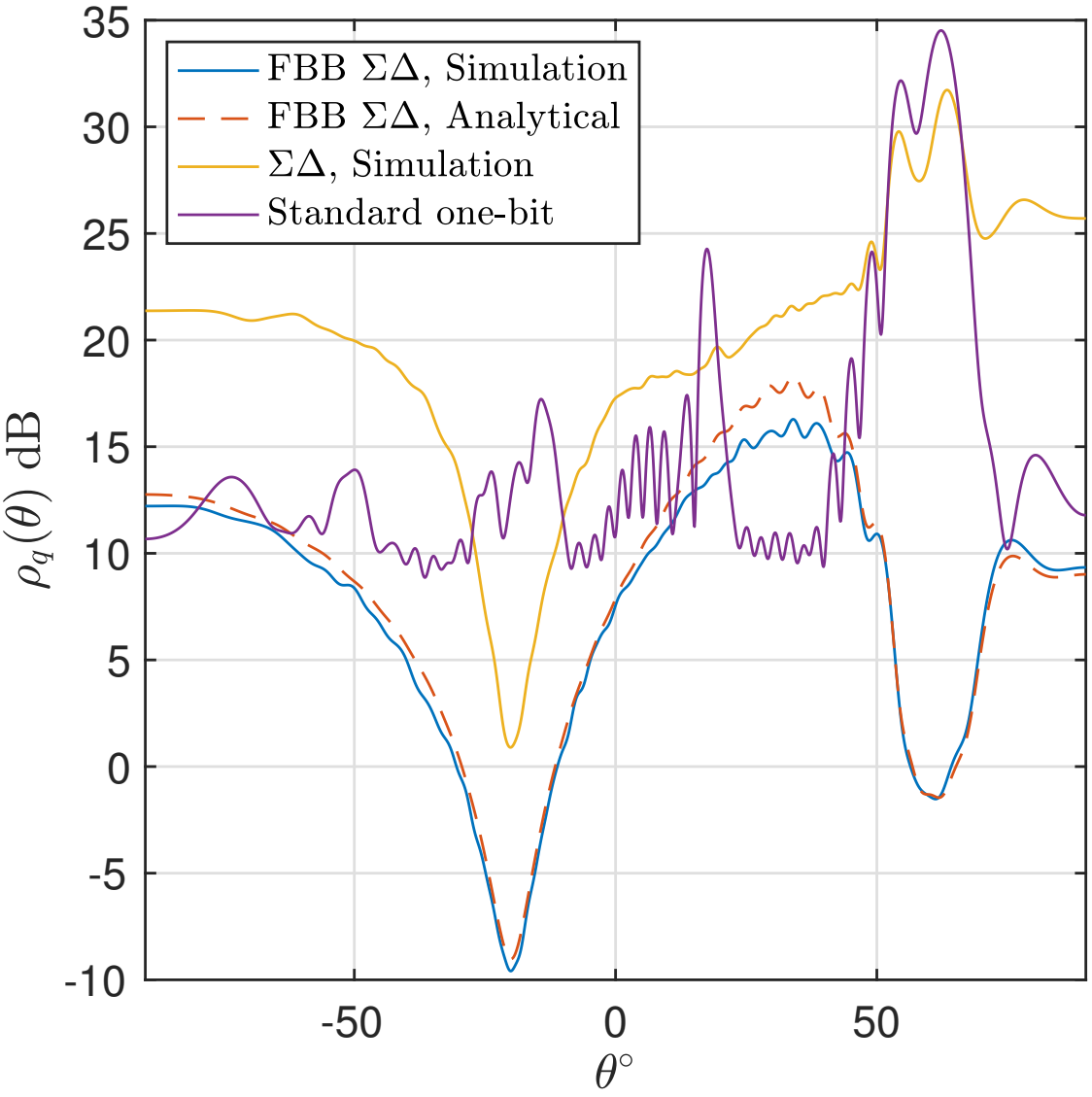

Figure 4.12: Spatial spectrum of the quantization noise for the FBB ΣΔ, regular ΣΔ, and standard one-bit architectures when $d = \lambda/4$, $p = 0$ dB, $q = 20$ dB, and $N = 50$.

which is defined as

$$\rho_q(\theta) \triangleq \frac{1}{M}\boldsymbol{a}(\theta)^H \boldsymbol{R}\ \boldsymbol{a}(\theta), \tag{4.92}$$

where $\boldsymbol{R}$ is the covariance matrix of the quantization noise for each approach (standard one-bit, $\Sigma\Delta$, or FBB $\Sigma\Delta$). In this figure, we set the order of the FBB $\Sigma\Delta$ filter at $N = 50$. We see that the quantization noise power for the FBB $\Sigma\Delta$ array is substantially lower over the angles where the user is present, while the effect is the opposite for standard one-bit quantization – the quantization noise is higher for angles where the amplitude of the received signals is larger. In addition, we see that even the ordinary $\Sigma\Delta$ array suffers from large quantization noise in the presence of strong interference. We also observe that there is excellent agreement between the simulations and our theoretically derived expressions for both cases. Note that careful design of the quantizer output levels is a critical component in achieving the desired $\Sigma\Delta$ noise shaping characteristic shown here.

In Fig. 4.13, we compare the symbol error rate of the FBB $\Sigma\Delta$ array with that of a system with high-resolution ADCs and a system with high-resolution ADCs and no strong interference, i.e., $\boldsymbol{x}_{\mathrm{p}} = \sqrt{p}\boldsymbol{g}_U s_U + \boldsymbol{n}$, as a benchmark. The methods that do not allow FBB in the RF domain must attempt to cancel the interference digitally, after the quantization. Consequently, for the systems implemented with high-resolution ADCs, standard one-bit ADCs, and the original spatial $\Sigma\Delta$ architecture, we project the sampled signal onto the subspace orthogonal to the interference in the digital domain. Denoting the signals received by the standard one-bit and $\Sigma\Delta$ architectures by $\boldsymbol{y}_1$ and $\boldsymbol{y}_{\Sigma\Delta}$, respectively, the signals after the projection for the three methods are given by $\boldsymbol{B}^{-1}\boldsymbol{U}\boldsymbol{x}$, $\boldsymbol{B}^{-1}\boldsymbol{U}\boldsymbol{y}_1$, and $\boldsymbol{B}^{-1}\boldsymbol{U}\boldsymbol{y}_{\Sigma\Delta}$. We assume perfect channel state information (CSI) is available and use the maximum ratio combiner (MRC) at the BS to decode the 8-PSK symbols. Fig. 4.13 shows the superior performance of the one-bit FBB $\Sigma\Delta$ architecture which achieves performance equivalent to that of a system with only high resolution ADCs. This performance is achieved with only minimal additional

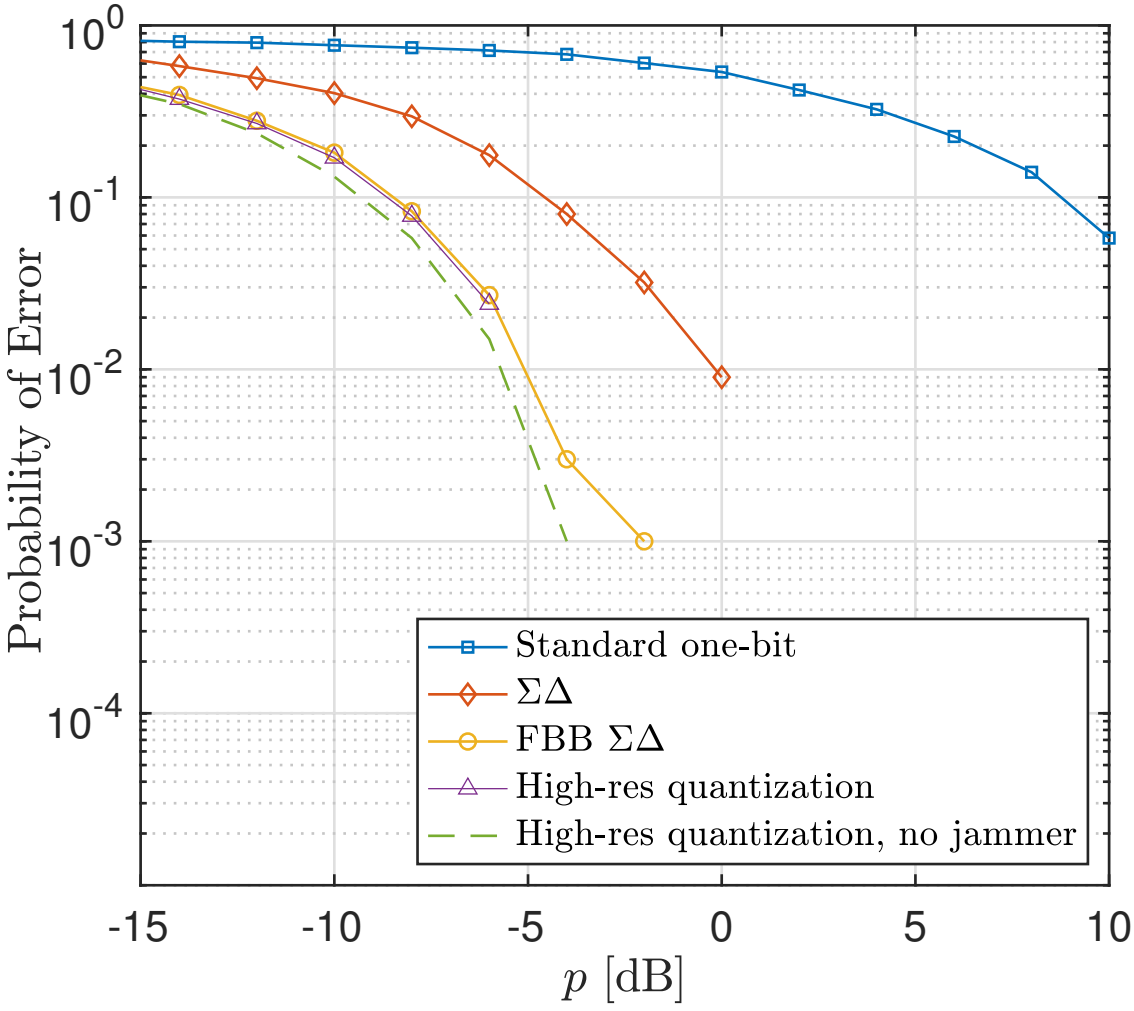

Figure 4.13: Symbol error rate versus p for a system with $\theta_{0_u} = -20°$ and $\theta_{0_j} = 60°$, $d = \lambda/4$, $q = 20$ dB, $N = 50$.

hardware in the analog domain, and thus has significantly reduced complexity and energy consumption compared with a system employing high-resolution ADCs.

4.7 Summary

In this chapter, we studied the performance of massive MIMO systems employing spatial one-bit $\Sigma\Delta$ quantization. Using an element-wise Bussgang approach, we derived an equivalent linear model in order to analytically characterize the spectral efficiency of a massive MIMO base station with a $\Sigma\Delta$ array, and we compared the results with the performance achieved by an array that employs standard one-bit quantization. Our results demonstrated that the spatial $\Sigma\Delta$ architecture can scale down the quantization noise power proportional to the square of the spatial oversampling rate. This can be interpreted as scaling down the quantization noise power proportional to the inverse square of the number of antennas at the BS for space-constrained arrays. This result gains more importance by noting that in standard one-bit quantization, the quantization noise power grows proportional to the inverse of the spatial oversampling rate, or equivalently, proportional to the number of antennas at the BS in space-constrained arrays. Furthermore, it was shown how this capability allows the spatial $\Sigma\Delta$ architecture to bridge the SE gap between infinite resolution and standard one-bit quantized systems. For the ZF receiver, the spatial $\Sigma\Delta$ architecture can outperform standard one-bit quantization by about 50%, and achieve almost the same performance as an infinite resolution system for the MRC receiver. While these results were obtained by assuming the availability of perfect CSI at the BS, we also showed that the spatial $\Sigma\Delta$ architecture is able to alleviate the adverse impact of quantization noise in the presence of channel estimation error.

We presented a new spatial one-bit FBB $\Sigma\Delta$ architecture for mitigating strong interference in massive MIMO systems with one-bit quantization. We showed that this simple architecture

can effectively compensate for the vulnerability of one-bit ADCs against strong interference. The critical challenges in designing this architecture are to find the appropriate output levels for the one-bit quantizers and the values for the feedback weights. A recursive algorithm was proposed to specify the quantizers' output levels. The feedback weights were designed by adopting an algorithm used previously for a temporal FBB $\Sigma\Delta$ implementation. However, the behaviour of the feedback weights indicates that they amount to a spatial beamformer pointing in the direction(s) of the interference, and hence could be designed by a less complicated approach. Interesting directions for future work include studying the impact of angle estimation errors on the performance of the FBB $\Sigma\Delta$ architecture, or using the approach for combined quantization noise shaping and transmit beampattern design for the downlink with low-resolution digital-to-analog converters.

