Performance of Linear Receivers for Wideband Massive MIMO with One-Bit ADCs

Christopher Mollén①, Junil Choi②, Erik G. Larsson①, and Robert W. Heath Jr.②
① Linköping University, Dept. of Electrical Engineering, 581 83 Linköping, Sweden
② University of Texas at Austin, Dept. of Electrical and Computer Engineering, Austin, TX 78712, USA

***Abstract*—The power consumption of analog-to-digital converters (ADCs) grows linearly in the number of antennas in massive MIMO base stations. To reduce power consumption, one-bit ADCs can be used. It is believed that the nonlinear distortion of one-bit ADCs makes channel estimation and symbol equalization in such systems computationally complex and resource demanding. In this paper, it is shown that low-complexity linear channel estimation and symbol equalization are feasible in massive MIMO with one-bit ADCs when the number of channel taps is large. The effective SINR of the received symbol estimates of a maximum-ratio combiner with estimated channel state information is 4 dB lower in a system with one-bit ADCs than in an equivalent unquantized system.**

I. Background and Motivation

The idea of massive MIMO (Multiple-Input Multiple-Output) is to increase the number of antenna chains at the base station, so that a large number of users can be served over the same time-frequency resource. Compared to conventional systems that exploit small numbers of antennas, there are many benefits: increased spectral efficiency and reduced *radiated* power. To reduce *consumed* power, however, the power consumption of each antenna chain has to decrease. The analog-to-digital converter (ADC) in each chain consumes a significant part of the total power. Since the power consumption of ADCs is proportional to the number of quantization levels, quantization has to be made coarser. From a power efficiency point of view, the coarsest ADCs—one-bit ADCs—are the most desirable.

It has been shown that the capacity of a single-input single-output (SISO) system with one-bit ADCs only decreases by a factor $2/\pi$ at low SNR compared to an unquantized system [1]; this is also true in MIMO systems [2], [3]. It has also been shown that the capacity of a noise-free MIMO system with one-bit ADCs grows linearly in the number of receive antennas [3]. In massive MIMO, it is thus theoretically possible to reach high data rates, also with one-bit ADCs. In [2], [3], perfect channel state knowledge at the base station was assumed. In reality, the channel state information has to be estimated and is never perfect, especially if the received signals are quantized by one-bit ADCs.

Prior work has presented solutions both for equalization and for channel estimation with one-bit ADCs, see [4]–[8] for example. The complexity of these equalization methods, however, becomes increasingly difficult to handle in practical wideband scenarios, where the numbers of antennas and users are large and where the channel is frequency-selective. The channel estimation techniques proposed for non-sparse channels also require long training sequences to obtain good enough channel state knowledge. A typical mobile channel has a coherence interval that could be only a few hundred symbols long. Since massive MIMO is usually implemented in time-division duplex, both the downlink and uplink have to fit in this short interval along with the training sequences [9]. The training sequences therefore have to be short.

In [10], we showed that the transmitted symbols can be recovered error-free in the limit of infinitely many antennas. In this paper, we establish capacity lower bounds for finite number of antennas. The main contributions of this work are:

- It is proposed to use a low-complexity maximum-ratio combiner (MRC) for symbol detection and a low-complexity LMMSE (Linear Minimum-Mean-Square Error) technique with short pilot sequences for channel estimation in the massive MIMO uplink with one-bit ADCs. This implementation is the kind of architecture commonly considered for massive MIMO with high resolution quantization [11]. The implementation is thus feasible in terms of complexity.
- A technique similar to [12] is used to derive an analytical expression for the rate achievable with MRC using estimated channel state information in a massive MIMO system with one-bit ADCs. If the quantized system and the unquantized system have perfect channel state information, the power loss incurred by using one-bit ADCs is equal to the $2/\pi \approx -2$ dB limit. If the two systems estimate the channel with equally short training sequences, the power loss increases to approximately -4 dB.
- It is shown that frequency-selective channels can be helpful in massive MIMO systems with one-bit quantizers, in that such channels improve the performance of linear detectors.

