

# UC Irvine

## UC Irvine Previously Published Works

### Title

Optimal Design of Energy and Spectral Efficiency Tradeoff in One-Bit Massive MIMO Systems

### Permalink

<https://escholarship.org/uc/item/8p4290h9>

### Authors

Li, Yongzhi  
Tao, Cheng  
Mezghani, Amine  
[et al.](#)

### Publication Date

2016-12-10

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Optimal Design of Energy and Spectral Efficiency Tradeoff in One-Bit Massive MIMO Systems

Yongzhi Li, Cheng Tao, *Member, IEEE*, Amine Mezghani, A. Lee Swindlehurst, *Fellow, IEEE*  
Gonzalo Seco-Granados, *Senior Member, IEEE*, and Liu Liu

**Abstract**—This paper considers a single-cell massive multiple-input multiple-output (MIMO) system equipped with a base station (BS) that uses one-bit quantization and investigates the energy efficiency (EE) and spectral efficiency (SE) trade-off. We first propose a new precoding scheme and downlink power allocation strategy that results in uplink-downlink SINR duality for one-bit MIMO systems. Taking into account the effect of the imperfect channel state information, we obtain approximate closed-form expressions for the uplink and downlink achievable rates under duality with maximum ratio combining/matched-filter and zero-forcing processing. We then focus on joint optimization of the competing SE and EE objectives over the number of users, pilot training duration and operating power, using the weighted product method to obtain the EE/SE Pareto boundary. Numerical results are presented to verify our analytical results and demonstrate the fundamental tradeoff between EE and SE for different parameter settings.

**Index Terms**—massive MIMO, one-bit quantization, resource allocation, spectral efficiency, energy efficiency.

## I. INTRODUCTION

In the past decade, there has been considerable work focused on improving wireless system spectral efficiency (SE) and throughput triggered by spectrum scarcity and the demand for higher rates for multimedia applications. Massive multiple-input multiple-output (MIMO) technology is considered to be an important component of 5th generation (5G) wireless communication systems, and has been shown to potentially achieve increases in SE by orders of magnitude over contemporary systems. The main idea of massive MIMO is based

Y. Li, C. Tao, and L. Liu are with the Institute of Broadband Wireless Mobile Communications, School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, 100044, China (email: liyongzhi@bjtu.edu.cn; chtao@bjtu.edu.cn; liuliu@bjtu.edu.cn).

G. Seco-Granados is with the Telecommunications and Systems Engineering Department, Universitat Autònoma de Barcelona, Barcelona 08193, Spain (e-mail: gonzalo.seco@uab.es).

A. Mezghani and A. L. Swindlehurst are with the Center for Pervasive Communications and Computing, University of California, Irvine, CA 92697 USA (e-mail: amezghan@uci.edu; swindle@uci.edu). A. L. Swindlehurst is also a Hans Fischer Senior Fellow of the Institute for Advanced Study at the Technical University of Munich.

The research was supported in part by the National 863 Project Granted No. 2014AA01A706, Beijing Natural Science Foundation project Grant No. 2016023, Fundamental Research Funds for the Central Universities under grant 2015JBM011, the NSFC project under grant No. 61471027, the Research Fund of National Mobile Communications Research Laboratory, Southeast University No. 2014D05, and Beijing Natural Science Foundation project under grant No.4152043. A. Swindlehurst was supported by the National Science Foundation under Grant ECCS-1547155, and by the Technische Universität München Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement No. 291763, and by the European Union under the Marie Curie COFUND Program.

on equipping the base stations (BSs) with many antenna elements, providing unprecedented spatial degrees of freedom for simultaneously serving multiple user terminals on the same time-frequency channel [1]–[4].

However, with a large number of antenna elements deployed at the BS, system cost and power consumption will be excessive if each antenna element and corresponding radio frequency (RF) chain is equipped with high-resolution and power-hungry analog-to-digital converter/digital-to-analog converters (ADC/DACs). In addition, as huge bandwidths and correspondingly high sampling rates will be required in next generation wireless systems, high-speed ADCs are either unavailable or too costly for practical implementation [5]. Therefore, the use of low-resolution ADC/DACs, especially those that operate with only one bit, has been suggested for massive MIMO systems [?], [6]–[9]. One-bit ADC/DACs significantly reduce power consumption and cost since they consist of a simple comparator and do not require automatic gain control or linear amplifiers, and hence greatly simplify the analog RF front end. It has been shown that MIMO capacity is not severely reduced by the coarse quantization at low signal-to-noise ratios (SNRs) [10]; the power penalty due to one-bit quantization is approximately equal to only  $\pi/2$  (1.96dB) in the low SNR region [11]. In addition, [8] and [9] showed that, compared with the conventional MIMO systems with perfect hardware implementation, the loss of SE in one-bit MIMO systems can be compensated for by with deploying around 2.2-2.5 times more antennas at the BS. Therefore the use of one-bit ADC/DACs can make massive MIMO systems more viable in practice.

While SE is the most common performance metric for designing wireless systems [12], [13], recent efforts to promote energy efficient or "green" communications have recently attracted considerable research interest [14], [15]. There has been a large amount of work devoted to maximizing the EE for massive MIMO in various scenarios, including single cell and multi-cell deployments [16], [17]. SE and EE are typically competing objectives; increasing one usually leads to degradations in the other. Since both are important factors in wireless systems, it is important to consider both when optimizing the system design. Multiple objective optimization approaches are useful for obtaining reasonable operating points on the Pareto performance boundary that balance the two criteria [18]–[20].

There has been limited prior work on the EE-SE tradeoff in various wireless communication systems [21]–[24]. The authors of [21] proposed a new paradigm for the EE-SE tradeoff in OFDMA cellular networks and showed that a

significant amount of bandwidth can be saved with a slight increase in energy consumption by using the proposed paradigm. In [22], the relationship between EE and SE for downlink multiuser distributed antennas systems (DAS) is investigated, and an algorithm to allocate the available power to balance EE and SE is also proposed. A full duplex relay system was considered in [23], and several power scaling schemes were proposed that achieve a good trade-off between SE and EE. The work of [24] investigated the SE-EE tradeoff using multi-objective optimization subject to a maximum total transmit power constraint for downlink massive MIMO systems for several different parameter settings.

In this paper, we consider a single-cell massive MIMO system with a BS equipped with one-bit ADC/DACs and we simultaneously address the EE-SE tradeoff for both uplink and downlink. Our goal is to provide insight on how the number of active terminals, the pilot training length and the system operating power affect the EE and SE for different linear processing schemes. Our specific contributions are summarized below.

- Since uplink-downlink duality is helpful when jointly designing the precoder and receiver, we prove the uplink-downlink SINR duality for one-bit MIMO systems operating at low SNR. On the basis of this duality, we propose a new precoding scheme and downlink power allocation strategy that achieve the same SINR in both the uplink and downlink. In addition, we show that by using our proposed power allocation strategy, the required range of power amplifier gains is small, especially for a large number of transmit antennas.
- We quantify the theoretical rate achievable in both the uplink and downlink using MRC/MF and ZF processing based on linear minimum mean-squared error (LMMSE) channel estimation. Closed-form expressions for the uplink and downlink achievable rates with MRC/MF and ZF are obtained as well, which enable us to simultaneously examine the uplink and downlink performance to evaluate the system SE and EE.
- Using a standard multiple objective optimization approach, we investigate the optimal number of active terminals, pilot training length and system operating power that achieve a good tradeoff between SE and EE, and we show the benefit of doing so via several numerical examples. We further show that the power penalty in one-bit systems can be compensated for by deploying 2-2.5 times more antennas for MRC/MF processing, while for ZF processing, more and more antennas are required as the SNR increases.

The rest of this paper is organized as follows. In the next section, we present the assumed system architecture and signal model for both the uplink and downlink. In Section III, we investigate channel estimation and data transmission for one-bit MIMO systems using the Bussgang decomposition. The uplink-downlink SINR duality in one-bit MIMO systems is proved in Section IV. In Section V, we provide approximate expressions for the uplink and downlink achievable rate. We focus on the energy and spectral efficiency trade-off in Section

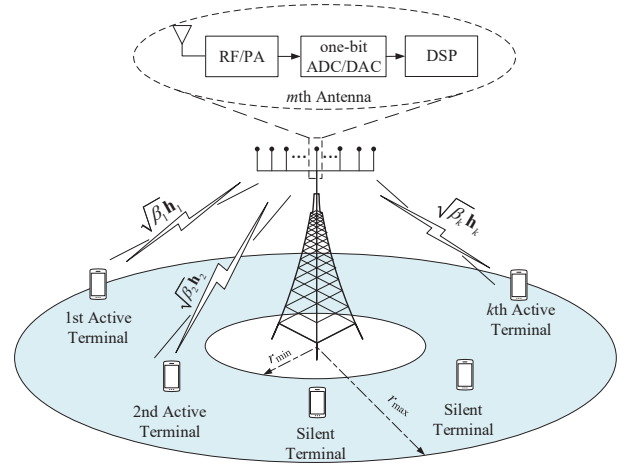


Fig. 1. System architecture of one-bit massive MIMO system.

VI, and investigate the effect of optimal resource allocation in improving both the SE and EE. Simulation results are presented in Section VI and we conclude the paper in Section VII.

*Notation:* The following notation is used throughout the paper. Bold uppercase (lowercase) letters denote matrices (vectors);  $(\cdot)^*$ ,  $(\cdot)^T$ , and  $(\cdot)^H$  denote complex conjugate, transpose, and Hermitian transpose operations, respectively;  $\|\cdot\|$  represents the 2-norm of a vector;  $\text{tr}(\cdot)$  represents the trace of a matrix;  $\text{diag}\{\mathbf{X}\}$  denotes a diagonal matrix containing only the diagonal entries of  $\mathbf{X}$ ;  $\otimes$  represents the Kronecker product;  $[\mathbf{X}]_{ij}$  denotes the  $(i, j)$ th entry of  $\mathbf{X}$ ;  $\mathbf{x} \sim \mathcal{CN}(\mathbf{a}, \mathbf{B})$  indicates that  $\mathbf{x}$  is a complex Gaussian vector with mean  $\mathbf{a}$  and covariance matrix  $\mathbf{B}$ ;  $\mathbb{E}\{\cdot\}$  and  $\text{Var}\{\cdot\}$  denote the expected value and variance of a random variable, respectively.

## II. SYSTEM MODEL

As depicted in Fig. 1, we consider a single-cell massive MIMO system, where each antenna of the base station (BS) is assumed to be equipped with one-bit ADC/DACs. For the uplink, the one-bit ADCs convert the real and imaginary part of the received signal from the analog domain, while for the downlink transmission, the one-bit DACs convert the precoded data symbols fed to the analog domain. We assume the BS is deployed with a large array of  $M$  antennas and communicates with  $K \gg 1$  out of  $K_{\max}$  single-antenna terminals at a time, where we refer to the  $K$  terminals as *active terminals* and to the remaining  $K_{\max} - K$  as *silent terminals*. The  $K$  active terminals are randomly selected in a round-robin fashion, such that the subset of active terminals changes over time, and hence the index of the terminal  $k \in \{1, \dots, K\}$  corresponds to different terminals at different times. The geographical position of the  $k$ th terminal is assumed to be a random variable drawn from a circular uniform distribution with maximum radius  $r_{\max}$  and minimum radius  $r_{\min}$ .