Chapter 5

Impact of Mutual Coupling

One of the challenges of deploying massive MIMO systems is to fit a large number of antennas in a constrained physical space [73]. In this scenario, adding more antennas leads to increasing spatial correlation and mutual coupling among adjacent antennas which can cause severe performance degradation. Hence, it is not possible to arbitrarily increase the number of antennas at the BS to offset the SE loss. In [70], it is observed that contrary to the Rayleigh fading case, the user's channels are no longer asymptotically orthogonal in the space-constrained scenario. The SE and energy efficiency of massive MIMO systems with space-constrained arrays are shown in [74, 75] to be concave functions of the number of antennas. This demonstrates that in space-constrained arrays, it is not beneficial to increase the number of antennas beyond a certain threshold within a limited physical space. In [76], the increased correlation among adjacent antennas in the space-constrained array is leveraged to reduce the RF chain complexity. The SE of a space-constrained array with linear receivers for Rayleigh and Rician fading channels is studied in [77] and [78], respectively.

The Mixed-ADC architecture is a suitable approach to bridge the gap between massive MIMO with low- and high-resolution ADCs in space-constrained arrays. First, it does not

require an increase in the number of antennas and therefore avoids the severe performance degradation due to mutual coupling. Moreover, it can take advantage of intrinsic correlation among adjacent antennas of a space-constrained array as done in [76]. In this chapter, we analyze the performance of mixed-ADC massive MIMO with space-constrained arrays. This phenomenon can eliminate the need for round-robin training for channel estimation in mixed-ADC architectures. We first analyze the channel estimation performance of such systems. Since the channel coefficients are correlated under the space-constrained assumption, high-resolution ADCs in a mixed-ADC architecture are added to improve the channel estimation accuracy of all channel coefficients. Based on this channel estimate, we exploit a maximum-ratio-combiner (MRC) at the BS and derive a closed-form approximation for the SE which explicitly shows the contribution of the low- and high-resolution ADCs on the SE of a mixed-ADC architecture with a space-constrained array. It is shown that by equally spacing the high-resolution ADCs over a uniform linear array, one can dramatically reduce the performance gap between a system with all low-resolution and all high-resolution ADCs.

For spatial $\Sigma\Delta$ architecture which relies on spatial oversampling (antenna spacing less than half a wavelength), the impact of mutual coupling may become significant as the antenna spacing decreases. Unlike temporal oversampling, there is a limit to the amount of spatial oversampling that can be achieved, due to the physical dimensions of the antennas. In the last section of this chapter, we show that the one-bit Sigma-Delta array is particularly advantageous in space-constrained scenarios, and can still provide significant gains even in the presence of mutual coupling when the antennas are closely spaced. For very small antenna spacings, the noise shaping gain is offset by the loss due to mutual coupling, and the performance remains relatively constant; this is in contrast to a standard high-resolution ADC architecture without $\Sigma\Delta$, where the performance degrades monotonically as the antennas move closer together.

5.1 Channel Model and Mutual Coupling

Consider the uplink of a single-cell multi-user MIMO system consisting of K single-antenna users that send their signals simultaneously to a BS equipped with a uniform linear array (ULA) with M antennas. The $M \times 1$ signal received at the BS from the K users is given by

$$\boldsymbol{x} = \boldsymbol{G}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{s} + \boldsymbol{n}, \tag{5.1}$$

where $\boldsymbol{G} = [\boldsymbol{g}_1, \cdots, \boldsymbol{g}_K] \in \mathbb{C}^{M \times K}$ is the channel matrix between the users and the BS, and $\boldsymbol{P}$ is the diagonal matrix whose kth diagonal element, p_k, represents the average transmitted power of the kth user. The symbol vector transmitted by the users is denoted by $\boldsymbol{s} \in \mathbb{C}^{K \times 1}$ where $\mathbb{E}\left\{\boldsymbol{s}\boldsymbol{s}^H\right\} = \boldsymbol{I}_K$ and is drawn from a CSCG codebook independent of the other users, and, $\boldsymbol{n} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{n}})$ denotes additive CSCG receiver noise at the BS. The channel matrix can be decomposed as

$$\boldsymbol{G} = \boldsymbol{T}\boldsymbol{A}\boldsymbol{H}\boldsymbol{D}^{\frac{1}{2}}, \tag{5.2}$$

where $\boldsymbol{T}$ is the mutual coupling matrix which is defined as

$$\boldsymbol{T} = \left(\boldsymbol{I}_M + \frac{1}{R}\boldsymbol{Z}\right)^{-1} \tag{5.3}$$

where R denotes low noise amplifier (LNA) input impedance. For thin half wavelength dipoles, the elements of $\boldsymbol{Z}$ can be characterized as [79]

$$[\boldsymbol{Z}]_{ij} = 30\Big(2\mathrm{Ci}(2\pi d_{ij}) - \mathrm{Ci}(\xi_{ij} + \pi) - \mathrm{Ci}(\xi_{ij} - \pi) + j\left(-2\mathrm{Si}(2\pi d_{ij}) + \mathrm{Si}(\xi_{ij} + \pi) + \mathrm{Si}(\xi_{ij} - \pi)\right)\Big), \; i \neq j$$

$$[\boldsymbol{Z}]_{ii} = 30\big(\gamma + \log(2\pi) - \mathrm{Ci}(2\pi) + j\mathrm{Si}(2\pi)\big), \tag{5.4}$$

where d_{ij} is the distance between antennas i and j normalized by the wavelength, $\xi_{ij} = \pi\sqrt{1+4d_{ij}^2}$, and γ denotes the Euler-Mascheroni constant.

We consider a physical channel model where the angular domain is comprised of L finite directions, as in [65]. Hence, $\boldsymbol{A}$ is an $M \times L$ matrix whose ℓth column denotes the array steering vector corresponding to the direction of arrival θ_ℓ and is given by

$$\boldsymbol{a}(\theta_\ell) = \frac{1}{\sqrt{L}}[1, e^{-j2\pi d_{12}\sin(\theta_\ell)}, \cdots, e^{-j2\pi d_{1M}\sin(\theta_\ell)}]^T, \tag{5.5}$$

and $\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K] \in \mathbb{C}^{L\times K}$ is the propagation response matrix and represents the fast fading whose elements are distributed identically and independently as $\mathcal{CN}(0,1)$. Geometric attenuation and shadow fading are modeled by the diagonal matrix $\boldsymbol{D}$ whose kth diagonal element is denoted by β_k. In the remainder of the chapter, we assume that $\boldsymbol{A}$ and $\boldsymbol{D}$ are priori known at the BS [77, 65]. By considering the mutual coupling effect, the additive noise covariance matrix, $\boldsymbol{R_n}$, can be derived as

$$\boldsymbol{R_n} = \boldsymbol{T}\big(\sigma_i^2(\boldsymbol{Z}\boldsymbol{Z}^H + R_N^2\boldsymbol{I}_M) - 2R_N\Re\mathrm{e}(\rho^*\boldsymbol{Z}) + 4\mathrm{k}TB\Re\mathrm{e}(\boldsymbol{Z})\big)\boldsymbol{T}^H, \tag{5.6}$$

with $\mathbb{E}\{\boldsymbol{i}_N\boldsymbol{i}_N^H\} = \sigma_i^2\boldsymbol{I}_M$, $\mathbb{E}\{\boldsymbol{u}_N\boldsymbol{u}_N^H\} = \sigma_u^2\boldsymbol{I}_M$, $R_N = \frac{\sigma_u}{\sigma_i}$, $\frac{\mathbb{E}\{\boldsymbol{u}_N\boldsymbol{i}_N^H\}}{\sigma_u\sigma_i} = \rho\boldsymbol{I}_M$, where $\boldsymbol{i}_N$ and $\boldsymbol{u}_N$ denote the equivalent noise current and voltage of the LNA and k, $\boldsymbol{T}$, and $\boldsymbol{B}$ represent the Boltzman constant, environment temperature, and bandwidth, respectively.

5.2 Mixed-ADC Architecture

We consider a mixed-ADC architecture at the BS in which M_0 antennas are connected to high-resolution ADCs while M_1 antennas are fed to low-resolution ADCs. As a result, by partitioning the channel matrix $\boldsymbol{G}$, we can rewrite (5.1) as

$$\begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}_0 \\ \boldsymbol{G}_1 \end{bmatrix} \boldsymbol{P}^{\frac{1}{2}} \boldsymbol{s} + \begin{bmatrix} \boldsymbol{n}_0 \\ \boldsymbol{n}_1 \end{bmatrix}. \tag{5.7}$$

In (5.7), $\boldsymbol{G}_0 = \boldsymbol{T}_0 \boldsymbol{A}\boldsymbol{H}\boldsymbol{D}^{\frac{1}{2}} \in \mathbb{C}^{M_0 \times K}$ ($\boldsymbol{G}_1 = \boldsymbol{T}_1 \boldsymbol{A}\boldsymbol{H}\boldsymbol{D}^{\frac{1}{2}} \in \mathbb{C}^{M_1 \times K}$) contains the channel coefficients from the users to the M_0 (M_1) antennas connected to high-resolution (low-resolution) ADCs and $\boldsymbol{n}_0$ ($\boldsymbol{n}_1$) denotes the corresponding elements of $\boldsymbol{n}$, and similarly for $\boldsymbol{T}_0$ and $\boldsymbol{T}_1$. Therefore, the received signal at the BS is

$$\boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_0 \\ Q(\boldsymbol{x}_1) \end{bmatrix}, \tag{5.8}$$

where $Q(\cdot)$ represents the element-wise quantization operation. We adopt the additive quantization noise model (AQNM) for the quantizer output [22, 27]

$$Q(\boldsymbol{x}_1) = \alpha \boldsymbol{x}_1 + \boldsymbol{q}, \tag{5.9}$$

where $0 < \alpha < 1$ is a linear gain dependent on the number of quantization bits and $\boldsymbol{q}$ denotes the quantization noise which is uncorrelated with $\boldsymbol{y}_1$ and whose covariance matrix is given by

$$\boldsymbol{R}_{\boldsymbol{q}} = \alpha(1-\alpha)\,\mathrm{diag}\left(\boldsymbol{R}_{\boldsymbol{x}_1}\right), \tag{5.10}$$

where $\boldsymbol{R}_{\boldsymbol{x}_1}$ denotes the covariance of $\boldsymbol{x}_1$.

We assume a block-fading model where each channel remains constant in a coherence interval of length T symbols and changes independently between different intervals. At the beginning of each coherence interval, the users send their η-tuple mutually orthogonal pilot sequences ($K \leq \eta \leq T$) to the BS for channel estimation. The remaining $T - \eta$ symbols are dedicated to uplink data transmission.

5.2.1 Training Phase

In this section, we investigate the linear minimum mean squared error (LMMSE) channel estimator for different ADC architectures at the BS. In all scenarios, the pilot sequences are drawn from an $\eta \times K$ matrix $\boldsymbol{\Phi}$, where the kth column of $\boldsymbol{\Phi}$, $\boldsymbol{\phi}_k$, is the kth user's pilot sequence and $\boldsymbol{\Phi}^H\boldsymbol{\Phi} = \boldsymbol{I}_K$. Therefore, the $M \times \eta$ received signal at the BS becomes

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_0 \\ \boldsymbol{X}_1 \end{bmatrix} = \sqrt{\eta} \begin{bmatrix} \boldsymbol{G}_0 \\ \alpha \boldsymbol{G}_1 \end{bmatrix} \boldsymbol{P}^{\frac{1}{2}} \boldsymbol{\Phi}^{\boldsymbol{T}} + \begin{bmatrix} \boldsymbol{N}_0 \\ \alpha \boldsymbol{N}_1 + \boldsymbol{N}_q \end{bmatrix}, \tag{5.11}$$

where $\boldsymbol{N}_0$ ($\boldsymbol{N}_1$) is an $M_0 \times \eta$ ($M_1 \times \eta$) additive noise matrix and $\boldsymbol{N}_q$ denotes the quantization noise. The LMMSE estimate of the k-th user propagation response, $\boldsymbol{h}_k$, is

$$\hat{\boldsymbol{h}}_k = \sqrt{\eta p \beta_k} \boldsymbol{B}^H \left(\eta p_k \beta_k \boldsymbol{B}\boldsymbol{B}^H + \boldsymbol{R}_{\boldsymbol{w}} \right)^{-1} \boldsymbol{X} \boldsymbol{\phi}_k^*, \tag{5.12}$$

where

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{T}_0 \boldsymbol{A} \\ \alpha \boldsymbol{T}_1 \boldsymbol{A} \end{bmatrix}, \ \boldsymbol{R}_{\boldsymbol{w}} = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{n}_0} & \alpha \boldsymbol{R}_{\boldsymbol{n}_0 \boldsymbol{n}_1} \\ \alpha \boldsymbol{R}_{\boldsymbol{n}_1 \boldsymbol{n}_0} & (\alpha^2 \boldsymbol{R}_{\boldsymbol{n}_1} + \boldsymbol{R}_{\boldsymbol{q}}) \end{bmatrix} \tag{5.13}$$

$$\boldsymbol{R}_{\boldsymbol{q}} = \alpha(1-\alpha)\mathrm{diag}\left(\left(\sum_{k=1}^{K} p_k \beta_k \right) \boldsymbol{T}_1 \boldsymbol{A} \boldsymbol{A}^H \boldsymbol{T}_1^H + \boldsymbol{R}_{\boldsymbol{n}_1} \right). \tag{5.14}$$

By defining the channel estimation error, $\boldsymbol{\epsilon} = \boldsymbol{h}_k - \hat{\boldsymbol{h}}$, we have

$$\boldsymbol{R}_{\hat{\boldsymbol{h}}_k} = \eta p_k \beta_k \boldsymbol{B}^H \left(\eta p_k \beta_k \boldsymbol{B}\boldsymbol{B}^H + \boldsymbol{R}_{\boldsymbol{w}} \right)^{-1} \boldsymbol{B} \tag{5.15}$$

$$\boldsymbol{R}_{\boldsymbol{\epsilon}_k} = \left(\boldsymbol{I}_P + \eta p_k \beta_k \boldsymbol{B}^H \boldsymbol{R}_{\boldsymbol{w}}^{-1} \boldsymbol{B} \right)^{-1} \tag{5.16}$$

where $\boldsymbol{R}_{\hat{\boldsymbol{h}}_k}$ and $\boldsymbol{R}_{\boldsymbol{\epsilon}_k}$ denote the covariance matrix of $\hat{\boldsymbol{h}}_k$ and $\boldsymbol{\epsilon}$, respectively.