II. System Model

The uplink of the massive MIMO system in Figure 1 is considered. The base station is equipped with M antennas and there are K single-antenna users. All signals are modeled in complex baseband and are sampled at Nyquist-rate with perfect synchronization.

In symbol duration n, base station antenna m receives:

$$y_m[n] \triangleq \sum_{k=1}^{K} \sum_{\ell=0}^{L-1} \sqrt{P_k} g_{mk}[\ell] x_k[n-\ell] + z_m[n], \tag{1}$$

where $x_k[n]$ is the zero-mean transmit signal from user k, whose power $\mathsf{E}\big[|x_k[n]|^2\big] = 1$, P_k is the transmit power of user k and $z_m[n] \sim \mathcal{CN}(0, N_0)$ is a random variable that

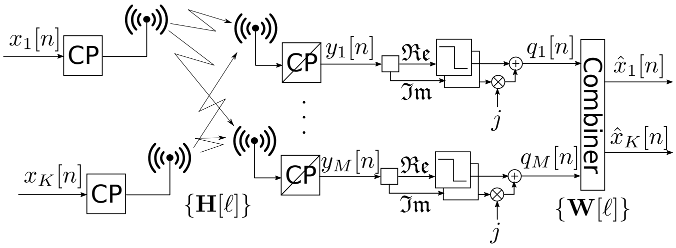

Fig. 1. The system model of a massive MIMO uplink with one-bit ADCs.

models the thermal noise of the base station hardware. It is assumed that $z_m[n]$ is IID over n and m and independent of all other variables. We assume that the L-tap impulse response $\{g_{mk}[\ell]\}$ of the channel between user k and antenna m can be divided into two parts:

$$g_{mk}[\ell] = \sqrt{\beta_k} h_{mk}[\ell]. \tag{2}$$

The small-scale fading has to be estimated by the base station, only its mean $\mathsf{E}\big[h_{mk}[\ell]\big] = 0$ and variance is *a priori* known:

$$\mathsf{E}\big[|h_{mk}[\ell]|^2\big] = \frac{1}{L}, \quad \forall \ell. \tag{3}$$

The base station is assumed to know the large-scale fading β_k, which generally changes so slowly over time that an accurate estimate easily can be obtained in most cases.

In a wideband system, the number of channel taps L can be large—in the order of tens. For example, a system that uses 15 MHz of bandwidth over a channel with 1 µs of maximum excess delay, which corresponds to a moderately frequency-selective channel, has $L = 15$ channel taps, c.f. [13], where the "Extended Typical Urban Model" has a maximum excess delay of 5 µs.

Upon reception, the signals are quantized into:

$$q_m[n] \triangleq \frac{1}{\sqrt{2}} \operatorname{sign}\big(\mathfrak{Re}(y_m[n])\big) + j\frac{1}{\sqrt{2}} \operatorname{sign}\big(\mathfrak{Im}(y_m[n])\big). \tag{4}$$

We assume that the in-phase and quadrature signals are separately sampled, each by identical one-bit ADCs, and that the threshold of the quantization is zero. Other thresholds are studied in [14], [15]. The scaling of the quantized signal is arbitrary but chosen such that $q_m[n]$ has unit power.

The users also transmit a cyclic prefix that is $L-1$ symbol long:

$$x_k[n] = x_k[N+n], \quad -L < n < 0, \tag{5}$$

which is introduced to simplify the mathematical exposition.

III. Quantization

In this section, some properties of the quantization with one-bit ADCs are derived. These results are used later in the channel estimation and the rate analysis.