Let  $\mathbf{g}_k \in \mathbb{C}^{M \times 1}$  denote the channel vector between the  $k$ th active terminal and the BS, which includes the effects of independent fast fading, geometric attenuation and log-normal shadow fading:

$$\mathbf{g}_k = \sqrt{\beta_k} \mathbf{h}_k, \quad (1)$$

where  $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$  represents the fast fading channel of the  $k$ th active terminal and is assumed to be Gaussian distributed with zero mean and unit variance, and  $\beta_k$  is the large-scale fading coefficient that models both the geometric attenuation and shadow fading. The large scale fading coefficient of the  $k$ th active terminal is modeled as

$$\beta_k = \frac{\bar{d}}{(d_k/r_{\min})^\kappa}, \quad (2)$$

where  $\bar{d}$  is a non-logarithmic shadowing value,  $\kappa$  is the path-loss exponent and  $d_k$  represents the distance between the  $k$ th user and the BS. Due to the slow-varying nature of the long term channel statistics, we assume the large scale fading coefficients are known at the BS [1], [16].

We assume that the one-bit massive MIMO system operates in time-division duplex (TDD) mode and the channels of the  $K$  active terminals undergo block fading in which the channel state remains constant over a coherence interval of length  $T$  symbols. Assuming reciprocal uplink and downlink channels, the BS can process both the uplink and downlink signals using the uplink channel measurement.

#### A. Uplink Signal Model

During the uplink phase, we assume the  $K$  active terminals simultaneously transmit independent data symbols to the BS. Thus the uplink received signal at the BS is

$$\mathbf{y}_u = \sum_{k=1}^K \sqrt{p_k} \mathbf{g}_k s_k + \mathbf{n}_u, \quad (3)$$

where the subscript ‘u’ stands for ‘uplink’,  $p_k$  and  $s_k$  are respectively the transmit power and transmit data symbol of the  $k$ th active terminal. In addition we assume  $\mathbb{E}\{|s_k|^2\} = 1$  and  $\mathbb{E}\{s_i s_k^H\} = 0$  for  $i \neq k$ . The additive noise is assumed to be Gaussian:  $\mathbf{n}_u \sim \mathcal{CN}(0, \mathbf{I}) \in \mathbb{C}^{M \times 1}$ . The total uplink transmit power is then given by  $P_{\text{total}}^{(\text{UL})} = \sum_{k=1}^K p_k$ .

In most existing work, fixed and equal uplink transmit power is assumed for each active terminal. This may result in a large rate disparity between the nearest and furthest active terminals due to path loss. However, we assume a statistically-aware power control strategy in which the data symbols from the  $k$ th terminal have transmit power  $p_k = \beta_k^{-1} \rho_u$ , where  $\rho_u$  is a system design parameter we refer to as the system *operating power* [12]. The quantized signal obtained after the one-bit ADCs is represented as

$$\begin{aligned} \mathbf{r}_u &= \mathcal{Q}(\mathbf{y}_u) = \mathcal{Q}\left(\sum_{k=1}^K \sqrt{p_k} \mathbf{g}_k s_k + \mathbf{n}_u\right) \\ &= \mathcal{Q}\left(\sum_{k=1}^K \mathbf{g}_k^{\text{eff}} s_k + \mathbf{n}_u\right), \end{aligned} \quad (4)$$

where  $\mathbf{g}_k^{\text{eff}} = \sqrt{p_k} \mathbf{g}_k = \sqrt{\rho_u} \mathbf{h}_k$  represents the effective power-controlled channel of the  $k$ th active terminal. Since the large scale fading coefficients can be compensated for by adjusting the transmit power, we model the effective channel as  $\mathbf{g}_k^{\text{eff}} \sim \mathcal{CN}(0, \rho_u \mathbf{I})$ , which implies that a uniform terminal experience can be guaranteed since each active terminal has

the same channel statistics. The operator  $\mathcal{Q}(\cdot)$  represents the one-bit quantization, which is applied separately to the real and imaginary parts as  $\mathcal{Q}(\cdot) = \frac{1}{\sqrt{2}}(\text{sign}(\Re(\cdot)) + j\text{sign}(\Im(\cdot)))$ . Thus, the output set of the one-bit quantization is equivalent to the quadrature phase-shift keying (QPSK) constellation points  $\Omega = \frac{1}{\sqrt{2}}\{1 + j, 1 - j, -1 + j, -1 - j\}$ .

#### B. Downlink Signal Model

For the downlink, we assume that the data symbol  $x_k$  for the  $k$ -th active terminal is multiplied by a precoder  $\mathbf{t}_k$ , and then the real and imaginary parts are quantized by a one-bit DAC:

$$\mathbf{r}_d = \mathcal{Q}(\mathbf{y}_d) = \mathcal{Q}\left(\sum_{k=1}^K \mathbf{t}_k x_k\right), \quad (5)$$

where the subscript ‘d’ stands for ‘downlink’, the data symbols  $x_k$  are zero mean with unit variance and satisfy  $\mathbb{E}\{|x_k|^2\} = 1$  and  $\mathbb{E}\{x_i x_k^H\} = 0$  for  $i \neq k$ . Again, each element of  $\mathbf{r}_d$  lies in the set  $\Omega$  due to the one-bit quantization. As we assume that the system operates in TDD mode, the received signal at the  $k$ th active terminal is given by

$$\begin{aligned} u_{d,k} &= \mathbf{g}_k^T \mathbf{Q} \mathbf{r}_d + n_{d,k} \\ &= \mathbf{g}_k^T \mathbf{Q} \mathcal{Q}\left(\sum_{k=1}^K \mathbf{t}_k x_k\right) + n_{d,k}, \end{aligned} \quad (6)$$

where  $n_{d,k} \sim \mathcal{CN}(0, 1)$  is additive white Gaussian noise at the  $k$ th active terminal,  $\mathbf{Q}$  is a diagonal matrix that represents the power allocated to each BS antenna, and hence the total transmit power for the downlink is  $P_{\text{total}}^{(\text{DL})} = \text{tr}(\mathbf{Q} \mathbf{Q}^H)$ .

Note that contrary to conventional MIMO systems where the transmit power allocation can be considered jointly with the precoding vector, the power allocated to each BS antenna in a one-bit MIMO system should be considered separately since the power of each element of the output signal  $\mathbf{r}_d$  is normalized due to the one-bit quantization. We later show how to design the precoding vector  $\mathbf{t}_k$  and the transmit power matrix  $\mathbf{Q}$  to achieve the same achievable rate performance in the downlink as in the uplink.

### III. CHANNEL ESTIMATION AND DATA TRANSMISSION FOR ONE-BIT MIMO SYSTEMS

We assume each coherence interval  $T$  is divided into three parts: the first  $\tau = \tau_0 K$  symbols reserved for pilot training, and then the remaining  $T - \tau$  symbols are split between the uplink and downlink data transmission phases. The factor  $\tau_0$  represents the number of pilots relative to the number of active terminals. In conventional MIMO systems, the value that maximizes SE is  $\tau_0 = 1$ , while for one-bit systems SE is in general maximized for  $\tau_0 > 1$  [8]. We denote  $\gamma \in (0, 1)$  as the fraction of the coherence interval allocated for uplink transmission and  $1 - \gamma$  the fraction for downlink transmission.

### A. Channel Estimation

We assume the CSI is estimated based on pilot sequences of length  $\tau$  symbols:

$$\mathbf{Y}_t = \sum_{k=1}^K \sqrt{p_k} \mathbf{g}_k \phi_k^T + \mathbf{N}_t = \sum_{k=1}^K \mathbf{g}_k^{\text{eff}} \phi_k^T + \mathbf{N}_t, \quad (7)$$

where the subscript ‘t’ stands for ‘training’,  $\mathbf{Y}_t \in \mathbb{C}^{M \times \tau}$  is the received training signal matrix,  $p_k$  and  $\phi_k \in \mathbb{C}^{\tau \times 1}$  are the pilot transmit power and the pilot sequence of the  $k$ th user, respectively. In this paper, all pilot sequences are assumed to be mutual orthogonal, i.e.,  $\phi_k^T \phi_k^* = \tau$  and  $\phi_i^T \phi_k^* = 0$  for  $i \neq k$ .

To match the matrix form of (7) to the vector form of (4), we vectorize the received signal as

$$\mathbf{y}_t = \text{vec}(\mathbf{Y}_t) = \sum_{k=1}^K \bar{\phi}_k \mathbf{g}_k^{\text{eff}} + \mathbf{n}_t, \quad (8)$$

where  $\bar{\phi}_k = (\phi_k \otimes \mathbf{I}_M)$  and  $\mathbf{n}_t = \text{vec}(\mathbf{N}_t)$ . After one-bit ADCs, the quantized signal is expressed as

$$\mathbf{r}_t = \mathcal{Q}(\mathbf{y}_t) = \mathcal{Q} \left( \sum_{k=1}^K \bar{\phi}_k \mathbf{g}_k^{\text{eff}} + \mathbf{n}_t \right), \quad (9)$$

where each element of  $\mathbf{r}_t$  takes values from the set  $\Omega$ . We would like to remark that the output of (9),  $\mathbf{r}_t$ , is independent of any real-valued scaling factor applied to  $\mathbf{y}_t$ . Using the Bussgang decomposition [25], we can reformulate the nonlinear quantization using a statistically equivalent linear operator that will simplify the channel estimator and the resulting analysis:

$$\mathbf{r}_t = \mathcal{Q}(\mathbf{y}_t) = \mathbf{A}_t \mathbf{y}_t + \boldsymbol{\eta}_t = \sum_{k=1}^K \tilde{\phi}_k \mathbf{g}_k^{\text{eff}} + \tilde{\mathbf{n}}_t, \quad (10)$$

where  $\boldsymbol{\eta}_t$  is the quantization noise,  $\tilde{\phi}_k = \mathbf{A}_t \bar{\phi}_k$ ,  $\tilde{\mathbf{n}}_t = \mathbf{A}_t \mathbf{n}_t + \boldsymbol{\eta}_t$  is the total effective noise, and  $\mathbf{A}_t$  is chosen to make  $\boldsymbol{\eta}_t$  uncorrelated with  $\mathbf{y}_t$  [25], or equivalently, to minimize the power of the equivalent quantization noise. If the received signal  $\mathbf{y}_t$  is Gaussian, then  $\mathbf{A}_t$  is given by [8]

$$\begin{aligned} \mathbf{A}_t &= \sqrt{\frac{2}{\pi}} \text{diag}(\mathbf{C}_{\mathbf{y}_t})^{-\frac{1}{2}} \\ &= \sqrt{\frac{2}{\pi}} \text{diag} \left( \sum_{k=1}^K \rho_u (\phi_k \phi_k^H \otimes \mathbf{I}_M) + \mathbf{I}_{M\tau} \right)^{-\frac{1}{2}}, \end{aligned} \quad (11)$$

where  $\mathbf{C}_{\mathbf{y}_t}$  denotes the covariance matrix of  $\mathbf{y}_t$ .