In contrast with independent Rayleigh fading channel where each high-resolution ADC only can improve the estimation accuracy of the channel associated with the antenna connected to it [26], in this model each high-resolution observation sees information about all L channel coefficients coefficients due to the coupling matrix $\boldsymbol{T}$ and array steering matrix $\boldsymbol{A}$. Hence, the high-resolution ADCs will have a stronger contribution to increasing the estimation accuracy of all channel coefficients due to the correlation among the observations. In Section 5.2.3, we numerically show to what extent adding a small number of high-resolution ADCs improves the channel estimation accuracy.

5.2.2 Spectral Efficiency

In the data transmission phase, all users simultaneously send their data symbols to the BS. Without loss of generality, we index the antennas such that the first N are connected to high-resolution ADCs. From equation (5.1), the received signal at the BS after quantization can be approximated as [27]:

$$\begin{bmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{bmatrix} \approx \begin{bmatrix} \boldsymbol{G}_0 \boldsymbol{P}^{\frac{1}{2}} \boldsymbol{s} + \boldsymbol{n}_0 \\ \alpha \boldsymbol{G}_1 \boldsymbol{P}^{\frac{1}{2}} \boldsymbol{s} + \alpha \boldsymbol{n}_1 + \boldsymbol{q}_d \end{bmatrix}, \tag{5.17}$$

where $\boldsymbol{q}_d$ denotes the quantization noise whose covariance matrix is given by:

$$\boldsymbol{R}_{\boldsymbol{q}_d} = \alpha\,(1-\alpha)\,\mathrm{diag}\left(\boldsymbol{G}_1\boldsymbol{P}\boldsymbol{G}_1^H + \boldsymbol{R}_{n_1}\right). \tag{5.18}$$

For data detection, the BS uses the MRC receiver $\boldsymbol{W} = \hat{\boldsymbol{G}}$ where $\hat{\boldsymbol{G}} = [\hat{\boldsymbol{g}}_1, ..., \hat{\boldsymbol{g}}_K]$ denotes the estimate of the channel. Therefore, the resulting signal at the BS is

$$\hat{\boldsymbol{s}} = \boldsymbol{W}^H \begin{bmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{G}}_0 \\ \hat{\boldsymbol{G}}_1 \end{bmatrix}^H \begin{bmatrix} \boldsymbol{G}_0 \\ \alpha\boldsymbol{G}_1 \end{bmatrix} \boldsymbol{P}^{\frac{1}{2}}\boldsymbol{s} + \begin{bmatrix} \hat{\boldsymbol{G}}_0 \\ \hat{\boldsymbol{G}}_1 \end{bmatrix}^H \begin{bmatrix} \boldsymbol{n}_0 \\ \alpha\boldsymbol{n}_1 + \boldsymbol{q}_d \end{bmatrix}. \tag{5.19}$$

The kth element of $\hat{\boldsymbol{s}}$ is

$$\begin{aligned} \hat{s}_k = \sqrt{p_k}\left(\hat{\boldsymbol{g}}_{0k}^H\hat{\boldsymbol{g}}_{0k} + \alpha\hat{\boldsymbol{g}}_{1k}^H\hat{\boldsymbol{g}}_{1k}\right)s_k + \sqrt{p_k}\sum_{i=1,i\neq k}^{K}\left(\hat{\boldsymbol{g}}_{0k}^H\hat{\boldsymbol{g}}_{0i} + \alpha\hat{\boldsymbol{g}}_{1k}^H\hat{\boldsymbol{g}}_{1i}\right)s_i + \\ \sqrt{p_k}\sum_{i=1}^{K}\left(\hat{\boldsymbol{g}}_{0k}^H\boldsymbol{\varepsilon}_{0i} + \alpha\hat{\boldsymbol{g}}_{1k}^H\boldsymbol{\varepsilon}_{1i}\right)s_i + \hat{\boldsymbol{g}}_{0k}^H\boldsymbol{n}_0 + \alpha\hat{\boldsymbol{g}}_{1k}^H\boldsymbol{n}_1 + \hat{\boldsymbol{g}}_{1k}^H\boldsymbol{q}_d, \end{aligned} \tag{5.20}$$

where $\hat{\boldsymbol{g}}_{0k}, \hat{\boldsymbol{g}}_{1k}, \boldsymbol{\varepsilon}_{0k} \triangleq \boldsymbol{g}_{0k} - \hat{\boldsymbol{g}}_{0k}, \boldsymbol{\varepsilon}_{1k} \triangleq \boldsymbol{g}_{1k} - \hat{\boldsymbol{g}}_{1k}$ are the kth columns of $\hat{\boldsymbol{G}}_0$, $\hat{\boldsymbol{G}}_1$, $\boldsymbol{\mathcal{E}}_0 \triangleq \boldsymbol{G}_0 - \hat{\boldsymbol{G}}_0$, and $\boldsymbol{\mathcal{E}}_1 \triangleq \boldsymbol{G}_1 - \hat{\boldsymbol{G}}_1$, respectively. The BS treats $\hat{\boldsymbol{g}}_{0k}^H\hat{\boldsymbol{g}}_{0k} + \alpha\hat{\boldsymbol{g}}_{1k}^H\hat{\boldsymbol{g}}_{1k}$ as the desired channel and the other terms of (5.20) as Gaussian noise when decoding the signal. Consequently, a lower bound for the ergodic achievable SE at the kth user can be written as [27]

$$\mathcal{S}_k = \left(1 - \frac{\eta}{T}\right)\mathbb{E}\left\{\log_2\left(1 + \frac{p_k \mid \hat{\boldsymbol{g}}_{0k}^H\hat{\boldsymbol{g}}_{0k} + \alpha\hat{\boldsymbol{g}}_{1k}^H\hat{\boldsymbol{g}}_{1k} \mid^2}{\Omega}\right)\right\}, \tag{5.21}$$

where

$$\Omega = p_k \sum_{i=1, i\neq k}^{K} |\hat{\boldsymbol{g}}_{0k}^H \hat{\boldsymbol{g}}_{0i} + \alpha \hat{\boldsymbol{g}}_{1k}^H \hat{\boldsymbol{g}}_{1i}|^2 + \sum_{i=1}^{K} p_i \beta_i \Bigg(\hat{\boldsymbol{g}}_{0k}^H \boldsymbol{A}_0 \boldsymbol{R}_{\boldsymbol{\epsilon}_i} \boldsymbol{A}_0^H \hat{\boldsymbol{g}}_{0k}^H + \alpha^2 \hat{\boldsymbol{g}}_{1k}^H \boldsymbol{A}_1 \boldsymbol{R}_{\boldsymbol{\epsilon}_i} \boldsymbol{A}_1^H \hat{\boldsymbol{g}}_{1k}^H + 2\alpha \Re \left\{ \hat{\boldsymbol{g}}_{0k}^H \boldsymbol{A}_0 \boldsymbol{R}_{\boldsymbol{\epsilon}_i} \boldsymbol{A}_1^H \hat{\boldsymbol{g}}_{1k}^H \right\} \Bigg) + \sigma_n^2 \left(\|\hat{\boldsymbol{g}}_{0k}\|^2 + \alpha^2 \|\hat{\boldsymbol{g}}_{1k}\|^2 \right) + \alpha (1-\alpha) \hat{\boldsymbol{g}}_{1k}^H \mathrm{diag} \left(p_k \left(\hat{\boldsymbol{G}}_1 - \boldsymbol{\mathcal{E}}_1 \right) \left(\hat{\boldsymbol{G}}_1 - \boldsymbol{\mathcal{E}}_1 \right)^H + \boldsymbol{R}_{\boldsymbol{n}_1} \right) \hat{\boldsymbol{g}}_{1k} \quad (5.22)$$

with $\boldsymbol{A}_t \triangleq \boldsymbol{T}_t \boldsymbol{A}$.

An approximation for the achievable SE of a mixed-ADC system with mutual coupling and MRC detection is presented in the next theorem.

Theorem 5.1. *For a mixed-ADC massive MIMO system with MRC detection and an array of thin dipoles, the SE of the kth user is*

$$\mathcal{S}_k \approx \left(1 - \frac{\eta}{T}\right) \log_2 \left(1 + \frac{\mathcal{A}}{\mathcal{I} + \mathcal{B} + \mathcal{N} + \alpha(1-\alpha)\mathcal{Q}} \right), \quad (5.23)$$

when mutual coupling effects are considered, where $\mathcal{A}$, $\mathcal{I}$, $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{Q}$ are

$$\mathcal{A} = p_k \beta_k \Bigg(|\mathrm{Tr}(\Gamma_{0k})|^2 + \mathrm{Tr}(\Gamma_{0k}^2) + \alpha^2 \left(|\mathrm{Tr}(\Gamma_{1k})|^2 + \mathrm{Tr}(\Gamma_{1k}^2) \right) + 2\alpha \Re \left\{ \mathrm{Tr}(\Gamma_{0k}\Gamma_{1k}) \mathrm{Tr}(\Gamma_{0k}) \mathrm{Tr}(\Gamma_{1k}) \right\} \Bigg) \quad (5.24)$$

$$\mathcal{I} = \sum_{i\neq k}^{K} p_i \beta_i \left(\mathrm{Tr}\left(\Gamma_{0k}\Gamma_{0i}\right) + \alpha^2 \mathrm{Tr}(\Gamma_{1k}\Gamma_{1i}) + 2\alpha \Re \left\{ \mathrm{Tr}(\Gamma_{0k}\Gamma_{1i}) \right\} \right) \quad (5.25)$$

$$\mathcal{B} = \sum_{i=k}^{K} p_i \beta_i \left(\mathrm{Tr}\left(\Gamma_{0k}\Gamma_{0i}^{\epsilon}\right) + \alpha^2 \mathrm{Tr}(\Gamma_{1k}\Gamma_{1i}^{\epsilon}) + 2\alpha \Re \left\{ \mathrm{Tr}(\Gamma_{0k}\Gamma_{1i}^{\epsilon}) \right\} \right) \quad (5.26)$$

$$\mathcal{N} = \mathrm{Tr}\left(\boldsymbol{R}_{\hat{\boldsymbol{h}}_k}(\boldsymbol{TA})^H \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{n}_0} & \alpha \boldsymbol{R}_{\boldsymbol{n}_0\boldsymbol{n}_1} \\ \alpha \boldsymbol{R}_{\boldsymbol{n}_1\boldsymbol{n}_0} & \alpha^2 \boldsymbol{R}_{\boldsymbol{n}_1} \end{bmatrix} \boldsymbol{TA}\right) \tag{5.27}$$

$$\mathcal{Q} = \left(\sum_{i\neq k}^{K} p_i\beta_i\right)\mathrm{Tr}(\tilde{\Gamma}_{1k} \odot (\boldsymbol{A}_1\boldsymbol{A}_1^H)) + 2p\beta_k\mathrm{Tr}(\tilde{\Gamma}_{1k} \odot \tilde{\Gamma}_{1k}) + p_k\beta_k\mathrm{Tr}(\tilde{\Gamma}_{1k} \odot \tilde{\Gamma}_{1k}^{\epsilon}) + \mathrm{Tr}\left(\boldsymbol{R}_{\hat{\boldsymbol{h}}_k}(\boldsymbol{T_1A})^H\boldsymbol{R}_{\boldsymbol{n}_1}\boldsymbol{T_1A}\right) \tag{5.28}$$

$$\Gamma_{tk} = \boldsymbol{R}_{\hat{\boldsymbol{h}}_k}\boldsymbol{A}_t^H\boldsymbol{A}_t,\ t \in \{0,1\},\ \Gamma_{tk}^{\epsilon} = \boldsymbol{R}_{\boldsymbol{\epsilon}_k}\boldsymbol{A}_t^H\boldsymbol{A}_t,\ t \in \{0,1\}$$

$$\tilde{\Gamma}_{tk} = \boldsymbol{A}_t\boldsymbol{R}_{\hat{\boldsymbol{h}}_k}\boldsymbol{A}_t^H,\ t \in \{0,1\},\ \tilde{\Gamma}_{tk}^{\epsilon} = \boldsymbol{A}_t\boldsymbol{R}_{\boldsymbol{\epsilon}_k}\boldsymbol{A}_t^H,\ t \in \{0,1\}.$$

Proof. From [27, 71], an approximation for (5.21) can be calculated as

$$\mathcal{S}_k = \left(1 - \frac{\eta}{T}\right)\log_2\left(1 + \frac{p_k\mathbb{E}\left\{|\ \hat{\boldsymbol{g}}_{0k}^H\hat{\boldsymbol{g}}_{0k} + \alpha\hat{\boldsymbol{g}}_{1k}^H\hat{\boldsymbol{g}}_{1k}\ |^2\right\}}{\mathbb{E}\left\{\Omega\right\}}\right). \tag{5.29}$$

Using Lemma 2 of [11], the expected values of the desired signal, interference, and noise can be easily calculated. To calculate the expected value of the last term in (5.22), we have

$$\mathbb{E}\left\{\hat{\boldsymbol{g}}_{1k}^H\mathrm{diag}\left(p_k\left(\hat{\boldsymbol{G}}_1 - \boldsymbol{\mathcal{E}}_1\right)\left(\hat{\boldsymbol{G}}_1 - \boldsymbol{\mathcal{E}}_1\right)^H + \boldsymbol{R}_{n1}\right)\hat{\boldsymbol{g}}_{1k}\right\} = p_k\sum_{m=1}^{M_1}\sum_{i\neq k}^{K}\mathbb{E}\left\{|\hat{g}_{1km}|^2\right\}\mathbb{E}\left\{|g_{1im}|^2\right\} + p_k\sum_{m=1}^{M_1}\mathbb{E}\left\{|\hat{g}_{1km}|^4\right\} + \mathbb{E}\left\{|\hat{g}_{1km}|^2|\varepsilon_{1km}|^2\right\} + \mathbb{E}\left\{\hat{\boldsymbol{g}}_{1k}^H\boldsymbol{R}_{\boldsymbol{n}_1}\hat{\boldsymbol{g}}_{1k}\right\}.$$

Given that $\mathbb{E}\left\{|g_{1im}|^2\right\} = \beta_i \left[\boldsymbol{A}_1 \boldsymbol{A}_1^H\right]_{mm}$, we need to calculate the expected values of $\sum_{m=1}^{M} |\hat{g}_{1km}|^2$, $\sum_{m=1}^{M} |\hat{g}_{1km}|^4$, and $\sum_{m=1}^{M} |\hat{g}_{1km}|^2 |\varepsilon_{1km}|^2$. Since $\hat{g}_{1km}$ is a complex normal random variable with variance $\mathbb{E}\left\{|\hat{g}_{1km}|^2\right\} = \beta_k \left[\tilde{\Gamma}_{1k}\right]_{mm}$, we have $\mathbb{E}\left\{|\hat{g}_{1km}|^4\right\} = 2\beta_k^2 \left[\tilde{\Gamma}_{1k}\right]_{mm}^2$. Hence

$$\sum_{m=1}^{M_1} \mathbb{E}\left\{|\hat{g}_{1km}|^4\right\} = 2\beta_k^2 \mathrm{Tr}(\tilde{\Gamma}_{1k} \odot \tilde{\Gamma}_{1k}). \tag{5.30}$$

The expected value of $\sum_{m=1}^{M} |\hat{g}_{1km}|^2 |\varepsilon_{1km}|^2$ can be derived similarly. ■

In the next section, we show how adding a small number of high-resolution ADCs impacts the channel estimation accuracy and SE of the system.

5.2.3 Numerical Results

In this section, we numerically investigate the performance of the space-constrained mixed-ADC massive MIMO system in terms of channel estimation error and SE. We assume static-aware power control in the network [51], so that $p_k = p_0/\beta_k$. Furthermore, we assume $K = \eta = 10$, and a uniform linear array with uniformly distributed angle of arrivals, i.e., $\theta_\ell = -\pi/3 + (\ell - 1)\pi/L,\ \ell = 1,\ \ldots,\ L$ as in [65]. The normalized distance between antenna elements is denoted by $d = d_{\mathrm{ap}}/M$ where d_{ap} is the normalized array length. In all figures, we assume $T = 200$, $L = 20$, and SNR $\triangleq \frac{p_0}{\boldsymbol{R}_{\boldsymbol{n}}(1,1)}$.

Fig. 5.1 shows the advantage of the mixed-ADC architecture in the channel estimation phase for a system with $M = 100$ antennas and $d_{\mathrm{ap}} = 10$. As a benchmark, we have sketched the channel estimation error, i.e., $\frac{1}{L}\mathrm{Tr}(\boldsymbol{R}_{\boldsymbol{\epsilon}})$, of a system with all 1-bit ADCs and all-high-resolution ADCs. "Co-located" denotes a mixed-ADC system with $M_0 = 20$ high-resolution ADCs which are embedded in the array side by side. "Equidistant" refers to a mixed-ADC system where the high-resolution ADCs are equally spaced over the array. It can be seen that the mixed-ADC architecture with equidistant spacing reduces the gap between the

all-high-resolution architecture and the all-1-bit architecture by one half at high SNRs.

As mentioned above, one approach to compensate for SE loss when all antennas are connected to low-resolution ADCs is to increase the number of antennas. However, this is true only for the case that there is no constraint on the size of the array. As we show in the next example, when the array size is constrained, increasing the number of antennas can lead to saturation in SE. Therefore, the only viable approach to achieve a predefined SE seems to be through a mixed-ADC architecture.

Fig. 5.2 shows the impact of increasing the number of antennas for different array lengths d_{ap}. It can be seen that as long as the antenna spacing is greater than half a wavelength, the SE increases monotonically as a function of the number of antennas at the BS. Eventually, the SE of the system saturates and is maximized at a certain number of antennas.

5.3 ΣΔ architecture

The channel estimation analysis of massive MIMO systems with ΣΔ architecture is studied in [47]. This analysis could be easily extended to the case that mutual coupling is taken into account. It is shown that, especially for low to moderate SNRs, ΣΔ architecture can provide channel estimation accuracy almost similar to that of high-resolution massive MIMO architectures. Therefore, we assume perfect CSI for spectral efficiency analysis of ΣΔ architecture with mutual coupling.

5.3.1 Spectral Efficiency

Due to the complicated structure of the mutual coupling matrix in (5.3) and the quantization noise shaping matrix $\boldsymbol{U}^{-1}$, a closed-form expression for the spectral efficiency (SE), if it exists,

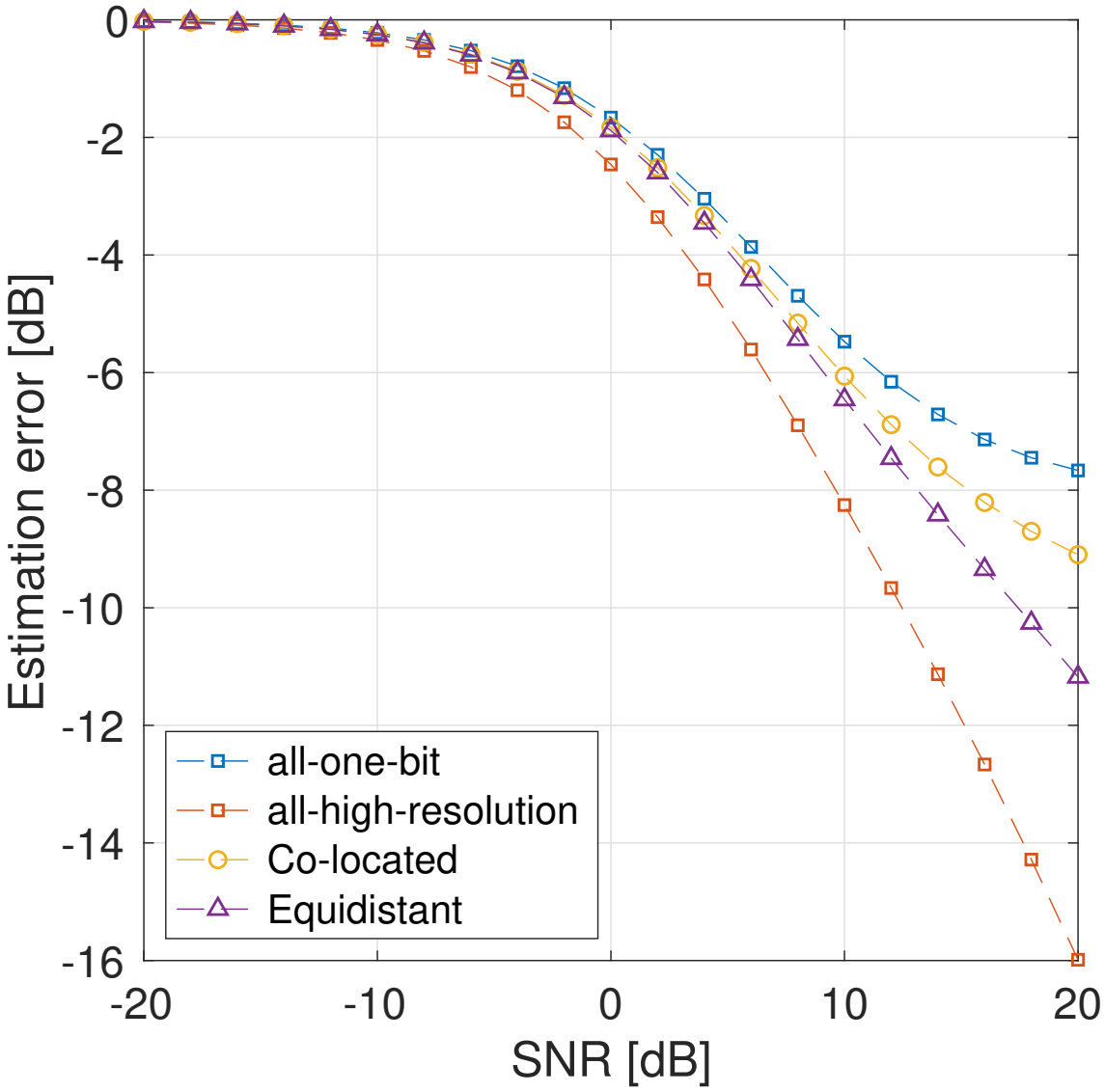

Figure 5.1: Channel estimation error versus SNR.

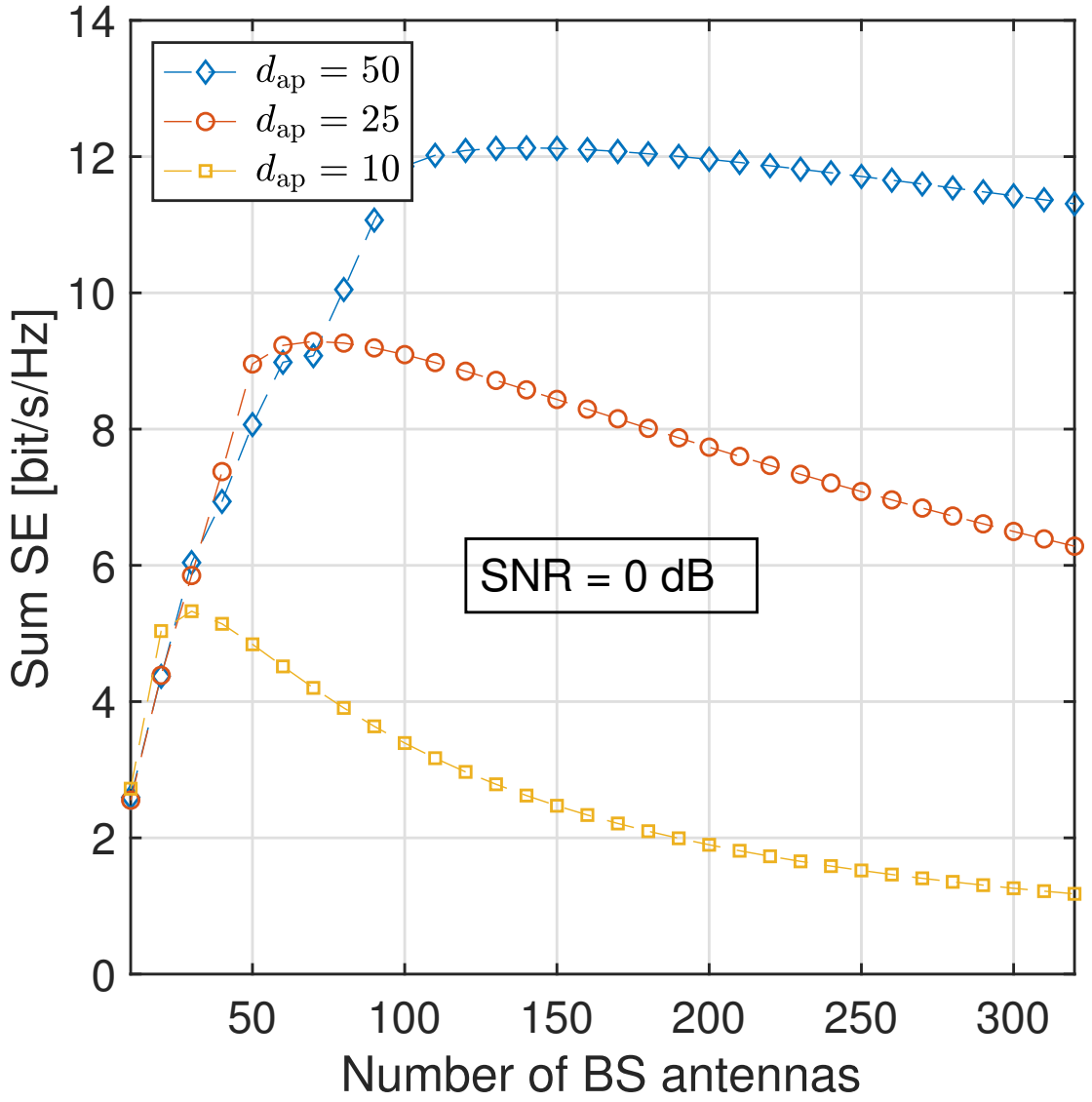

Figure 5.2: Sum SE versus the number of BS antennas, M, for SNR = 0 dB, and different d_{ap}. 20% of the antennas are connected to high-resolution ADCs.

would likely not provide significant insight into its behavior with respect to antenna spacing, nor would it provide a tool for the purpose of optimization. Hence, in the next section we numerically evaluate the SE of the system.