We can see from (11) that  $\mathbf{A}_t$  is related to the diagonal terms of  $\mathbf{C}_{\mathbf{y}_t}$  and therefore, to the pilot sequences. In order to obtain a simple expression for  $\mathbf{A}_t$ , we consider pilot sequences composed of submatrices of the discrete Fourier transform (DFT) operator [26]. The benefits of using DFT pilot sequences are: i) all elements of the pilot sequences have the same magnitude, which simplifies peak transmit power constraints, and ii) the diagonal terms of  $\phi_k \phi_k^H$  are always equal to 1, which results in the following simple expression for  $\mathbf{A}_t$ :

$$\mathbf{A}_t = \sqrt{\frac{2}{\pi} \frac{1}{K \rho_u + 1}} \mathbf{I} \triangleq \alpha \mathbf{I}. \quad (12)$$

Since the one-bit quantization always results in a signal  $\mathbf{r}_t$  with unit variance, the individual transmit power  $p_k$  of each user is difficult to estimate [27]. Therefore we estimate the effective channels  $\mathbf{g}_k^{\text{eff}}$  including the uplink power-control instead of the true channel  $\mathbf{g}_k$ . In this way, we can estimate the channel without taking the transmit power  $p_k$  into account. According to [28] and the fact that  $\eta_p$  is uncorrelated with the channel  $\mathbf{g}_k^{\text{eff}}$  (see [8] for a detailed proof), the LMMSE channel estimate of  $\mathbf{g}_k^{\text{eff}}$  can be expressed as

$$\hat{\mathbf{g}}_k^{\text{eff}} = \rho_u \tilde{\phi}_k^H \mathbf{C}_{\mathbf{r}_t}^{-1} \mathbf{r}_t, \quad (13)$$

where  $\mathbf{C}_{\mathbf{r}_t}$  is the auto-correlation matrix of  $\mathbf{r}_t$ . It has been shown in [29] that for one-bit ADCs, the arcsine law can be used to obtain  $\mathbf{C}_{\mathbf{r}_t}$ :

$$\begin{aligned} \mathbf{C}_{\mathbf{r}_t} &= \frac{2}{\pi} \left( \arcsin \left( \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-\frac{1}{2}} \Re(\mathbf{C}_{\mathbf{y}_t}) \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-\frac{1}{2}} \right) \right. \\ &\quad \left. + j \arcsin \left( \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-\frac{1}{2}} \Im(\mathbf{C}_{\mathbf{y}_t}) \boldsymbol{\Sigma}_{\mathbf{y}_t}^{-\frac{1}{2}} \right) \right), \end{aligned} \quad (14)$$

where we define  $\boldsymbol{\Sigma}_{\mathbf{y}_t} = \text{diag}(\mathbf{C}_{\mathbf{y}_t})$ . The covariance matrix of the estimation error  $\boldsymbol{\varepsilon}_k^{\text{eff}} = \mathbf{g}_k^{\text{eff}} - \hat{\mathbf{g}}_k^{\text{eff}}$  is then given by

$$\begin{aligned} \mathbf{C}_{\boldsymbol{\varepsilon}_k^{\text{eff}}} &= \mathbb{E} \left\{ (\mathbf{g}_k^{\text{eff}} - \hat{\mathbf{g}}_k^{\text{eff}}) (\mathbf{g}_k^{\text{eff}} - \hat{\mathbf{g}}_k^{\text{eff}})^H \right\} \\ &= \rho_u \mathbf{I}_M - \rho_u^2 \tilde{\phi}_k^H \mathbf{C}_{\mathbf{r}_t}^{-1} \tilde{\phi}_k \end{aligned} \quad (15)$$

and the mean-squared error (MSE) is  $\text{MSE}_k = \text{tr}(\mathbf{C}_{\boldsymbol{\varepsilon}_k^{\text{eff}}})$ .

Note that  $\mathbf{C}_{\mathbf{y}_t}$  is a circulant matrix, and hence  $\mathbf{C}_{\mathbf{r}_t}$  is also circulant since the arcsine operation is performed element-wise. Note also that the inverse of a circulant matrix is also circulant, and that circulant matrices can be diagonalized by a DFT matrix. Therefore, another advantage of using the DFT matrix as pilot sequences is that, by using the channel estimator in (13), the elements of the channel estimate are mutual uncorrelated. Thus, the covariance matrix of the estimation error in (15) is diagonal.

### B. Data Transmission

During the data transmission phase, the BS considers the channel estimates as the true channel and employs a linear receiver to decode the uplink signals from the  $K$  active terminals, and a linear precoder to broadcast the downlink signal to the  $K$  terminals.

1) *Linear Receiver*: During the uplink, the  $K$  users simultaneously transmit their data symbols to the BS. We again employ the Bussgang decomposition and reformulate the nonlinear quantizer  $\mathcal{Q}(\cdot)$  as a linear function, and the quantized signal at the BS can be expressed as

$$\begin{aligned} \mathbf{r}_u &= \mathcal{Q}(\mathbf{y}_u) = \mathcal{Q} \left( \sum_{k=1}^K \sqrt{p_k} \mathbf{g}_k s_k + \mathbf{n}_u \right) \\ &= \mathbf{A}_u \sum_{k=1}^K \mathbf{g}_k^{\text{eff}} s_k + \mathbf{A}_u \mathbf{n}_u + \boldsymbol{\eta}_u, \end{aligned} \quad (16)$$

where the same definitions as in the previous sections apply, but replacing the subscript ‘t’ with ‘u’. Again, according to the Bussgang decomposition, the matrix  $\mathbf{A}_u$  for a Gaussian

input is given by

$$\begin{aligned} \mathbf{A}_u &= \sqrt{\frac{2}{\pi}} \text{diag}(\mathbf{C}_{y_u})^{-\frac{1}{2}} \\ &= \sqrt{\frac{2}{\pi}} \text{diag} \left( \sum_{k=1}^K \mathbf{g}_k^{\text{eff}} (\mathbf{g}_k^{\text{eff}})^H + \mathbf{I}_M \right)^{-\frac{1}{2}} \\ &\cong \sqrt{\frac{2}{\pi}} \sqrt{\frac{1}{K\rho_u + 1}} \mathbf{I}_M = \alpha \mathbf{I}_M, \end{aligned} \quad (17)$$

where the approximation is reasonable since we assume  $K \gg 1$  and i.i.d. effective channel coefficients.

Note that unlike (11), where the pilot sequences are fixed and the expectation operation is taken over the channel vectors, the expectation in (17) is taken over the data symbols for each individual channel realization. As such, the effect that the quantization noise sometimes appears to coherently combine with the data symbols as observed in [9] can be avoided.

With a linear receiver, the quantized signal  $\mathbf{r}_u$  is separated into  $K$  streams via multiplication by the matrix  $\mathbf{W}^T$  as follows:

$$\begin{aligned} \hat{\mathbf{s}} &= \mathbf{W}^T \mathbf{r}_u \\ &= \mathbf{W}^T \mathbf{A}_u \sum_{k=1}^K \mathbf{g}_k^{\text{eff}} s_k + \mathbf{W}^T \mathbf{A}_u \mathbf{n}_u + \mathbf{W}^T \boldsymbol{\eta}_u. \end{aligned} \quad (18)$$

Then the  $k$ th stream is used to decode the signal transmitted from the  $k$ th active terminal, which can be expressed as

$$\begin{aligned} \hat{s}_k &= \mathbf{w}_k^T \mathbf{A}_u \mathbf{g}_k^{\text{eff}} s_k + \mathbf{w}_k^T \sum_{i \neq k}^K \mathbf{A}_u \mathbf{g}_i^{\text{eff}} s_i \\ &\quad + \mathbf{w}_k^T \mathbf{A}_u \mathbf{n}_u + \mathbf{w}_k^T \boldsymbol{\eta}_u, \end{aligned} \quad (19)$$

where  $\mathbf{w}_k$  is the  $k$ th column of  $\mathbf{W}$ . The last three terms in (19) respectively correspond to user interference, AWGN noise and quantization noise.

2) *Linear Precoding*: Similar to the uplink, in the downlink we can reformulate the precoded signal model of (5) using the Bussgang decomposition:

$$\mathbf{r}_d = \mathcal{Q} \left( \sum_{k=1}^K \mathbf{t}_k x_k \right) = \mathbf{A}_d \sum_{k=1}^K \mathbf{t}_k x_k + \boldsymbol{\eta}_d, \quad (20)$$

where the same definitions as in the previous sections apply, but replacing the subscript ‘t’ with ‘d’. The matrix  $\mathbf{A}_d$  for a Gaussian input is given by

$$\mathbf{A}_d = \sqrt{\frac{2}{\pi}} \text{diag}(\mathbf{C}_{y_d})^{-\frac{1}{2}} = \sqrt{\frac{2}{\pi}} \text{diag} \left( \sum_{k=1}^K \mathbf{t}_k \mathbf{t}_k^H \right)^{-\frac{1}{2}}. \quad (21)$$

Substituting (20) into (6), the received signal at the  $k$ th terminal can be expressed as

$$\begin{aligned} u_{d,k} &= \mathbf{g}_k^T \mathbf{Q} \mathcal{Q} \left( \sum_{k=1}^K \mathbf{t}_k x_k \right) + n_{d,k} \\ &= \mathbf{g}_k^T \mathbf{Q} \mathbf{A}_d \mathbf{t}_k x_k + \mathbf{g}_k^T \mathbf{Q} \mathbf{A}_d \sum_{i \neq k}^K \mathbf{t}_i x_i + \mathbf{g}_k^T \mathbf{Q} \boldsymbol{\eta}_d + n_{d,k}. \end{aligned} \quad (22)$$

The last three terms in (22) respectively correspond to multiuser interference, quantization noise and AWGN noise.

#### IV. UPLINK-DOWNLINK SINR DUALITY IN ONE-BIT MIMO SYSTEMS

The receiver and precoder can be designed in a variety of ways to satisfy specific uplink and downlink SINR requirements. In general, it is more difficult to optimize the downlink precoder than the uplink receiver, and the uplink-downlink SINR duality theorem is a key tool which substantially simplifies the problem of joint precoder and receiver design. The duality between the downlink (broadcast) and the uplink (multiple access) channels in unquantized MIMO systems has been proved in previous work [30]. As with conventional MIMO systems, establishing SINR duality for the one-bit case is useful because, for example, it allows one to transform the more difficult problem of downlink precoder design and power allocation to the easier dual uplink problem. Thus, in the following, we establish the conditions for which the uplink and downlink duality holds in one-bit systems. As we will see later, this uplink-downlink duality allows us to derive an EE-SE tradeoff that is valid for both links.