The received $\Sigma\Delta$-quantized signal, $\boldsymbol{y}$, at the BS is

$$\boldsymbol{y} = Q(\boldsymbol{x}) = \boldsymbol{G}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{s} + \boldsymbol{n} + \boldsymbol{U}^{-1}\boldsymbol{q} \ . \tag{5.31}$$

The total effective noise $\boldsymbol{\eta} = \boldsymbol{n} + \boldsymbol{U}^{-1}\boldsymbol{q}$ has covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}} = \boldsymbol{R}_{\boldsymbol{n}} + \boldsymbol{U}^{-1}\boldsymbol{R}_{\boldsymbol{q}}\boldsymbol{U}^{-H}$. We assume the BS employs a linear receiver $\boldsymbol{W}$, and we will consider the case of maximum ratio combining (MRC) and zero-forcing (ZF). For MRC, we do not account for the fact that $\boldsymbol{R}_{\boldsymbol{\eta}}$ is spatially colored, since pre-whitening $\boldsymbol{G}$ destroys the approximate orthogonality of the array response and increases the inter-user interference. Thus, for MRC we set $\boldsymbol{W} = \boldsymbol{G}$. However, knowledge of $\boldsymbol{R}_{\boldsymbol{\eta}}$ can be exploited by the ZF receiver, and thus we assume the pre-whitened solution $\boldsymbol{W} = \boldsymbol{R}_{\boldsymbol{\eta}}^{-1}\boldsymbol{G}(\boldsymbol{G}^H\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}\boldsymbol{G})^{-1}$.

For either receiver, the detected symbol vector is

$$\hat{\boldsymbol{s}} = \boldsymbol{W}^H\boldsymbol{y} = \boldsymbol{W}^H\boldsymbol{G}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{s} + \boldsymbol{W}^H\boldsymbol{n} + \boldsymbol{W}^H\boldsymbol{U}^{-1}\boldsymbol{q}. \tag{5.32}$$

The k-th detected symbol can be written as

$$\hat{s}_k = \sqrt{p_k}\boldsymbol{w}_k^H\boldsymbol{g}_k s_k + \sqrt{p_k}\boldsymbol{w}_k^H \sum_{i \neq k} \boldsymbol{g}_i s_i + \boldsymbol{w}_k^H\boldsymbol{n} + \boldsymbol{w}_k^H\boldsymbol{U}^{-1}\boldsymbol{q}, \tag{5.33}$$

where $\boldsymbol{w}_k$ is the kth column of $\boldsymbol{W}$. We assume the BS treats $\boldsymbol{w}_k^H\boldsymbol{g}_k$ as the desired signal and the other terms of (5.33) as worst-case Gaussian noise when decoding the signal. Consequently, a lower bound for the ergodic achievable SE at the kth user can be written as [71]

$$\mathcal{S}_k = \mathbb{E}\left[\log_2\left(1 + \frac{p_k \left|\boldsymbol{w}_k^H \boldsymbol{g}_k\right|^2}{\sum_{i\neq k} p_i |\boldsymbol{w}_k^H \boldsymbol{g}_i|^2 + \left|\boldsymbol{w}_k^H \boldsymbol{n}\right|^2 + \left|\boldsymbol{w}_k^H \boldsymbol{U}^{-1} \boldsymbol{q}\right|^2}\right)\right] . \tag{5.34}$$

5.3.2 Numerical Results

In this section, we numerically evaluate the SE performance of the ΣΔ massive MIMO system for various scenarios. Similar to previous section, we assume static-aware power control, i.e., $p_k = p_0/\beta_k$. In all of the cases considered, unless otherwise noted, we assume $K = 10$ users and equally spaced antennas with normalized spacing d. The DoAs for each user are drawn uniformly from the interval $[\theta_0 - \delta, \theta_0 + \delta]$, and the center angle of the ΣΔ array is steered towards $\phi = 2\pi d \sin(\theta_0)$.

To highlight the impact of mutual coupling, we will compare the performance when mutual coupling is included to that when it is hypothetically absent. Simulating the case without mutual coupling amounts to setting $\boldsymbol{Z} = R\boldsymbol{I}$ in (5.3) and (5.6), which leads to $\boldsymbol{T} = \frac{1}{2}\boldsymbol{I}$ and $\boldsymbol{R_n} = \sigma_n^2 \boldsymbol{I}$, with the noise power given by

$$\sigma_n^2 = \frac{1}{4}\left[\sigma_i^2\left(R^2 + R_N^2 - 2R_N R\Re(\rho)\right) + 4kTBR\right].$$

The factor of 1/2 in $\boldsymbol{T}$ results from the fact that $\boldsymbol{x}$ in (5.1) is the voltage on a load matched to the antenna impedance. This voltage is half of the antenna open-circuit voltage, and given that the load represents the input to the LNA, it is the signal available for further processing. Thus, the per-antenna and per-user reference signal-to-noise-ratio (SNR) in the

absence of mutual coupling is given by

$$\text{SNR} \triangleq \frac{1}{4}\frac{p_0}{\sigma_n^2}. \tag{5.35}$$

The circuit parameters used in (5.3) and (5.6) are defined as $\sigma_i^2 = 2kTB/R$, and $\sigma_u^2 = 2kTBR$, leading to $R_N = R$ where $R = 50\ \Omega$, $T = 290$ K, $\rho = 0$, and $B = 20$ MHz. This leads to a value of $\sigma_n^2 = 2kTBR$, where the factor of 2 appears because we are accounting for noise in both the antennas and the LNAs. We further assume CSCG symbols and 10^4 Monte Carlo trials for the simulations.

In Fig. 5.3, we investigate the impact of the mutual coupling matrix, $\boldsymbol{T}$, on the spatial spectrum of the quantization noise when $\theta_0 = 0$. To do so, we define the quantization noise power density as

$$\rho_q(\theta) \triangleq \frac{1}{c(\theta)}\mathbb{E}\left[\left|\boldsymbol{a}(\theta)^H \boldsymbol{T}^H \boldsymbol{U}^{-1}\boldsymbol{q}\right|^2\right], \tag{5.36}$$

where $c(\theta) = \|\boldsymbol{T}\boldsymbol{a}(\theta)\|^2$ is a normalizing factor and $\theta \in [-90°, 90°]$ denotes the DoA. We see that the noise shaping characteristic of the $\Sigma\Delta$ array is not significantly affected by the mutual coupling, except for the case of $d = \lambda/8$, where we see a small shift in the quantization noise spectrum.

In Fig. 5.4, we show the effect of antenna spacing on the SE of a system with an MRC receiver. We see that, when there is no constraint on the size of the array, better performance for the standard one-bit architecture can be achieved by moving the antennas farther apart. We see that the standard one-bit architecture outperforms the $\Sigma\Delta$ array when $d > \lambda/2$, due to the fact that increasing the antenna spacing increases the quantization noise power for the $\Sigma\Delta$ architecture across the DoA sector of interest, as observed in Fig. 5.3. Furthermore, we see that the SE for the $\Sigma\Delta$ architecture is not monotonic and $d = \lambda/2$ provides the best performance, which corresponds to no oversampling. The optimal value of d for the $\Sigma\Delta$

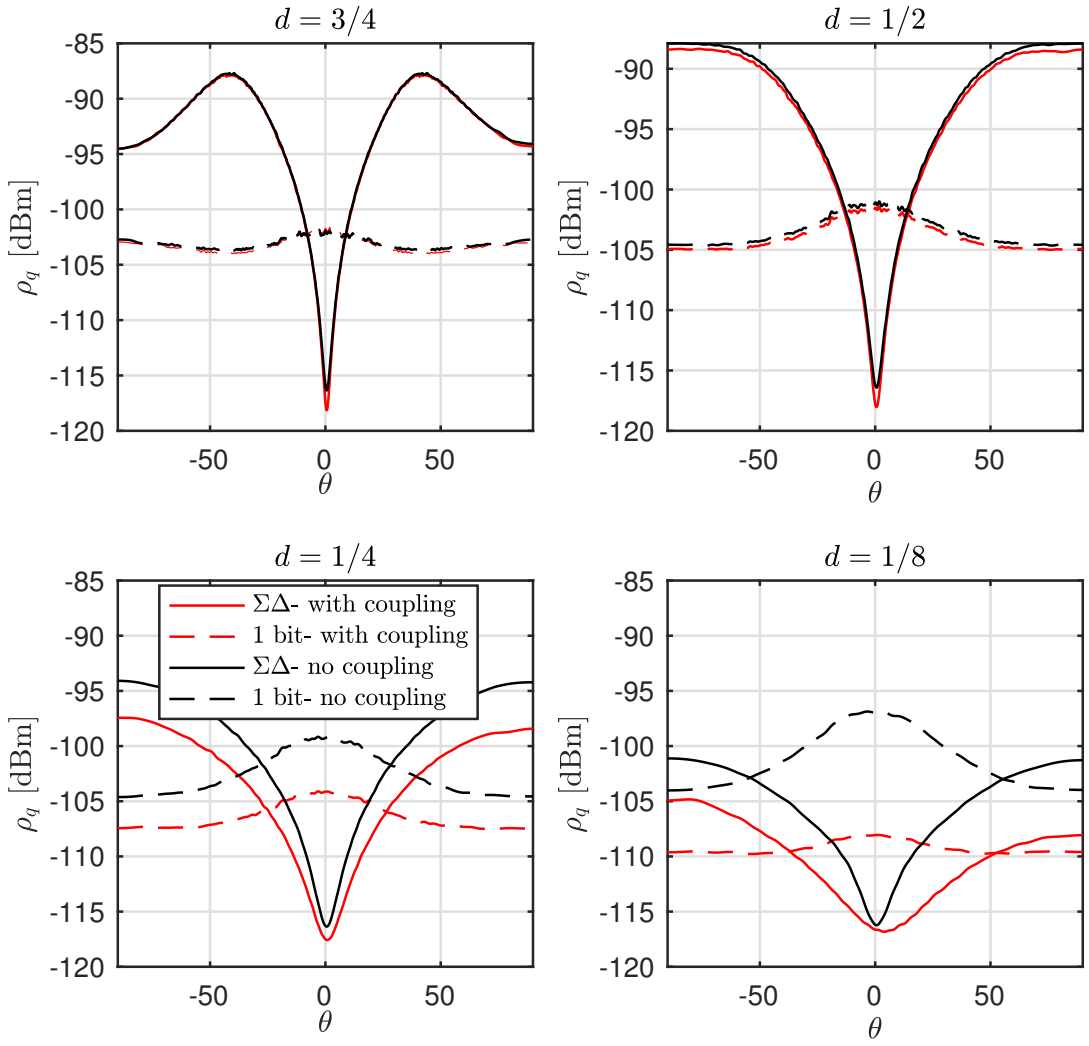

Figure 5.3: Quantization noise power density for a system with $\theta_0 = 0°$, $2\delta = 40°$, SNR = 0 dB, $M = 100$, $L = 15$.

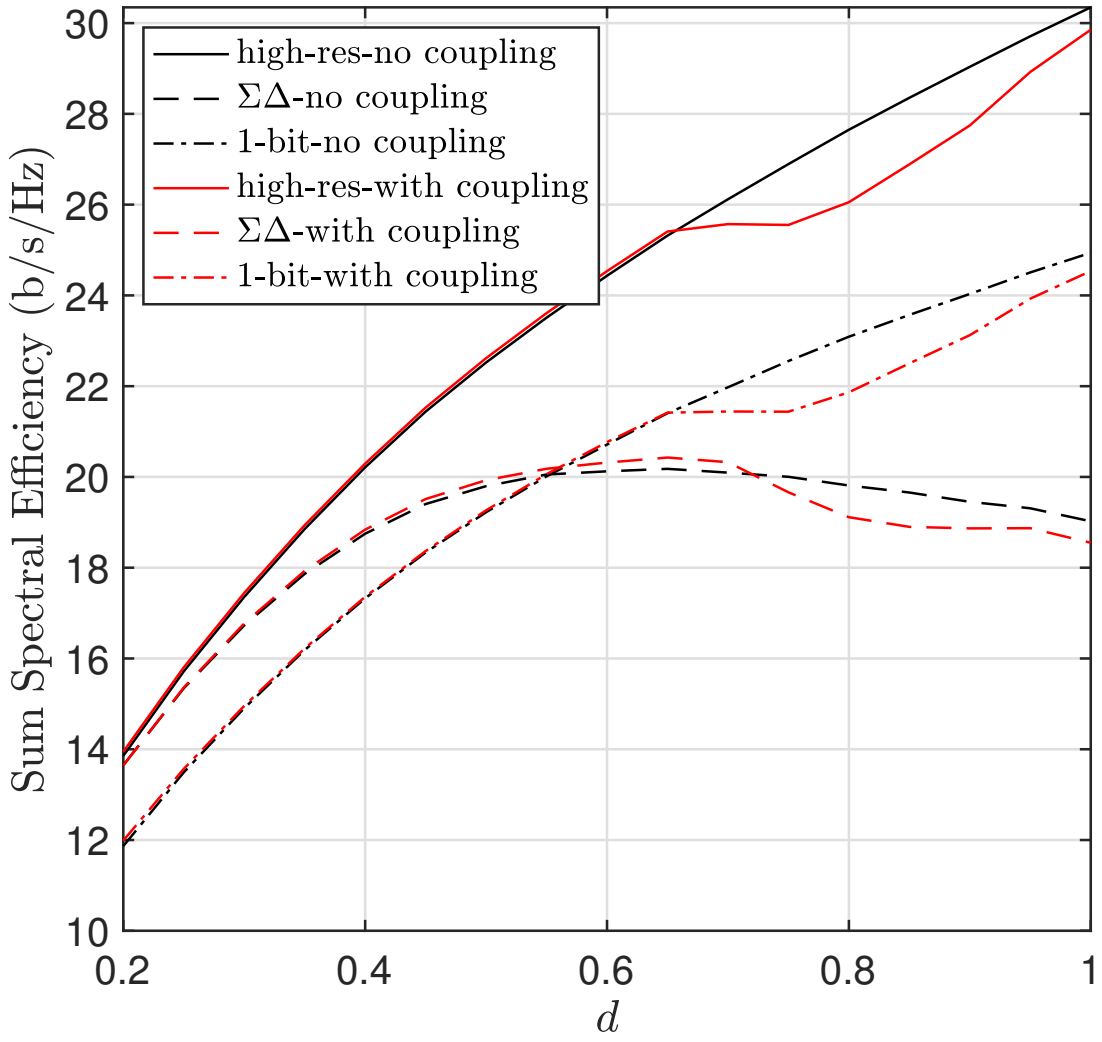

Figure 5.4: SE versus antenna spacing for a system with an MRC receiver and $\theta_0 = -10°$, angular sector $2\delta = 40°$, SNR = 10 dB, $M = 100$, $L = 15$.

array will of course decrease if the sector of user DoAs was widened.