According to the signal model of (19) and (22), the uplink and downlink SINR of the  $k$ th active user can be expressed as (23) and (24) shown on the top of next page, where  $\mathbf{C}_{\eta_p}$  denotes the covariance matrix of the quantizer noise  $\boldsymbol{\eta}_p$

$$\mathbf{C}_{\eta_p} = \mathbf{C}_{r_p} - \mathbf{A}_p \mathbf{C}_{y_p} \mathbf{A}_p^H, \quad (25)$$

where the subscript ‘p’ represents either the training, uplink or downlink phase,  $p \in \{t, u, d\}$ . According to the arcsine law, we can rewrite  $\mathbf{C}_{\eta_p}$  as

$$\mathbf{C}_{\eta_p} = \frac{2}{\pi} (\arcsin(\mathbf{X}_p) + j \arcsin(\mathbf{Y}_p)) - \frac{2}{\pi} (\mathbf{X}_p + j \mathbf{Y}_p), \quad (26)$$

where

$$\mathbf{X}_p = \text{diag}(\mathbf{C}_{y_p})^{-\frac{1}{2}} \Re(\mathbf{C}_{y_p}) \text{diag}(\mathbf{C}_{y_p})^{-\frac{1}{2}}, \quad (27)$$

$$\mathbf{Y}_p = \text{diag}(\mathbf{C}_{y_p})^{-\frac{1}{2}} \Im(\mathbf{C}_{y_p}) \text{diag}(\mathbf{C}_{y_p})^{-\frac{1}{2}}. \quad (28)$$

Due to the nonlinear arcsine function, it is complicated to prove the SINR duality directly. However, we note the approximation that

$$\frac{2}{\pi} \arcsin(a) \cong \begin{cases} 1, & a = 1 \\ 2a/\pi, & a < 1. \end{cases} \quad (29)$$

Thus if the non-diagonal elements of  $\mathbf{X}_p$  and  $\mathbf{Y}_p$  are much smaller than 1, we can approximate (26) as

$$\mathbf{C}_{\eta_p} \cong (1 - 2/\pi) \mathbf{I}, \quad (30)$$

which implies that the quantizer noise is approximately uncorrelated. By using this approximation, we arrive at Theorem 1 below.

*Theorem 1*: Assume a one-bit MIMO system with the uplink linear receiver  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_K]$ , user power allocation vector  $\mathbf{p} = [p_1, \dots, p_K]^T$  with  $\|\mathbf{p}\|_1 = P_{\text{total}}^{(\text{UL})}$  and uncorrelated quantizer noise. If the  $k$ th user achieves  $\text{SINR}_k^{(\text{UL})} = \gamma_k$  in the uplink, then the same  $\text{SINR}_k^{(\text{DL})} = \gamma_k$  can also be achieved using equal total power in the downlink if the downlink precoders and BS antenna power allocation is chosen

$$\text{SINR}_k^{(\text{UL})} = \frac{p_k |\mathbf{w}_k^T \mathbf{A}_u \mathbf{g}_k|^2}{\sum_{i \neq k}^K p_i |\mathbf{w}_k^T \mathbf{A}_u \mathbf{g}_i|^2 + \|\mathbf{w}_k^T \mathbf{A}_u\|_2^2 + \mathbf{w}_k^T \mathbf{C}_{\eta_u} \mathbf{w}_k^*}, \quad (23)$$

$$\text{SINR}_k^{(\text{DL})} = \frac{|\mathbf{g}_k^T \mathbf{Q} \mathbf{A}_d \mathbf{t}_k|^2}{\sum_{i \neq k}^K |\mathbf{g}_k^T \mathbf{Q} \mathbf{A}_d \mathbf{t}_i|^2 + \mathbf{g}_k^T \mathbf{Q} \mathbf{C}_{\eta_d} \mathbf{Q}^H \mathbf{g}_k^* + 1}. \quad (24)$$

as follows:

$$\mathbf{t}_i = \sqrt{\frac{q_i}{\|\mathbf{A}_u \mathbf{w}_i\|_2^2}} \mathbf{A}_u \mathbf{w}_i \triangleq \sqrt{q_i} \hat{\mathbf{t}}_i \quad (31)$$

$$\mathbf{Q} = \text{diag} \left( \sum_{i=1}^K \mathbf{t}_i \mathbf{t}_i^H \right)^{\frac{1}{2}}, \quad (32)$$

where  $\mathbf{A}_u$  is defined in (17) and  $q_i$  is the  $i$ th element of vector  $\mathbf{q}$  defined by

$$\mathbf{q} = \frac{\pi}{2} (\mathbf{I} - \mathbf{D} \Psi)^{-1} \mathbf{D} \mathbf{1}_K, \quad (33)$$

where  $\|\mathbf{q}\|_1 = P_{\text{total}}^{(\text{DL})} = P_{\text{total}}^{(\text{UL})}$ ,  $\mathbf{D}$  is a  $K \times K$  diagonal matrix whose  $k$ th diagonal element is

$$[\mathbf{D}]_{k,k} = \gamma_k |\mathbf{g}_k^T \hat{\mathbf{t}}_k|^{-2}, \quad (34)$$

and  $\Psi$  is a  $K \times K$  matrix, whose  $(k, i)$ th element is

$$[\Psi]_{k,i} = \begin{cases} |\mathbf{g}_k^T \hat{\mathbf{t}}_i|^2 + \left(\frac{\pi}{2} - 1\right) \hat{\mathbf{t}}_i^H \text{diag}(\mathbf{g}_k^* \mathbf{g}_k^T) \hat{\mathbf{t}}_i & i \neq k \\ \left(\frac{\pi}{2} - 1\right) \hat{\mathbf{t}}_i^H \text{diag}(\mathbf{g}_k^* \mathbf{g}_k^T) \hat{\mathbf{t}}_i & i = k \end{cases} \quad (35)$$

*Proof:* See Appendix A. ■

*Corollary 1:* Assume a one-bit MIMO system with the downlink linear precoder  $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_K]$ , BS antenna power allocation  $\mathbf{Q} = \text{diag} \left( \sum_{i=1}^K \mathbf{t}_i \mathbf{t}_i^H \right)^{\frac{1}{2}}$  with  $\text{tr}(\mathbf{Q} \mathbf{Q}^H) = P_{\text{total}}^{(\text{DL})}$  and uncorrelated quantizer noise. If the  $k$ th user achieves the  $\text{SINR}_k^{(\text{DL})} = \gamma_k$  in the downlink, then the same  $\text{SINR}_k^{(\text{UL})} = \gamma_k$  can also be achieved using equal total power in the uplink if the uplink receivers and user power allocation vector is chosen as

$$\mathbf{w}_i = \mathbf{A}_u^{-1} \mathbf{t}_i / \|\mathbf{t}_i\|_2 \quad (36)$$

$$\mathbf{p} = \frac{\pi}{2} (\mathbf{I} - \mathbf{D} \Psi^T)^{-1} \mathbf{D} \mathbf{1}_K, \quad (37)$$

where  $\|\mathbf{p}\|_1 = P_{\text{total}}^{(\text{UL})} = P_{\text{total}}^{(\text{DL})}$ . The matrices  $\mathbf{A}_u$ ,  $\mathbf{D}$  and  $\Psi$  are defined in (17), (34) and (35), respectively.

Note that the uplink-downlink SINR duality holds for arbitrary user power allocation and linear receivers, and the key point is that the quantizer noise has to be uncorrelated. Since it is expected that the number of users is large and the SNR per-antenna is low due to the availability of a large array gain in massive MIMO systems, the assumption of uncorrelated quantizer noise using the approximation in (30) is quite reasonable. Therefore, we conclude that uplink-downlink SINR duality holds in massive MIMO systems operating at low SNR.

## V. ACHIEVABLE RATE APPROXIMATION IN ONE-BIT MASSIVE MIMO SYSTEMS

In one-bit MIMO systems, the quantizer noise for the both uplink and downlink is not Gaussian due to the one-bit quantization. However the entropy of the uncorrelated effective noise is upper-bounded by the entropy of Gaussian noise, and this minimizes the mutual information of the input and output [31]. Therefore, a lower bound for the achievable rate can be found by modeling the quantizer noise  $\eta_u$  and  $\eta_d$  as Gaussian with the same covariance matrix given in (25). Thus the ergodic achievable rate for both the uplink and downlink in one-bit MIMO systems is lower bounded by

$$\tilde{R}_k^{(\text{Link})} = \mathbb{E} \left\{ \log_2 \left( 1 + \text{SINR}_k^{(\text{Link})} \right) \right\}, \quad (38)$$

where  $\text{Link} \in \{\text{UL}, \text{DL}\}$ , and  $\text{SINR}_k^{(\text{UL})}$  and  $\text{SINR}_k^{(\text{DL})}$  are given in (23) and (24), respectively [16]. In this section we derive closed-form expressions for the achievable rate in the low SNR region.

The covariance matrix of the estimated effective channel  $\hat{\mathbf{g}}_k^{\text{eff}}$  in (13) is given by

$$\mathbf{C}_{\hat{\mathbf{g}}_k^{\text{eff}}} = \rho_u^2 \tilde{\Phi}_k^H \mathbf{C}_{\mathbf{r}_t} \tilde{\Phi}_k. \quad (39)$$

We see that the expression above involves the covariance matrix  $\mathbf{C}_{\mathbf{r}_t}$ , which is complicated due to the arcsine law of (14). However, using the approximation of (30) for low SNR and using the linear form of  $\mathbf{r}_t$  given in (10), the covariance matrix of  $\mathbf{C}_{\eta_t}$  can be approximated as

$$\begin{aligned} \mathbf{C}_{\mathbf{r}_t} &= \sum_{k=1}^K \rho_u \tilde{\Phi}_k \tilde{\Phi}_k^H + \mathbf{A}_t \mathbf{A}_t^H + \mathbf{C}_{\eta_t \eta_t} \\ &\cong \sum_{k=1}^K \rho_u \tilde{\Phi}_k \tilde{\Phi}_k^H + \left( \alpha^2 + 1 - \frac{2}{\pi} \right) \mathbf{I}. \end{aligned} \quad (40)$$

Note that although the elements of the channel estimate (13) are not Gaussian distributed due to the one-bit quantization, we can approximate them as Gaussian according to Cramér's central limit theorem [32] since each element of  $\hat{\mathbf{g}}_k^{\text{eff}}$  can be expressed as the summation of a large number of random variables. Therefore, in the sequel each element of the channel estimate  $\hat{\mathbf{g}}_k^{\text{eff}}$  is approximated as Gaussian with covariance matrix

$$\begin{aligned} \mathbf{C}_{\hat{\mathbf{g}}_k^{\text{eff}}} &\cong \rho_u^2 \tilde{\Phi}_k^H \left( \sum_{k=1}^K \rho_u \tilde{\Phi}_k \tilde{\Phi}_k^H + \left( \alpha^2 + 1 - \frac{2}{\pi} \right) \mathbf{I} \right)^{-1} \tilde{\Phi}_k \\ &= \frac{\alpha^2 \tau \rho_u^2}{\alpha^2 \tau \rho_u + \alpha^2 + 1 - 2/\pi} \mathbf{I} = \sigma^2 \mathbf{I}, \end{aligned} \quad (41)$$

which implies that, at low SNR, each element of the channel estimate  $\hat{\mathbf{g}}_k^{\text{eff}}$  is approximated as independent and identically

Gaussian distributed (i.i.d) with zero mean and variance  $\sigma^2$ .

### A. Uplink Transmission

In this section we focus on deriving a closed-form expression for the uplink achievable rate. Since there is no efficient way to directly calculate the achievable rate in (38), we borrow Lemma 1 from [33] to derive an approximate closed-form expression. For completeness, we provide it below.

*Lemma 1:* If  $X = \sum_{i=1}^M X_i$  and  $Y = \sum_{i=j}^M Y_j$  are both sums of nonnegative random variables  $X_i$  and  $Y_j$ , then the following approximation holds

$$\mathbb{E} \left\{ \log_2 \left( 1 + \frac{X}{Y} \right) \right\} \cong \log_2 \left( 1 + \frac{\mathbb{E}\{X\}}{\mathbb{E}\{Y\}} \right). \quad (42)$$

*Proof:* See [33]. ■

Note that the approximation in Lemma 1 does not require the random variables  $X$  and  $Y$  to be independent and becomes more accurate as  $M$  increases.

According to Lemma 1 and substituting  $\mathbf{A}_u = \alpha \mathbf{I}$ , we can approximate the uplink ergodic achievable rate  $\tilde{R}_k^{(\text{UL})}$  by

$$R_k^{(\text{UL})} = \log_2 \left( 1 + \frac{\alpha^2 \mathbb{E} \left\{ |\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}}|^2 \right\}}{\sum_{i \neq k}^K \alpha^2 \mathbb{E} \left\{ |\mathbf{w}_k^T \mathbf{g}_i^{\text{eff}}|^2 \right\} + \text{AQN}_k} \right), \quad (43)$$

where we define

$$\text{AQN}_k = \alpha^2 \mathbb{E} \left\{ \|\mathbf{w}_k\|_2^2 \right\} + \mathbb{E} \left\{ \mathbf{w}_k^T \mathbf{C}_{\eta_u} \mathbf{w}_k^* \right\}, \quad (44)$$

and the expectation operation is taken with respect to the channel realizations. In this paper, we will consider the performance of the standard MRC and ZF receivers, defined by

$$\mathbf{W}_{\text{MRC}}^T = \hat{\mathbf{G}}_{\text{eff}}^H \quad (45)$$

$$\mathbf{W}_{\text{ZF}}^T = \left( \hat{\mathbf{G}}_{\text{eff}}^H \hat{\mathbf{G}}_{\text{eff}}^H \right)^{-1} \hat{\mathbf{G}}_{\text{eff}}^H \quad (46)$$

respectively, where  $\hat{\mathbf{G}}_{\text{eff}} = [\hat{\mathbf{g}}_1^{\text{eff}}, \dots, \hat{\mathbf{g}}_K^{\text{eff}}]$ .

*Theorem 2:* For an MRC receiver based on the LMMSE channel estimate and operating at low SNR, the uplink achievable rate for each active terminal in a one-bit massive MIMO system can be approximated by

$$R_{\text{MRC}}^{(\text{UL})} = \log_2 \left( 1 + \frac{\alpha^2 (\sigma^2 M + \rho_u)}{\rho_u \alpha^2 (K-1) + \alpha^2 + 1 - 2/\pi} \right). \quad (47)$$

*Proof:* See Appendix B. ■

*Theorem 3:* For a ZF receiver based on the LMMSE channel estimate and operating at low SNR, the uplink achievable rate for each active terminal in a one-bit massive MIMO system can be approximated by

$$R_{\text{ZF}}^{(\text{UL})} = \log_2 \left( 1 + \frac{\alpha^2 \sigma^2 (M - K - 1) + \alpha^2 \rho_u}{\alpha^2 (K-1) (\rho_u - \sigma^2) + \alpha^2 + 1 - 2/\pi} \right). \quad (48)$$

*Proof:* See Appendix C. ■

Note that equal achievable rates can be guaranteed to all active terminals due to the power control strategies; thus we omit the subscript  $k$  in the achievable rate expressions of (47) and (48).

### B. Downlink Transmission

For conventional MIMO systems<sup>1</sup>, in order to satisfy the conditions for uplink-downlink duality, the precoding vectors are assumed to satisfy  $\mathbf{t}_k = \sqrt{q_k / \|\mathbf{w}_k\|^2} \mathbf{w}_k$ , where the scaling  $q_k$  is chosen to satisfy the constraint that the total transmission power of the uplink and downlink are equal. However, according to Theorem 1, in order to ensure the SINR duality holds, we consider the modified matched-filter (MF) and ZF precoding schemes for the  $k$ th user

$$\mathbf{t}_{k,\text{MF}} = \sqrt{\frac{q_k}{\|\mathbf{A}_u \mathbf{w}_{k,\text{MRC}}\|^2}} \mathbf{A}_u \mathbf{w}_{k,\text{MRC}}, \quad (49)$$

$$\mathbf{t}_{k,\text{ZF}} = \sqrt{\frac{q_k}{\|\mathbf{A}_u \mathbf{w}_{k,\text{ZF}}\|^2}} \mathbf{A}_u \mathbf{w}_{k,\text{ZF}}, \quad (50)$$

where  $\mathbf{w}_{k,\text{MRC}}$  and  $\mathbf{w}_{k,\text{ZF}}$  are the  $k$ th column of  $\mathbf{W}_{\text{MRC}}$  and  $\mathbf{W}_{\text{ZF}}$ , respectively.

*Theorem 4:* Employing the precoding strategies shown in (49) and (50) and allocating transmit power as in (32), the downlink achievable rate of each active terminal in one-bit MIMO systems can be approximated by (47) and (48), respectively.

*Proof:* This follows directly from the uplink-downlink SINR duality in Theorem 1. ■

## VI. OPTIMAL SYSTEM DESIGN OF EE-SE TRADEOFF FOR ONE-BIT MASSIVE MIMO

Most prior work focuses on optimizing either the spectral or the energy efficiency individually. Since the two are competing objectives, improving one inevitably leads to a degradation of the other. Since both are of importance for current wireless systems, here we investigate the joint optimization of both the spectral and energy efficiency for one-bit massive MIMO systems by properly allocating resources such as the number of active terminals  $K$ , the training length  $\tau$  and the system operating power  $\rho_u$  for a fixed number of BS antennas  $M$  and coherence interval  $T$ .

The SE is defined as the average achievable rate over the coherence interval  $T$ :

$$\mathcal{F}_{\text{SE}} = \frac{(T - \tau)}{T} K \left( \gamma R_A^{(\text{UL})} + (1 - \gamma) R_A^{(\text{DL})} \right), \quad (51)$$

where  $A \in \{\text{MRC}, \text{ZF}\}$ . The EE is defined as the average achievable rate achieved per-unit transmit power:

$$\mathcal{F}_{\text{EE}} = \frac{(T - \tau) K}{T} \frac{\gamma R_A^{(\text{UL})} + (1 - \gamma) R_A^{(\text{DL})}}{\gamma \mathbb{E}\{P_{\text{total}}^{(\text{UL})}\} + (1 - \gamma) \mathbb{E}\{P_{\text{total}}^{(\text{DL})}\}}. \quad (52)$$

Note that the total transmit power  $P_{\text{total}}^{(\text{UL})}$  changes since the active terminals are randomly selected and the large scale fading coefficients are different for each user. Therefore, instead of considering the instantaneous total transmit power, here we consider the average total transmit power. For different

<sup>1</sup>The term ‘‘conventional MIMO system’’ refers here to MIMO systems equipped with infinite-resolution ADC/DACs.



channel realizations, the average total transmit power can be expressed as

$$\begin{aligned} \mathbb{E} \left\{ P_{\text{total}}^{(\text{UL})} \right\} &= \sum_{k=1}^K p_k = \sum_{k=1}^K \mathbb{E} \{ \beta_k^{-1} \} \rho_u \\ &= K \rho_u \frac{r_{\max}^{\kappa+2} - r_{\min}^{\kappa+2}}{\bar{d}(1 + \kappa/2)(r_{\max}^2 - r_{\min}^2)r_{\min}^{\kappa}}. \end{aligned} \quad (53)$$

Since, according to the uplink-downlink duality derived in Theorem 1, any target rate achieved in the uplink can be achieved in the downlink while maintaining  $P_{\text{total}}^{(\text{UL})} = P_{\text{total}}^{(\text{DL})}$ , we drop the superscript '(UL)' and '(DL)' and rewrite the SE and EE as

$$\mathcal{F}_{\text{SE}} = \frac{(T - \tau)}{T} K R_A, \quad (54)$$

$$\mathcal{F}_{\text{EE}} = \frac{(T - \tau)}{T} \frac{R_A \bar{d}(1 + \kappa/2)(r_{\max}^2 - r_{\min}^2)r_{\min}^{\kappa}}{\rho_u (r_{\max}^{\kappa+2} - r_{\min}^{\kappa+2})}. \quad (55)$$

Next we investigate the number of users and operating power that jointly optimize the spectral and energy efficiency in one-bit massive MIMO systems. Since the SE and EE are competing objectives, we will use a multiobjective optimization approach to balance the two criteria and obtain a Pareto optimal solution. In particular, we will employ the weighted product method, expressed as

$$(K^*, \tau^*, \rho_u^*) = \arg \max_{K, \tau, \rho} (\mathcal{F}_{\text{SE}})^{w_{\text{SE}}} \cdot (\mathcal{F}_{\text{EE}})^{w_{\text{EE}}}, \quad (56)$$

where the weights  $w_{\text{SE}}$  and  $w_{\text{EE}}$  indicate the relative significance of the objective functions. The specific Pareto optimal operating point provided as the solution depends on which weights are used. For example, if  $w_{\text{SE}} \neq 0$  and  $w_{\text{EE}} = 0$ , then the optimization problem reduces to maximizing only the spectral efficiency at the expense of the energy efficiency. The opposite is true if  $w_{\text{SE}} = 0$  and  $w_{\text{EE}} \neq 0$ . The specific choice of the weights depends on the system design requirements. In the simulations that follow, we will show the Pareto boundary achieved by a wide range of weights, as well as the specific solution obtained with  $w_{\text{SE}} = w_{\text{EE}} = 1$ .

Note that, by using the uplink-downlink duality, it is guaranteed that the optimal solutions for the number of active terminals  $K^*$ , training length  $\tau^*$  and the operating power  $\rho_u^*$  is the same for both the uplink and downlink.