The SE results for the ZF receiver are shown in Fig. 5.5. We again observe the degradation of the $\Sigma\Delta$ performance as d increases, but in this case there is a more significant gain relative to standard one-bit quantization for smaller antenna spacings, and the optimal antenna spacing for the $\Sigma\Delta$ array is reduced to approximately $d = \lambda/3$.

While the standard one-bit architecture can outperform the $\Sigma\Delta$ approach when there is no constraint on the dimension of the array (large d), Figs. 5.6 and 5.7 demonstrate that the $\Sigma\Delta$ array provides a better result in space-constrained scenarios. For these simulations, we consider a case in which the antenna array has a limited aperture of $d_0 = 50$ and we increase the number of antennas from $M = 100$ to $M = 400$, which corresponds to a decrease in antenna spacing from $d = 1/2$ to $d = 1/8$. For the case of an MRC receiver in Fig. 5.6, the $\Sigma\Delta$ architecture achieves a spectral efficiency nearly equal to that of an array with full-resolution ADCs when $M \geq 250$. For the case of ZF, $\Sigma\Delta$ provides a dramatic gain in SE over standard one-bit quantization.

In Fig. 5.8, the optimal antenna spacing for the $\Sigma\Delta$ architecture with a ZF receiver is shown for different SNRs, where we have quantized d to the nearest value of $\lambda/10$. It can be seen that the optimal spacing is dependent on the SNR and DoA region width, δ. The optimal antenna spacing decreases as SNR increases, and also as the size of the DoA sector of interest increases. We expect this phenomenon since for wider DoA sectors, a wider noise shaping characteristic is required to achieve the best performance. The same general conclusion holds true for the case with the MRC receiver.

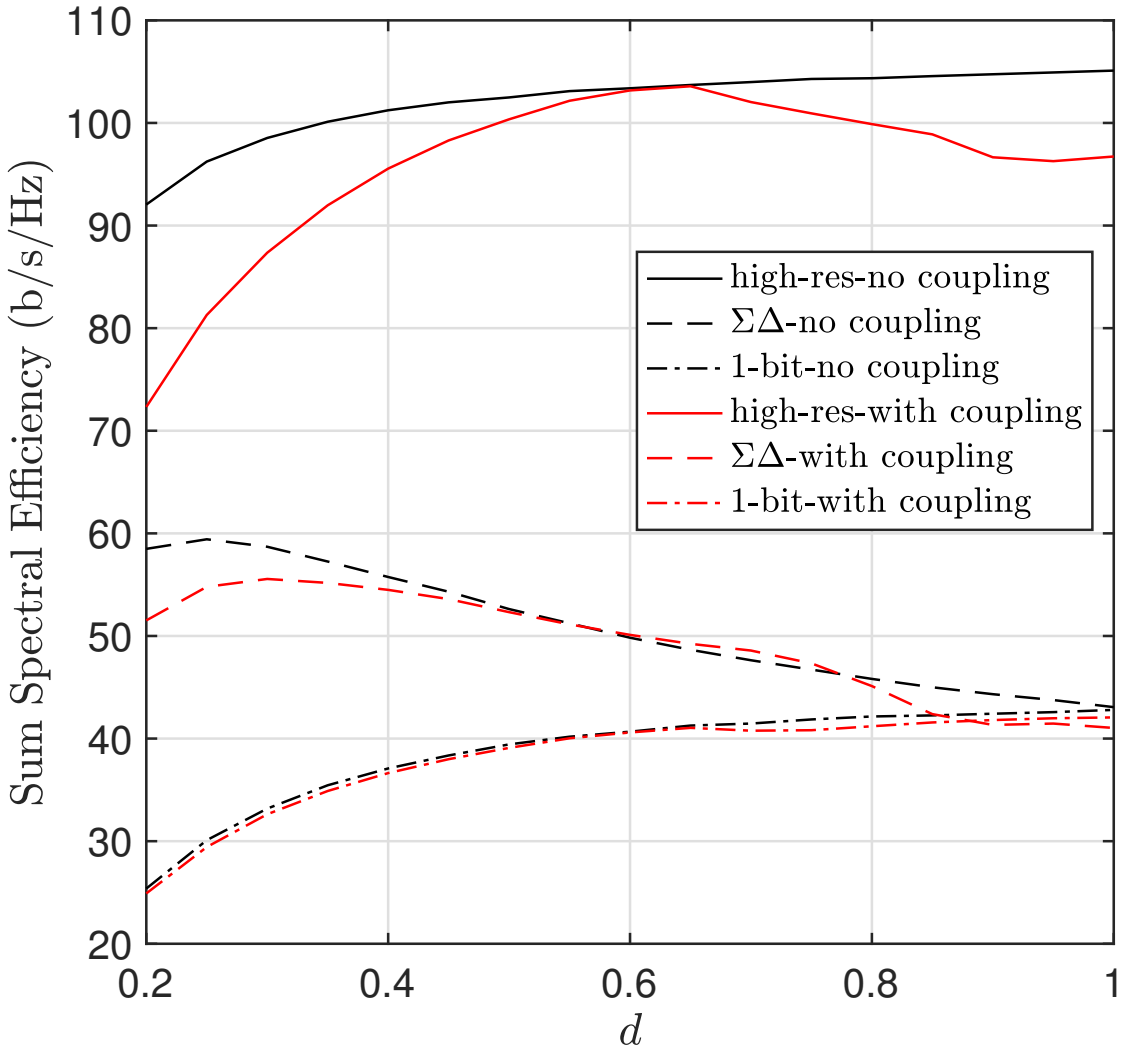

Figure 5.5: SE versus antenna spacing for a system with ZF receiver and $\theta_0 = -10°$, angular sector $2\delta = 40°$, SNR = 10 dB, $M = 100$, $L = 15$.

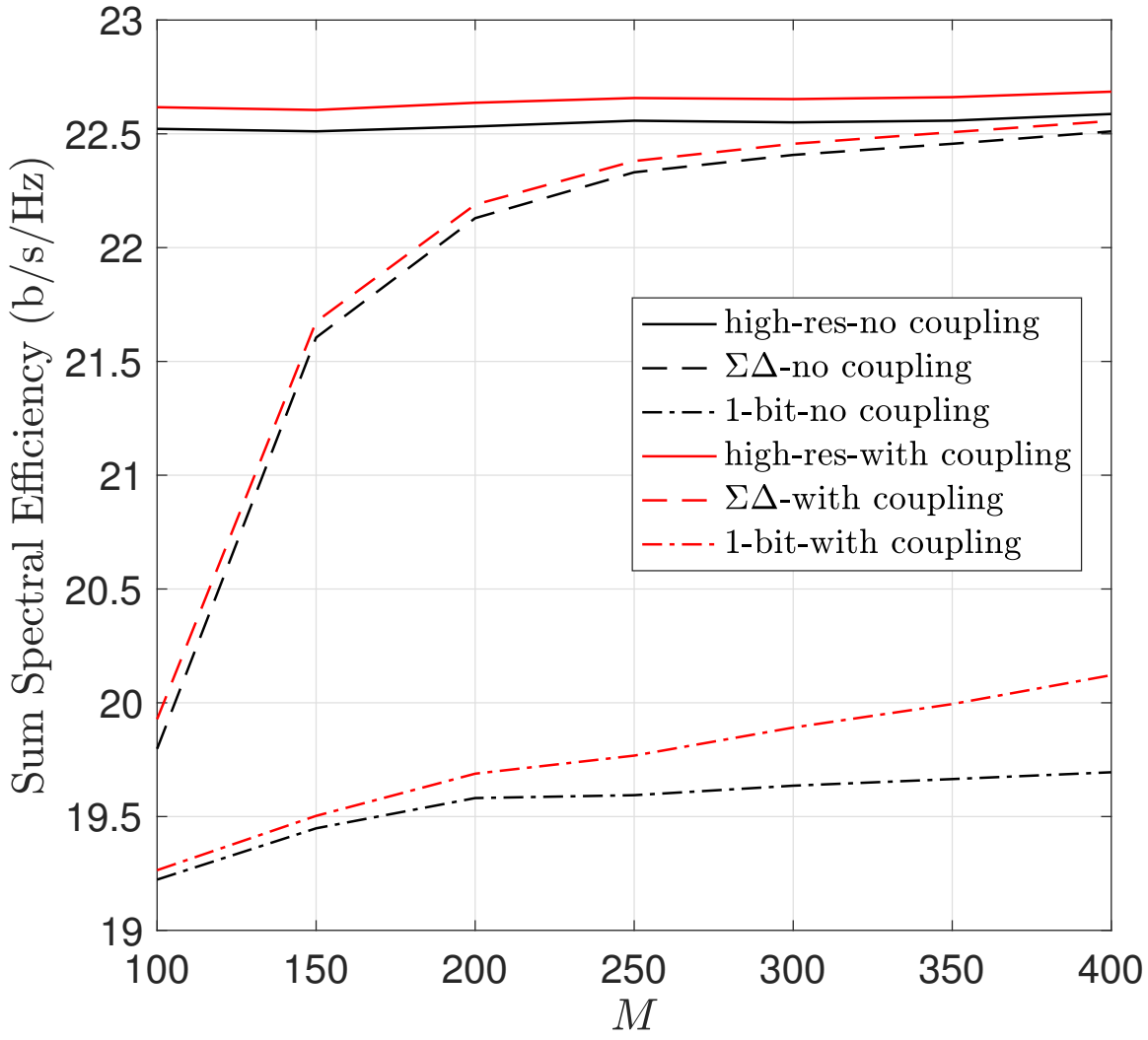

Figure 5.6: SE versus number of BS antennas for a system with MRC receiver and $\theta_0 = -10°$, angular sector $2\delta = 40°$, SNR = 10 dB, $d_0 = 50$, $L = 15$.