## VII. NUMERICAL RESULTS

For our simulations, we consider a cell with a radius of 500 meters and  $K_{\max}$  terminals distributed randomly and uniformly over the cell, with the exclusion of a central disk of radius 100 meters. According to [34], we choose  $\bar{d} = 10^{0.8}$  and  $\kappa = 3.8$  for a typical urban cellular environment.

We first evaluate the validity of the approximations for the achievable rate for the MRC and ZF receivers derived in Theorems 2 and 3 compared with the ergodic rate given in (38). Fig. 2 shows the sum spectral efficiency versus operating power  $\rho_u$  for different numbers of transmit antennas  $M = \{32, 64, 128\}$ . We choose the length of the coherence interval to be  $T = 200$ , the number of active terminals  $K = 8$ , and training length  $\tau = 16$ , such that the relative pilot

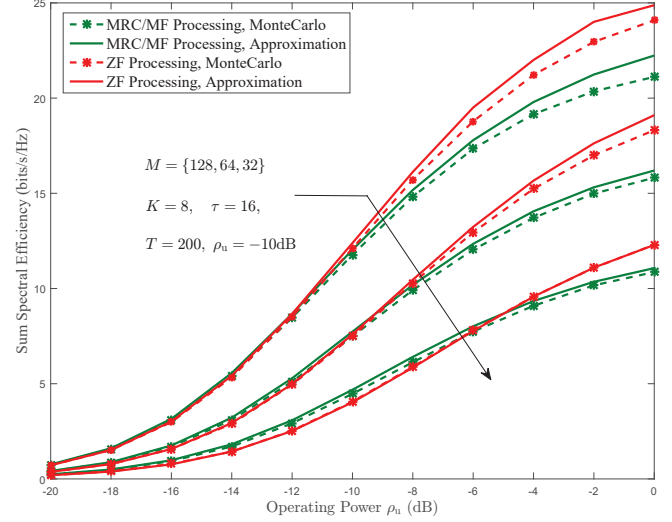


Fig. 2. Sum spectral efficiency versus operating power  $\rho_u$  with  $M = \{32, 64, 128\}$ ,  $K = 8$ ,  $T = 200$  and  $\tau = 16$  for MRC/MF and ZF processing.

length  $\tau_0 = 2$ . The straight lines represent the sum spectral efficiencies obtained numerically from (47) and (48), while the dashed lines with markers represent the ergodic sum spectral efficiencies obtained from (38). Apparently, for both MRC and ZF processing, the gap between the bounds and the empirical ergodic rates is small, especially for low operating power. This implies that the approximation for the achievable rate given in (47) and (48) is a good predictor of the performance of one-bit massive MIMO systems. Thus, in the following plots we will show only the approximation when evaluating performance.

In the downlink, as analyzed in previous sections, since each transmit antenna should be equipped with a power amplifier in order to implement the power allocation in (32), it is desirable that the power amplifier gains be confined to a limited range of values. Fig. 3 shows the cumulative density function of the downlink transmit power of the individual BS antennas for  $M = \{32, 64, 128\}$  and total transmit power  $P_{\text{total}}^{(\text{DL})} = P_{\text{total}}^{(\text{UL})} = 10\text{dB}$  for the modified MF and ZF precoding in (49) and (50), respectively. We assume the statistically-aware power control strategy is used in the uplink, while the power allocation as in (32) is implemented in the downlink to ensure the SINR duality. We can see that the ranges of the power amplifier gains are indeed small. We also observe that the ranges for both the modified MF and ZF precoding decrease as the number of transmit antennas increases. Interestingly, we note that for larger antenna arrays, both modified MF and ZF precoding have almost the same ranges of power amplifier gains.

Next we study the SE vs. EE tradeoff. In conventional massive MIMO systems, a common rule of thumb is that the number of BS antennas  $M$  should be roughly 10 times more than the number of active terminals  $K$  and the optimal pilot training length should be equal to the number of active terminals, i.e.,  $\tau = K$ . An interesting question is whether the aforementioned rule is still true for one-bit massive MIMO systems. To answer this question, Fig. 4 plots the SE versus

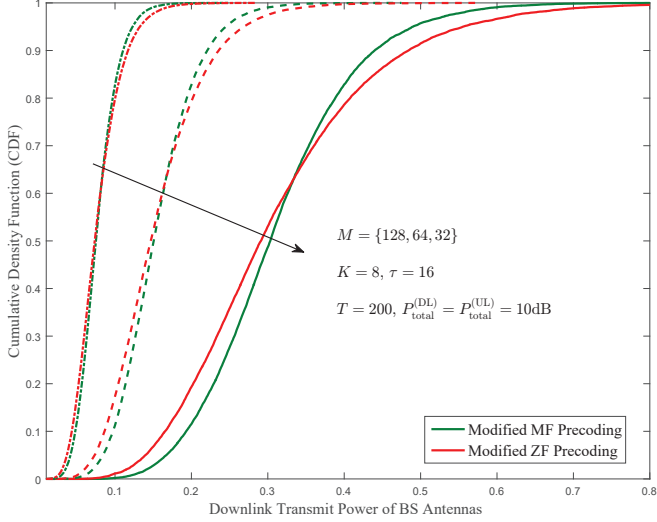
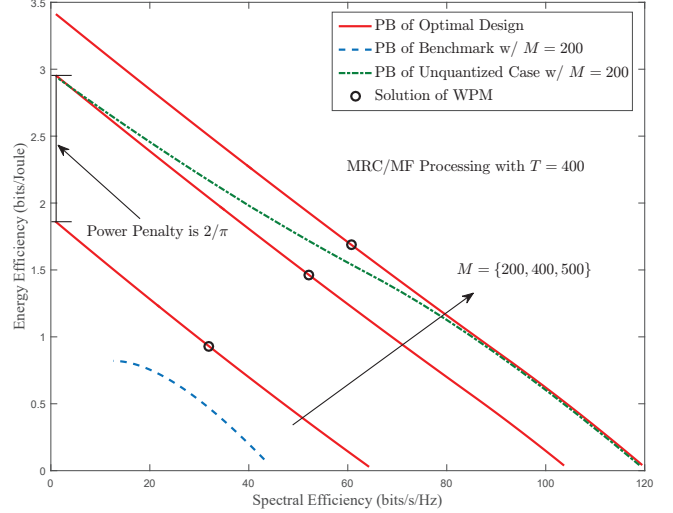


Fig. 3. Cumulative density function of the downlink transmit power of BS antennas over 500 channel realizations for different numbers of transmit antennas  $M$  and  $P_{\text{total}}^{(\text{DL})} = P_{\text{total}}^{(\text{UL})} = 10\text{dB}$ .

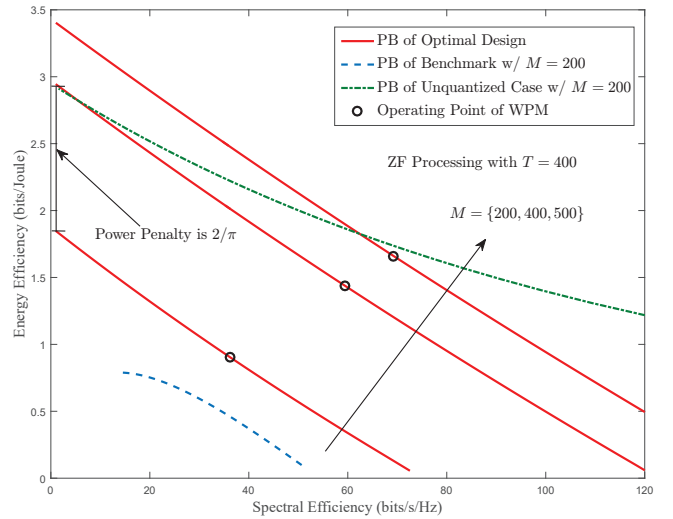
EE for different numbers of BS antennas  $M$  and a coherence interval  $T = 400$  for MRC/MF and ZF processing. The solid lines and the dash-dot line represent the Pareto boundary of the one-bit system design and the unquantized case [16], respectively, when the number of active terminals  $K$ , the training length  $\tau$  and the operating power  $\rho_u$  are jointly optimized. The dashed line represents the Pareto boundary of the benchmark case for one-bit systems, which is obtained by setting  $M = 200$ ,  $K = 0.1M$ ,  $\tau = K$  and only optimizing the operating power  $\rho_u$ .

The tradeoff between SE and EE is clearly visible and in this example is nearly linear for the one-bit case; e.g., increasing the SE by 10% leads to approximately a 10% reduction in EE. We also note that the non-optimized benchmark configuration leads to a considerably lower operating point in terms of both SE and EE; optimization of the system parameters provides a significant benefit. In both figures, we see the power penalty of  $2/\pi$  that separates the performance of conventional unquantized massive MIMO from that for one-bit systems at low SNR (low SE). The curves for the conventional unquantized system plotted for  $M = 200$  show different behavior for the MRC/MF and ZF methods. We see that for MRC/MF processing, the optimized one-bit system can achieve the same performance as the unquantized system when the number of antennas deployed ranges from twice as many ( $M = 400$ ) at low SE values (low operating SNR) to 2.5 times as many ( $M = 500$ ) for large SEs (high operating SNR). For ZF processing, twice as many antennas are also needed for low SNR/SE, but the amount of extra antennas grows beyond a factor of 2.5 for high SNR/SE. This is because at high SNR, unlike MRC/MF, the quality of the estimated CSI in the conventional system improves and ZF can eliminate more and more multiuser interference. The operating point achieved by (56) with  $w_{\text{SE}} = w_{\text{EE}} = 1$  shown by the small circle on each Pareto boundary yields a reasonable compromise between the SE and EE.

Fig. 5-6 show the optimal ratio of  $K^*/M$  and the optimal



(a)



(b)

Fig. 4. Pareto boundary of energy efficiency versus spectral efficiency for optimal design, benchmark and unquantized case with  $M = \{200, 400, 500\}$  and  $T = 400$  for (a) MRC/MF processing and (b) ZF processing.

relative training length  $\tau_0$ , which are obtained by (56) with  $w_{\text{SE}} = w_{\text{EE}} = 1$ , versus the number of BS antennas  $M$  for different coherence intervals  $T$ . We see from the plots that the optimal number of active terminals and pilot training length does not generally follow the assumption of  $K = 0.1M$  and  $\tau = K$ . The optimal loading factor  $K/M$  can be as high as 0.25 for MRC/MF when  $M = 100$  and  $T = 400$ , but it decreases as  $M$  increases or  $T$  decreases. The optimal length of the training interval is considerably larger for one-bit systems compared with conventional MIMO, where  $\tau_0 = 1$  is known to be optimal. For values of  $M$  below 200, 2-4 times as much training should be used depending on  $T$  and the type of processing used. For larger and larger  $M$ ,  $\tau_0$  converges towards unity.

Since our analysis is based on the assumption of low

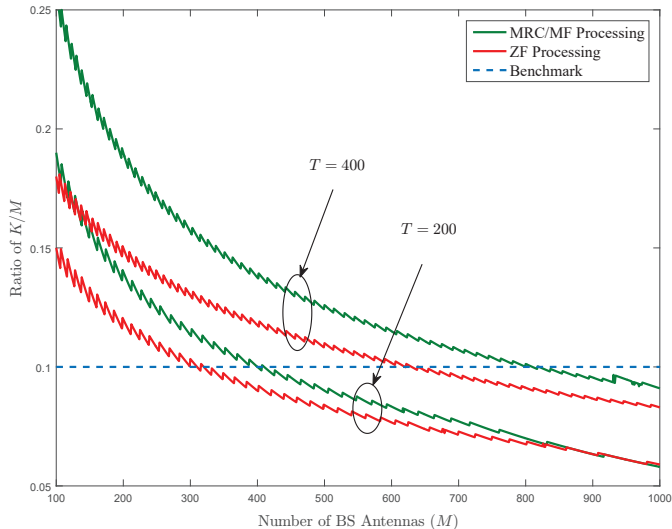


Fig. 5. Optimal ratio of  $K/M$  versus the number of BS antennas  $M$  for different coherence intervals  $T$  for MRC/MF and ZF processing.

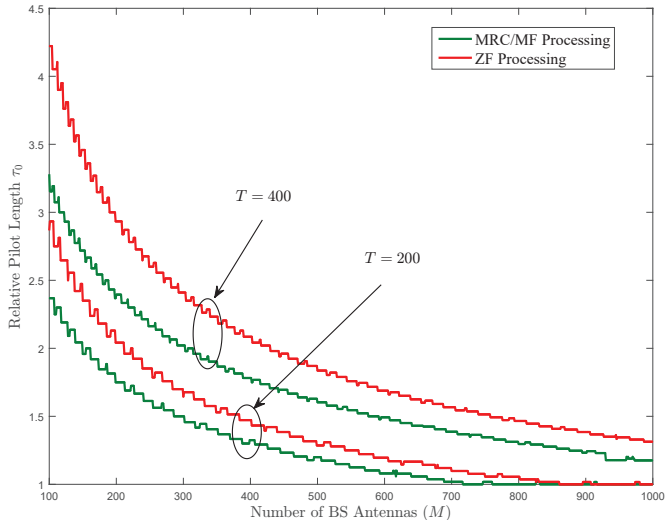


Fig. 6. Optimal relative pilot length  $\tau_0$  versus the number of BS antennas  $M$  for different coherence intervals  $T$  for MRC/MF and ZF processing.

operating power, we finally verify whether this assumption is reasonable or not. Fig. 7 illustrates the optimal operating power  $\rho_u$  versus the number of BS antennas  $M$  for MRC/MF and ZF processing for different coherence intervals. We see from the curves that for  $M$  and  $T$  larger than 100, the optimal operating SNR is less than -9 dB, which implies that our low operating power assumption is justified. Combining the observations from Fig. 5 with Fig. 7, it is interesting to note that, when one considers EE together with SE, MRC/MF processing competes favorably with ZF processing since it uses less power while serving more terminals with lower computational complexity.

## VIII. CONCLUSIONS

This paper has investigated the optimal design of single-cell massive MIMO systems employing one-bit ADC/DACs

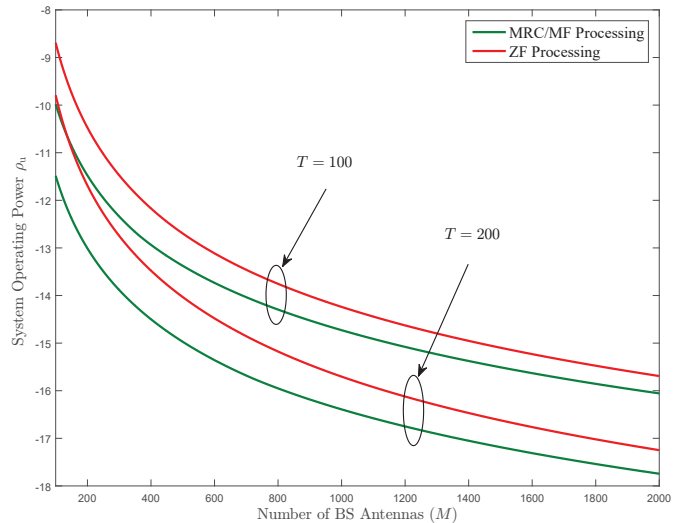


Fig. 7. Optimal system operating power  $\rho_u$  versus the number of BS antennas  $M$  for different coherence intervals  $T$  for MRC/MF and ZF processing.

at the BS. Using the Bussgang decomposition, we derived approximate closed-form expressions at low SNR for the uplink achievable rate in a one-bit massive MIMO system assuming MRC or ZF receivers that operate with imperfect CSI. We derived the conditions under which uplink-downlink SINR duality holds for one-bit massive MIMO systems, and under these conditions we examined the spectral and energy efficiency tradeoff. We optimized a weighted product of the two competing objectives over the number of active users, the length of the training data and the operating power and compared the resulting performance with an unoptimized benchmark system design and the optimized unquantized massive MIMO system. The results show that depending on the operating SNR, the one-bit systems require 2-2.5 times more antennas to achieve the same spectral/energy efficiency as a conventional unquantized system. We also illustrated that, in general, one-bit systems require more training than conventional systems at the optimal operating point, up to 2.5 times as much depending on the length of the coherence interval and the number of antennas. We also observed that MRC/MF processing can be competitive when compared with ZF when energy efficiency is taken into account, since it can service more users with less operating power and computational complexity.

## APPENDIX A

Suppose  $\gamma_k$  is the target SINR achieved by the  $k$ th active user for both uplink and downlink transmission. For the downlink, we assume the precoding vector is

$$\mathbf{t}_k = \sqrt{q_k} \frac{\mathbf{A}_u \mathbf{w}_k}{\|\mathbf{A}_u \mathbf{w}_k\|_2} \triangleq \sqrt{q_k} \hat{\mathbf{t}}_k. \quad (57)$$

where  $q_k = \|\mathbf{t}_k\|_2^2$ ,  $\hat{\mathbf{t}}_k = \mathbf{t}_k / \sqrt{q_k}$  with  $\|\hat{\mathbf{t}}_k\|_2^2 = 1$ , and the power allocation for the downlink is

$$\mathbf{Q} = \text{diag} \left( \sum_{i=1}^K \mathbf{t}_i \mathbf{t}_i^H \right)^{\frac{1}{2}}. \quad (58)$$

With this notation, we have

$$P_{\text{total}}^{(\text{DL})} = \text{tr}(\mathbf{Q}\mathbf{Q}^H) = \sum_{i=1}^K q_i, \quad (59)$$

which allows us to transfer the problem of power allocation for each antenna into the problem of power allocation for each precoding vector. Substituting (58) into (24) and assuming uncorrelated quantizer noise  $\boldsymbol{\eta}_d$  such that  $\mathbf{C}_{\boldsymbol{\eta}_d} = (1 - \frac{2}{\pi})\mathbf{I}$ , the downlink SINR can be expressed as

$$\begin{aligned} \gamma_k &= \frac{q_k |\mathbf{g}_k^T \hat{\mathbf{t}}_k|^2}{\sum_{i \neq k}^K q_i |\mathbf{g}_k^T \hat{\mathbf{t}}_i|^2 + (\frac{\pi}{2} - 1) \mathbf{g}_k^T \text{diag}(\sum_{i=1}^K q_i \hat{\mathbf{t}}_i \hat{\mathbf{t}}_i^H) \mathbf{g}_k^* + \frac{\pi}{2}} \\ &= \frac{q_k |\mathbf{g}_k^T \hat{\mathbf{t}}_k|^2}{\sum_{i \neq k}^K q_i |\mathbf{g}_k^T \hat{\mathbf{t}}_i|^2 + (\frac{\pi}{2} - 1) \sum_{i=1}^K q_i \hat{\mathbf{t}}_i^H \text{diag}(\mathbf{g}_k^* \mathbf{g}_k^T) \hat{\mathbf{t}}_i + \frac{\pi}{2}}. \end{aligned} \quad (60)$$

The second line in (60) holds thanks to the following identity for two matrices  $\mathbf{A}$ ,  $\mathbf{B}$ :

$$\text{tr}(\mathbf{A} \text{diag}(\mathbf{B})) = \text{tr}(\mathbf{B} \text{diag}(\mathbf{A})). \quad (61)$$

Then by using the definitions in (34) and (35), (60) can be rewritten as

$$[\mathbf{D}]_{k,k} = \frac{q_k}{\sum_{i=1}^K q_i [\boldsymbol{\Psi}]_{k,i} + \frac{\pi}{2}}. \quad (62)$$

Therefore, according to (66) we can express the downlink SINR conditions for all active users as

$$\mathbf{q} = \mathbf{D}\boldsymbol{\Psi}\mathbf{q} + \frac{\pi}{2}\mathbf{D}\mathbf{1}_K \Leftrightarrow \mathbf{q} = \frac{\pi}{2}(\mathbf{I} - \mathbf{D}\boldsymbol{\Psi})^{-1}\mathbf{D}\mathbf{1}_K, \quad (63)$$

where  $\mathbf{q} = [q_1, \dots, q_K]^T$  and  $\mathbf{1}_K$  denotes the all-one vector.

For the uplink transmission, since  $\frac{\mathbf{A}_u \mathbf{w}_k}{\|\mathbf{A}_u \mathbf{w}_k\|_2} = \hat{\mathbf{t}}_k$ , the uplink SINR in (23) can be rewritten as

$$\gamma_k = \frac{p_k |\hat{\mathbf{t}}_k^T \mathbf{g}_k|^2}{\sum_{i \neq k}^K p_i |\hat{\mathbf{t}}_k^T \mathbf{g}_i|^2 + 1 + (1 - \frac{2}{\pi}) \hat{\mathbf{t}}_k^T \mathbf{A}_u^{-2} \hat{\mathbf{t}}_k^*}. \quad (64)$$

After substituting (17) into (64), the denominator of (64) can be expressed as

$$\begin{aligned} &\sum_{i \neq k}^K p_i |\hat{\mathbf{t}}_k^T \mathbf{g}_i|^2 + 1 + \left(\frac{\pi}{2} - 1\right) \hat{\mathbf{t}}_k^T \text{diag} \left( \sum_{i=1}^K p_i \mathbf{g}_i \mathbf{g}_i^H + \mathbf{I} \right) \hat{\mathbf{t}}_k^* \\ &= \sum_{i \neq k}^K p_i |\hat{\mathbf{t}}_k^T \mathbf{g}_i|^2 + \left(\frac{\pi}{2} - 1\right) \sum_{i=1}^K p_i \hat{\mathbf{t}}_k^H \text{diag}(\mathbf{g}_i^* \mathbf{g}_i^T) \hat{\mathbf{t}}_k + \frac{\pi}{2}. \end{aligned} \quad (65)$$

By using the notation of (34) and (35), we can then express (64) as

$$[\mathbf{D}]_{k,k} = \frac{p_k}{\sum_{i=1}^K p_i [\boldsymbol{\Psi}]_{i,k} + \frac{\pi}{2}}. \quad (66)$$

After stacking it into vector form, we have

$$\mathbf{p} = \mathbf{D}\boldsymbol{\Psi}^T \mathbf{p} + \frac{\pi}{2}\mathbf{D}\mathbf{1}_K \Leftrightarrow \mathbf{p} = \frac{\pi}{2}(\mathbf{I} - \mathbf{D}\boldsymbol{\Psi}^T)^{-1}\mathbf{D}\mathbf{1}_K, \quad (67)$$

where  $\mathbf{p} = [p_1, \dots, p_K]^T$ . Using the matrix inversion lemma and the fact that  $\mathbf{D} = \mathbf{D}^T$ , we have  $(\mathbf{I} - \mathbf{D}\boldsymbol{\Psi}^T)^{-1}\mathbf{D} = ((\mathbf{I} - \mathbf{D}\boldsymbol{\Psi})^{-1}\mathbf{D})^T$ , which implies that the total transmit power for

the uplink and downlink are the same:

$$P_{\text{total}}^{(\text{UL})} = \|\mathbf{p}\|_1 = \|\mathbf{q}\|_1 = P_{\text{total}}^{(\text{DL})}. \quad (68)$$

This completes the proof.