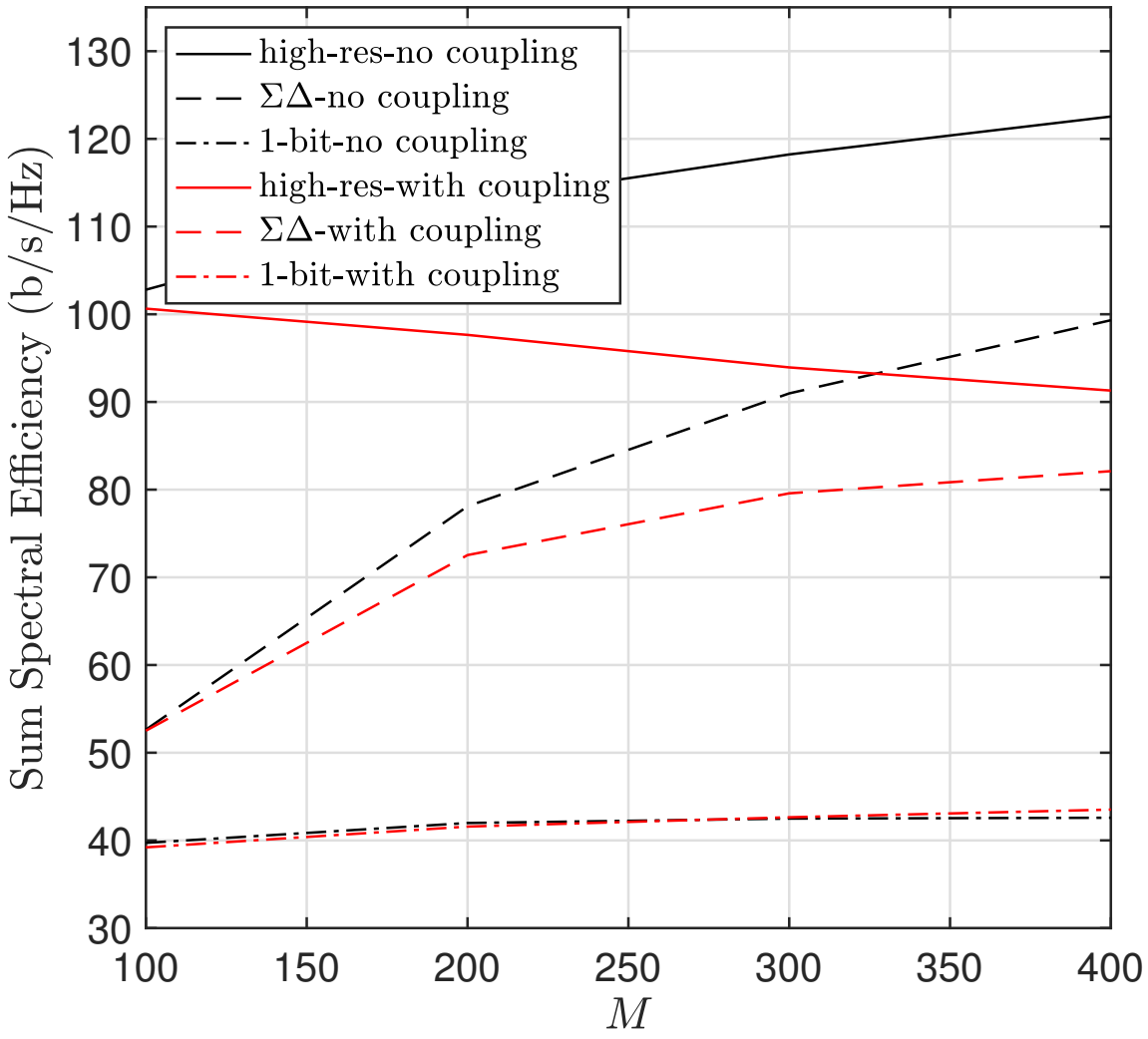

Figure 5.7: SE for a system with ZF receiver and $\theta_0 = -10°$, angular sector $2\delta = 40°$, SNR = 10 dB, $d_0 = 50$, $L = 15$.

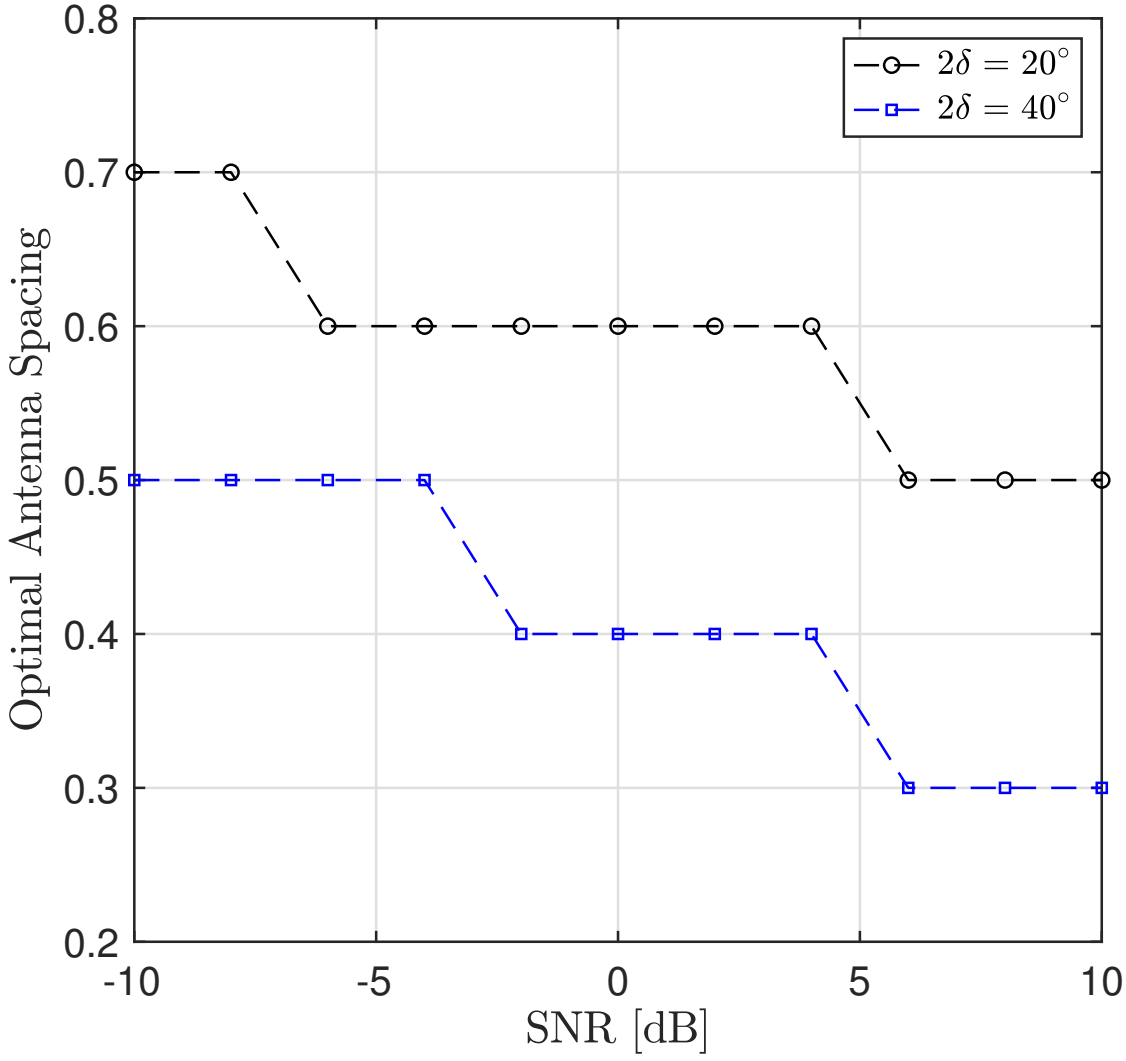

Figure 5.8: Optimal antenna spacing versus SNR for a system with ZF receiver and $\theta_0 = -10°$, $M = 100$, $L = 15$.

5.4 Summary

We have studied the effect of mutual coupling on the performance of mixed-ADC and one-bit $\Sigma\Delta$ massive MIMO systems. For mixed-ADC architecture, we see that by equally spacing a small number of high-resolution ADCs over the array, one can dramatically reduce the performance gap between a system with all low-resolution and all high-resolution ADCs. For $\Sigma\Delta$ architecture, it was shown that this architecture is most suitable for array deployments with an aperture size constraint. While the performance of standard one-bit quantization saturates as the number of antennas increases in a constrained-aperture array, the performance of the $\Sigma\Delta$ architecture tends to approach that of a system with high-resolution ADCs. This is due to the noise-shaping gain achieved by the $\Sigma\Delta$ architecture when the users are sectorized or the array is oversampled in space. It is worthwhile to note that the inevitable power loss due to mutual coupling can to some extent be alleviated using, for example, a matching network. This is a subject of future investigation.

Chapter 6

Conclusion

In this work, two main approaches in bridging the performance gap between massive MIMO systems with high-resolution and standard one-bit ADCs were investigated. The underlying theme of this dissertation has been taking the advantage of signal processing techniques in design and performance improvement of massive MIMO systems with coarse quantization.

In Chapter 2, we have studied the spectral efficiency of the mixed-ADC massive MIMO uplink under energy constraint and assuming MRC processing. We derived a closed form expression for the SE assuming the channel is estimated using both high-resolution and one-bit ADCs, without using the training-intensive round-robin approach. We then maximized the SE under a constraint on the power budget, and showed that in fact the mixed-ADC approach is not optimal. Typical power consumption models indicate that a system with all one-bit ADCs provides the highest SE for a given uplink power budget.

In Chapter 3, we studied the spectral efficiency of mixed-ADC massive MIMO systems with either MRC or ZF detection. We showed that properly accounting for the impact of the quantized receivers using the Bussgang decomposition is important for obtaining an accurate analysis of the SE. We introduced a joint channel estimation approach to leverage both

high-resolution ADCs and one-bit ADCs and our analytical and numerical results confirmed the benefit of joint channel estimation for low SNRs. Mixed-ADC detection with MRC and ZF detectors and antenna selection were also studied. Analytical expressions were derived for MRC detection and a numerical performance analysis was performed for ZF detection. It was shown that antenna selection provides a slight advantage for high SNRs while this advantage tends to disappear for low SNRs. We showed that the SNR, the number of high-resolution ADCs and the length of the coherence interval play a pivotal role in determining the performance of mixed-ADC systems. We showed that, in general, mixed-ADC architectures will have the greatest benefit compared to implementations with all low-resolution ADCs when ZF detection is used and the SNR is relatively high. In such cases, the gain of the mixed-ADC approach can be substantial. Gains are also possible for MRC, but they are not as significant, and require larger numbers of high-resolution ADCs to see a benefit compared with the ZF case. The more complicated mixed-ADC approach based on ADC switching and round-robin training can achieve the best performance in some cases, particularly when the coherence interval is long and more high-resolution ADCs are available to reduce the number of training interval repetitions. Otherwise, a mixed-ADC implementation without ADC switching and extra training is preferred.

In Chapter 4, the performance of massive MIMO systems employing spatial one-bit $\Sigma\Delta$ quantization was investigated. Using an element-wise Bussgang approach, we derived an equivalent linear model in order to analytically characterize the spectral efficiency of a massive MIMO base station with a $\Sigma\Delta$ array, and we compared the results with the performance achieved by an array that employs standard one-bit quantization. Our results demonstrated that the spatial $\Sigma\Delta$ architecture can scale down the quantization noise power proportional to the square of the spatial oversampling rate. This can be interpreted as scaling down the quantization noise power proportional to the inverse square of the number of antennas at the BS for space-constrained arrays. This result gains more importance by noting that in standard one-bit quantization, the quantization noise power grows proportional to the inverse of

the spatial oversampling rate, or equivalently, proportional to the number of antennas at the BS in space-constrained arrays. Furthermore, it was shown how this capability allows the spatial $\Sigma\Delta$ architecture to bridge the SE gap between infinite resolution and standard one-bit quantized systems. For the ZF receiver, the spatial $\Sigma\Delta$ architecture can outperform standard one-bit quantization by about 50%, and achieve almost the same performance as an infinite resolution system for the MRC receiver. While these results were obtained by assuming the availability of perfect CSI at the BS, we also showed that the spatial $\Sigma\Delta$ architecture is able to alleviate the adverse impact of quantization noise in the presence of channel estimation error. We presented a new spatial one-bit FBB $\Sigma\Delta$ architecture for mitigating strong interference in massive MIMO systems with one-bit quantization. We showed that this simple architecture can effectively compensate for the vulnerability of one-bit ADCs against strong interference. The critical challenges in designing this architecture are to find the appropriate output levels for the one-bit quantizers and the values for the feedback weights. A recursive algorithm was proposed to specify the quantizers' output levels. The feedback weights were designed by adopting an algorithm used previously for a temporal FBB $\Sigma\Delta$ implementation. However, the behaviour of the feedback weights indicates that they amount to a spatial beamformer pointing in the direction(s) of the interference, and hence could be designed by a less complicated approach. Interesting directions for future work include studying the impact of angle estimation errors on the performance of the FBB $\Sigma\Delta$ architecture, or using the approach for combined quantization noise shaping and transmit beampattern design for the downlink with low-resolution digital-to-analog converters.

Finally, in Chapter 5, we analyzed the effect of mutual coupling on the performance of mixed-ADC and one-bit $\Sigma\Delta$ massive MIMO systems. For mixed-ADC architecture, we see that by equally spacing a small number of high-resolution ADCs over the array, one can dramatically reduce the performance gap between a system with all low-resolution and all high-resolution ADCs. For $\Sigma\Delta$ architecture, it was shown that this architecture is most suitable for array deployments with an aperture size constraint. While the performance of standard one-

bit quantization saturates as the number of antennas increases in a constrained-aperture array, the performance of the $\Sigma\Delta$ architecture tends to approach that of a system with high-resolution ADCs. This is due to the noise-shaping gain achieved by the $\Sigma\Delta$ architecture when the users are sectorized or the array is oversampled in space. It is worthwhile to note that the inevitable power loss due to mutual coupling can to some extent be alleviated using, for example, a matching network. This is a subject of future investigation.

Bibliography

[1] T. L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11):3590–3600, 2010.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang. An overview of massive mimo: Benefits and challenges. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):742–758, 2014.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive mimo for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, 2014.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta. Energy and spectral efficiency of very large multiuser mimo systems. *IEEE Transactions on Communications*, 61(4):1436–1449, 2013.

[5] H. Yang and T. L. Marzetta. Performance of conjugate and zero-forcing beamforming in large-scale antenna systems. *IEEE Journal on Selected Areas in Communications*, 31(2):172–179, 2013.

[6] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah. Massive mimo systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits. *IEEE Transactions on Information Theory*, 60(11):7112–7139, 2014.

[7] C. Mollén, E. G. Larsson, and T. Eriksson. Waveforms for the massive mimo downlink: Amplifier efficiency, distortion, and performance. *IEEE Transactions on Communications*, 64(12):5050–5063, 2016.