## APPENDIX B

From (43), we first focus on deriving the terms  $\text{E}\left\{|\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}}|^2\right\}$  and  $\text{E}\left\{|\mathbf{w}_k^T \mathbf{g}_i^{\text{eff}}|^2\right\}$ . For the MRC receiver, we have

$$\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}} = (\hat{\mathbf{g}}_k^{\text{eff}})^H \hat{\mathbf{g}}_k^{\text{eff}} + (\hat{\mathbf{g}}_k^{\text{eff}})^H \boldsymbol{\varepsilon}_k. \quad (69)$$

Since the elements of the estimated effective channel  $(\hat{\mathbf{g}}_k^{\text{eff}})$  are i.i.d, according to [35] we can thus obtain

$$\begin{aligned} \text{E}\left\{|\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}}|^2\right\} &= \text{E}\left\{|\left(\hat{\mathbf{g}}_k^{\text{eff}}\right)^H \hat{\mathbf{g}}_k^{\text{eff}}|^2\right\} + \text{E}\left\{|\left(\hat{\mathbf{g}}_k^{\text{eff}}\right)^H \boldsymbol{\varepsilon}_k|^2\right\} \\ &= \sigma^4(M^2 + M) + \sigma^2(\rho_u - \sigma^2)M = \sigma^2 M(\sigma^2 M + \rho_u). \end{aligned} \quad (70)$$

For  $i \neq k$ , we have

$$\text{E}\left\{|\left(\hat{\mathbf{g}}_k^{\text{eff}}\right)^H \mathbf{g}_i^{\text{eff}}|^2\right\} = \text{Var}\left\{\left(\hat{\mathbf{g}}_k^{\text{eff}}\right)^H \mathbf{g}_i^{\text{eff}}\right\} = \rho_u \sigma^2 M. \quad (71)$$

Next we focus on the term  $\text{AQN}_k$ . Using (30), we rewrite  $\text{AQN}_k$  as

$$\text{AQN}_k = (\alpha_u^2 + 1 - 2/\pi) \text{E}\left\{\|\mathbf{w}_k^T\|_2^2\right\}. \quad (72)$$

With the MRC receiver  $\mathbf{w}_k^T = (\hat{\mathbf{g}}_k^{\text{eff}})^H$ , we have

$$\text{AQN}_k = (\alpha_u^2 + 1 - 2/\pi) \sigma^2 M. \quad (73)$$

Substituting (70)-(73) into (43), we can arrive at Theorem 2.

## APPENDIX C

For the ZF receiver, we have

$$\mathbf{W}_{\text{ZF}}^T \mathbf{G}_{\text{eff}} = \mathbf{W}_{\text{ZF}}^T (\hat{\mathbf{G}}_{\text{eff}} + \boldsymbol{\varepsilon}_{\text{eff}}) = \mathbf{I} + \mathbf{W}_{\text{ZF}}^T \boldsymbol{\varepsilon}_{\text{eff}}, \quad (74)$$

where  $\boldsymbol{\varepsilon}_{\text{eff}} = [\boldsymbol{\varepsilon}_1^{\text{eff}}, \dots, \boldsymbol{\varepsilon}_K^{\text{eff}}]$ . Therefore

$$\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}} = 1 + \mathbf{w}_k^T \boldsymbol{\varepsilon}_k^{\text{eff}}, \quad (75)$$

and  $\text{E}\left\{|\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}}|^2\right\}$  can be expressed as

$$\begin{aligned} \text{E}\left\{|\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}}|^2\right\} &= 1 + \text{E}\left\{|\mathbf{w}_k^T \boldsymbol{\varepsilon}_k^{\text{eff}}|^2\right\} \\ &= 1 + (\rho_u - \sigma^2) \text{E}\left\{\left[\left(\hat{\mathbf{G}}_{\text{eff}}^H \hat{\mathbf{G}}_{\text{eff}}\right)^{-1}\right]_{k,k}\right\}. \end{aligned} \quad (76)$$

Since the elements of  $\hat{\mathbf{G}}_{\text{eff}}$  are approximated as i.i.d Gaussian,  $(\hat{\mathbf{G}}_{\text{eff}}^H \hat{\mathbf{G}}_{\text{eff}})$  is a  $K \times K$  central Wishart matrix with  $M$  degrees of freedom. Thus

$$\text{E}\left\{|\mathbf{w}_k^T \mathbf{g}_k^{\text{eff}}|^2\right\} = 1 + \frac{\rho_u - \sigma^2}{\sigma^2(M - K)}. \quad (77)$$

For  $i \neq k$ , we have

$$\begin{aligned} \mathbf{w}_k^T \mathbf{g}_i^{\text{eff}} &= \text{E}\left\{|\mathbf{w}_k^T \boldsymbol{\varepsilon}_i^{\text{eff}}|^2\right\} \\ &= (\rho_u - \sigma^2) \text{E}\left\{\left[\left(\hat{\mathbf{G}}_{\text{eff}}^H \hat{\mathbf{G}}_{\text{eff}}\right)^{-1}\right]_{k,k}\right\} \end{aligned}$$

$$= \frac{\rho_u - \sigma^2}{\sigma^2(M - K)}. \quad (78)$$

Similarly, we can obtain

$$\text{AQN}_k = \frac{\alpha_u^2 + 1 - 2/\pi}{\zeta_k^2(M - K)}. \quad (79)$$

Substituting (76)-(79) into (43), we arrive at Theorem 3.

## REFERENCES

- [1] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 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*, vol. 8, no. 5, pp. 742–758, 2014.
- [3] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.
- [4] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan 2013.
- [5] B. Murmann, "ADC performance survey 1997-2016." [Online]. Available: <http://web.stanford.edu/~murmanc/adcsurvey.html>
- [6] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC," *Mathematics*, 2014.
- [7] J. Choi, J. Mo, and R. W. Heath, "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 2005–2018, May 2016.
- [8] Y. Li, C. Tao, L. Liu, A. Mezghani, G. Seco-Granados, and A. Swindlehurst, "Channel estimation and performance analysis in one-bit massive MIMO systems." [Online]. Available: <http://arxiv.org/abs/1504.03516>
- [9] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath Jr., "Uplink performance of wideband massive MIMO with one-bit ADCs," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 87–100, 2017.
- [10] A. Mezghani and J. Nosssek, "Analysis of Rayleigh-fading channels with 1-bit quantized output," in *IEEE International Symposium on Information Theory (ISIT)*, July 2008, pp. 260–264.
- [11] J. A. Nosssek and M. T. Ivrlač, "Capacity and coding for quantized MIMO systems," in *Proceedings of the international conference on Wireless communications and mobile computing*. ACM, 2006, pp. 1387–1392.
- [12] 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*, vol. 15, no. 2, pp. 1293–1308, Feb 2016.
- [13] H. Q. Ngo, M. Matthaiou, and E. G. Larsson, "Massive MIMO with optimal power and training duration allocation," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 605–608, Dec 2014.
- [14] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C. X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo, "Green radio: Radio techniques to enable energy-efficient wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, June 2011.
- [15] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–37, June 2011.
- [16] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, April 2013.
- [17] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, June 2015.
- [18] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Structural and multidisciplinary optimization*, vol. 26, no. 6, pp. 369–395, 2004.
- [19] J. Branke, K. Deb, K. Miettinen, and R. Slowiński, *Multiobjective optimization: interactive and evolutionary approaches*. Springer, 2008, vol. 5252.
- [20] E. Björnson, E. A. Jorswieck, M. Debbah, and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 14–23, 2014.
- [21] J. Tang, D. K. C. So, E. Alsusa, and K. A. Hamdi, "Resource efficiency: A new paradigm on energy efficiency and spectral efficiency tradeoff," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4656–4669, Aug 2014.
- [22] C. He, B. Sheng, P. Zhu, X. You, and G. Y. Li, "Energy- and spectral-efficiency tradeoff for distributed antenna systems with proportional fairness," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 5, pp. 894–902, May 2013.
- [23] Z. Zhang, Z. Chen, M. Shen, and B. Xia, "Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 848–863, April 2016.
- [24] Y. Hao, Z. Song, S. Hou, and H. Li, "Energy- and spectral-efficiency tradeoff in massive MIMO systems with inter-user interference," in *IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Aug 2015, pp. 553–557.
- [25] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," MIT Research Lab. Electronics, Tech. Rep. 216, 1952.
- [26] M. Biguesh and A. B. Gershman, "Downlink channel estimation in cellular systems with antenna arrays at base stations using channel probing with feedback," *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 1330–1339, 2004.
- [27] O. Bar-Shalom and A. Weiss, "DOA estimation using one-bit quantized measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 868–884, 2002.
- [28] S. M. Kay, *Fundamentals of statistical signal processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice Hall, 1993.
- [29] G. Jacovitti and A. Neri, "Estimation of the autocorrelation function of complex Gaussian stationary processes by amplitude clipped signals," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 239–245, Jan 1994.
- [30] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.
- [31] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, April 2003.
- [32] H. Cramér, *Random variables and probability distributions*. Cambridge University Press, 2004, vol. 36.
- [33] Q. Zhang, S. Jin, K.-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*, vol. 8, no. 5, pp. 966–981, Oct 2014.
- [34] W. Choi and J. G. Andrews, "The capacity gain from intercell scheduling in multi-antenna systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 714–725, 2008.
- [35] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Fundamental Trends Communication Information Theory*, vol. 1, no. 1, pp. 1–182, Jun 2004.