[8] C. Mollén, U. Gustavsson, T. Eriksson, and E. G. Larsson. Spatial characteristics of distortion radiated from antenna arrays with transceiver nonlinearities. *IEEE Transactions on Wireless Communications*, 17(10):6663–6679, 2018.

[9] C. Mollén, E. G. Larsson, U. Gustavsson, T. Eriksson, and R. W. Heath. Out-of-band radiation from large antenna arrays. *IEEE Communications Magazine*, 56(4):196–203, 2018.

[10] C. Mollén, U. Gustavsson, T. Eriksson, and E. G. Larsson. Impact of spatial filtering on distortion from low-noise amplifiers in massive mimo base stations. *IEEE Transactions on Communications*, 66(12):6050–6067, 2018.

[11] E. Björnson, M. Matthaiou, and M. Debbah. Massive mimo with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design. *IEEE Transactions on Wireless Communications*, 14(8):4353–4368, 2015.

[12] Q. Bai and J. A. Nossek. Energy efficiency maximization for 5g multi-antenna receivers. *Transactions on Emerging Telecommunications Technologies*, 26(1):3–14, 2015.

[13] W. B. Abbas, F. Gomez-Cuba, and M. Zorzi. Millimeter wave receiver efficiency: A comprehensive comparison of beamforming schemes with low resolution adcs. *IEEE Transactions on Wireless Communications*, 16(12):8131–8146, 2017.

[14] K. Roth, H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek. A comparison of hybrid beamforming and digital beamforming with low-resolution adcs for multiple users and imperfect csi. *IEEE Journal of Selected Topics in Signal Processing*, 12(3):484–498, 2018.

[15] 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*, 65(15):4075–4089, 2017.

[16] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath. Uplink performance of wideband massive mimo with one-bit adcs. *IEEE Transactions on Wireless Communications*, 16(1):87–100, 2017.

[17] 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*, 16(6):4038–4051, 2017.

[18] C. Studer and G. Durisi. Quantized massive mu-mimo-ofdm uplink. *IEEE Transactions on Communications*, 64(6):2387–2399, 2016.

[19] J. Mo and R. W. Heath. Capacity analysis of one-bit quantized mimo systems with transmitter channel state information. *IEEE Transactions on Signal Processing*, 63(20):5498–5512, 2015.

[20] M. Sarajli, L. Liu, and O. Edfors. When are low resolution adcs energy efficient in massive mimo? *IEEE Access*, 5:14837–14853, 2017.

[21] D. Verenzuela, E. Björnson, and M. Matthaiou. Hardware design and optimal adc resolution for uplink massive mimo systems. In *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5, 2016.

[22] L. Fan, S. Jin, C. Wen, and H. Zhang. Uplink achievable rate for massive mimo systems with low-resolution adc. *IEEE Communications Letters*, 19(12):2186–2189, 2015.

[23] J. Zhang, L. Dai, S. Sun, and Z. Wang. On the spectral efficiency of massive mimo systems with low-resolution adcs. *IEEE Communications Letters*, 20(5):842–845, 2016.

[24] N. Liang and W. Zhang. Mixed-adc massive mimo. *IEEE Journal on Selected Areas in Communications*, 34(4):983–997, 2016.

[25] N. Liang and W. Zhang. Mixed-adc massive mimo uplink in frequency-selective channels. *IEEE Transactions on Communications*, 64(11):4652–4666, 2016.

[26] W. Tan, S. Jin, C. Wen, and Y. Jing. Spectral efficiency of mixed-adc receivers for massive mimo systems. *IEEE Access*, 4:7841–7846, 2016.

[27] J. Zhang, L. Dai, Z. He, S. Jin, and X. Li. Performance analysis of mixed-adc massive mimo systems over rician fading channels. *IEEE Journal on Selected Areas in Communications*, 35(6):1327–1338, 2017.

[28] H. Pirzadeh and A. L. Swindlehurst. Spectral efficiency under energy constraint for mixed-adc mrc massive mimo. *IEEE Signal Processing Letters*, 24(12):1847–1851, 2017.

[29] T. Zhang, C. Wen, S. Jin, and T. Jiang. Mixed-adc massive mimo detectors: Performance analysis and design optimization. *IEEE Transactions on Wireless Communications*, 15(11):7738–7752, 2016.

[30] J. Liu, J. Xu, W. Xu, S. Jin, and X. Dong. Multiuser massive mimo relaying with mixed-adc receiver. *IEEE Signal Processing Letters*, 24(1):76–80, 2017.

[31] J. Park, S. Park, A. Yazdan, and R. W. Heath. Optimization of mixed-adc multi-antenna systems for cloud-ran deployments. *IEEE Transactions on Communications*, 65(9):3962–3975, 2017.

[32] A. B. Üçüncü and A. Ö. Yılmaz. Oversampling in one-bit quantized massive mimo systems and performance analysis. *IEEE Transactions on Wireless Communications*, 17(12):7952–7964, 2018.

[33] A. Gokceoglu, E. Björnson, E. G. Larsson, and M. Valkama. Spatio-temporal waveform design for multiuser massive mimo downlink with 1-bit receivers. *IEEE Journal of Selected Topics in Signal Processing*, 11(2):347–362, 2017.

[34] M. Schluter, M. Dorpinghaus, and G. P. Fettweis. Bounds on channel parameter estimation with 1-bit quantization and oversampling. In *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5, 2018.

[35] 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*, 42(12):773–784, 1995.

[36] 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*, 15(10):6556–6569, 2016.

[37] P. M. Aziz, H. V. Sorensen, and J. vn der Spiegel. An overview of sigma-delta converters. *IEEE Signal Processing Magazine*, 13(1):61–84, 1996.

[38] R. T. Baird and T. S. Fiez. Linearity enhancement of multibit /spl delta//spl sigma/ a/d and d/a converters using data weighted averaging. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 42(12):753–762, 1995.

[39] I. Galton. Delta-sigma data conversion in wireless transceivers. *IEEE Transactions on Microwave Theory and Techniques*, 50(1):302–315, 2002.

[40] D. P. Scholnik, J. O. Coleman, D. Bowling, and M. Neel. Spatio-temporal delta-sigma modulation for shared wideband transmit arrays. In *Proceedings of the 2004 IEEE Radar Conference (IEEE Cat. No.04CH37509)*, pages 85–90, 2004.

[41] J. D. Krieger, C. Yeang, and G. W. Wornell. Dense delta-sigma phased arrays. *IEEE Transactions on Antennas and Propagation*, 61(4):1825–1837, 2013.

[42] V. Venkateswaran and A. van der Veen. Multichannel sigma-delta adcs with integrated feedback beamformers to cancel interfering communication signals. *IEEE Transactions on Signal Processing*, 59(5):2211–2222, 2011.

[43] A. Nikoofard, J. Liang, M. Twieg, S. Handagala, A. Madanayake, L. Belostotski, and S. Mandal. Low-complexity n-port adcs using 2-d sigma-delta noise-shaping for n-element array receivers. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 301–304, 2017.

[44] 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 *2017 10th International Workshop on Multidimensional (nD) Systems (nDS)*, pages 1–6, 2017.

[45] R. M. Corey and A. C. Singer. Spatial sigma-delta signal acquisition for wideband beamforming arrays. In *WSA 2016; 20th International ITG Workshop on Smart Antennas*, pages 1–7, 2016.

[46] D. Barac and E. Lindqvist. Spatial sigma-delta modulation in a massive mimo cellular system. Master's thesis, Chalmers University of Technology, Chalmers University of Technology / Department of Computer Science and Engineering, 2016.

[47] S. Rao, A. L. Swindlehurst, and H. Pirzadeh. Massive mimo channel estimation with 1-bit spatial sigma-delta adcs. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4484–4488, 2019.

[48] M. Shao, W. K. Ma, Q. Li, and A. L. Swindlehurst. One-bit sigma-delta mimo precoding. *IEEE Journal of Selected Topics in Signal Processing*, 13(5):1046–1061, 2019.

[49] Julian J. Bussgang and Julian J. Bussgang. Crosscorrelation functions of amplitude-distorted gaussian signals. *Tech. Rep. 216, Research Lab. Electron*, 1952.

[50] Y. Li, C. Tao, A. Lee Swindlehurst, A. Mezghani, and L. Liu. Downlink achievable rate analysis in massive mimo systems with one-bit dacs. *IEEE Communications Letters*, 21(7):1669–1672, 2017.

[51] E. Björnson, E. G. Larsson, and M. Debbah. Massive mimo for maximal spectral efficiency: How many users and pilots should be allocated? *IEEE Transactions on Wireless Communications*, 15(2):1293–1308, 2016.

[52] G. Jacovitti and A. Neri. Estimation of the autocorrelation function of complex gaussian stationary processes by amplitude clipped signals. *IEEE Transactions on Information Theory*, 40(1):239–245, 1994.

[53] H. Pirzadeh and L. Swindlehurst. On the optimality of mixed-adc massive mimo with mrc detection. In *WSA 2017; 21th International ITG Workshop on Smart Antennas*, pages 1–6, 2017.

[54] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., USA, 1993.

[55] A. Garcia-Rodriguez, C. Masouros, and P. Rulikowski. Reduced switching connectivity for large scale antenna selection. *IEEE Transactions on Communications*, 65(5):2250–2263, 2017.

[56] Y. Gao, H. Vinck, and T. Kaiser. Massive mimo antenna selection: Switching architectures, capacity bounds, and optimal antenna selection algorithms. *IEEE Transactions on Signal Processing*, 66(5):1346–1360, 2018.

[57] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson. Multi-switch for antenna selection in massive mimo. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, 2015.

[58] A. Le Fevre and R. Flett. A 100 mb/s multi-lan crosspoint chip set for cable management. *IEEE Journal of Solid-State Circuits*, 32(7):1115–1121, 1997.

[59] Q. Shi and Y. Karasawa. Some applications of lauricella hypergeometric function f_a in performance analysis of wireless communications. *IEEE Communications Letters*, 16(5):581–584, 2012.

[60] H. A. David. *Order Statistics*. Wiley, USA, 1981.

[61] Saralees Nadarajah and Manisha Pal. Explicit expressions for moments of gamma order statistics. *Bulletin of the Brazilian Mathematical Society, New Series*, 39(1):45–60, 2008.

[62] L. Fan, D. Qiao, S. Jin, C. Wen, and M. Matthaiou. Optimal pilot length for uplink massive mimo systems with low-resolution adc. In *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5, 2016.

[63] D. Qiao, W. Tan, Y. Zhao, C. Wen, and S. Jin. Spectral efficiency for massive mimo zero-forcing receiver with low-resolution adc. In *2016 8th International Conference on Wireless Communications Signal Processing (WCSP)*, pages 1–6, 2016.

[64] H. Q. Ngo, M. Matthaiou, and E. G. Larsson. Massive mimo with optimal power and training duration allocation. *IEEE Wireless Communications Letters*, 3(6):605–608, 2014.

[65] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta. The multicell multiuser mimo uplink with very large antenna arrays and a finite-dimensional channel. *IEEE Transactions on Communications*, 61(6):2350–2361, 2013.

[66] J. Barrett and D. Lampard. An expansion for some second-order probability distributions and its application to noise problems. *IRE Transactions on Information Theory*, 1(1):10–15, 1955.

[67] R. C. Booton. Nonlinear control systems with random inputs. *IRE Transactions on Circuit Theory*, 1(1):9–18, 1954.

[68] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall Press, USA, 3rd edition, 2009.

[69] J. H. Van Vleck and D. Middleton. The spectrum of clipped noise. *Proceedings of the IEEE*, 54(1):2–19, 1966.

[70] C. Masouros and M. Matthaiou. Space-constrained massive mimo: Hitting the wall of favorable propagation. *IEEE Communications Letters*, 19(5):771–774, 2015.

[71] Q. Zhang, S. Jin, K. Wong, H. Zhu, and M. Matthaiou. Power scaling of uplink massive mimo systems with arbitrary-rank channel means. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):966–981, 2014.

[72] H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek. Space-constrained mixed-adc massive mimo. In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5, 2019.

[73] C. Masouros, M. Sellathurai, and T. Ratnarajah. Large-scale mimo transmitters in fixed physical spaces: The effect of transmit correlation and mutual coupling. *IEEE Transactions on Communications*, 61(7):2794–2804, 2013.

[74] S. Biswas, C. Masouros, and T. Ratnarajah. Performance analysis of large multiuser mimo systems with space-constrained 2-d antenna arrays. *IEEE Transactions on Wireless Communications*, 15(5):3492–3505, 2016.

[75] X. Ge, R. Zi, H. Wang, J. Zhang, and M. Jo. Multi-user massive mimo communication systems based on irregular antenna arrays. *IEEE Transactions on Wireless Communications*, 15(8):5287–5301, 2016.

[76] A. Garcia-Rodriguez and C. Masouros. Exploiting the increasing correlation of space constrained massive mimo for csi relaxation. *IEEE Transactions on Communications*, 64(4):1572–1587, 2016.

[77] J. Zhang, L. Dai, M. Matthaiou, C. Masouros, and S. Jin. On the spectral efficiency of space-constrained massive mimo with linear receivers. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6, 2016.

[78] H. Tataria, P. J. Smith, M. Matthaiou, and P. A. Dmochowski. Uplink analysis of large mu-mimo systems with space-constrained arrays in ricean fading. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–7, 2017.

[79] S. A. Schelkunoff. *Antennas: Theory and Practice*. John Wiley and Sons, USA, 1952.