We define the quantization noise as

$$e_m[n] \triangleq q_m[n] - \rho y_m[n], \tag{6}$$

where the scaling factor ρ is chosen to minimize the error variance

$$E \triangleq \mathsf{E}\big[|e_m[n]|^2\big]. \tag{7}$$

The error variance is minimized by the Wiener solution:

$$\rho = \frac{\mathsf{E}\big[y_m^*[n] q_m[n]\big]}{\mathsf{E}\big[|y_m[n]|^2\big]}. \tag{8}$$

Note that the distribution of $e_m[n]$ depends on the distribution of the received signal $y_m[n]$ in a nonlinear way, see (6), and that $e_m[n]$ is uncorrelated to $y_m[n]$ due to the choice of ρ (the orthogonality principle). Because $e_m[n]$ later is treated as additional noise and any information about the symbol that it might contain is discarded, it is referred to as quantization noise.

In the next step, we define the "expected received power given all transmit signals" and the "average received power":

$$\begin{aligned} P_{\text{rx}}[n] &\triangleq \mathsf{E}\left[|y_m[n]|^2 \;\middle|\; \{x_k[n]\}\right] \\ &= N_0 + \frac{1}{L}\sum_{k=1}^{K}\sum_{\ell=0}^{L-1} \beta_k P_k \big|x_k[n-\ell]\big|^2, \end{aligned} \tag{9}$$

$$\bar{P}_{\text{rx}} \triangleq \mathsf{E}\left[|y_m[n]|^2\right] = N_0 + \sum_{k=1}^{K} \beta_k P_k. \tag{10}$$

When the number of channel taps L in (9) is large, the two powers $P_{\text{rx}}[n]$ and $\bar{P}_{\text{rx}}$ are close to equal. The law of large numbers applied to the inner sum in (9) gives

$$P_{\text{rx}}[n] \xrightarrow{\text{a.s.}} \bar{P}_{\text{rx}}, \quad L \to \infty, \qquad \forall n. \tag{11}$$

Because of the cyclic prefix, the block length N cannot be shorter than L. We therefore assume that N grows together with L in (11). As we will see later, the convergence can be fast and the left-hand side in (11) is close to its limit also for L much smaller than usual block sizes.

Remark 1: Note that the number of terms in (9) is KL. Therefore, the relative difference between the expected received power given all transmit signals and its mean $|P_{\text{rx}}[n] - \bar{P}_{\text{rx}}|/P_{\text{rx}}[n]$ can become small, not only with increasing L, but also with increasing number of users K. This happens if there is no dominating user, i.e., some index k for which $\beta_k P_k \gg \beta_{k'} P_{k'}$ for all $k' \neq k$. Note that by doing power control such that P_k is chosen proportional to $\frac{1}{\beta_k}$, dominating users can be avoided. The expected received power given all transmit signals is thus close to its average also in some narrowband systems with a large number of users.

The next lemma gives the scaling factor and the variance of the quantization noise.

Lemma 1: If the fading is IID Rayleigh, i.e., $h_{mk}[\ell] \sim \mathcal{CN}(0, \frac{1}{L})$ for all m, k and ℓ, then the scaling factor defined in (8) is given by

$$\rho = \sqrt{\frac{2}{\pi}} \frac{\mathsf{E}\Big[\sqrt{P_{\text{rx}}[n]}\Big]}{\bar{P}_{\text{rx}}}, \tag{12}$$

and the quantization noise has the variance

$$E = 1 - \rho^2 \bar{P}_{\text{rx}}. \tag{13}$$

In a wideband system, the scaling factor approaches

$$\rho \to \sqrt{\frac{2}{\pi \bar{P}_{\text{rx}}}}, \quad L \to \infty \tag{14}$$

and the variance of the quantization noise approaches

$$E \to 1 - \frac{2}{\pi}, \quad L \to \infty. \tag{15}$$

Proof: The proof can be found in [16] and is omitted here for conciseness. ■

If $P_{\text{rx}}[n]$ were constant, the error variance in (15) would equal its limit. That is the reason this limit coincides with the mean-squared quantization error of a one-bit ADC in [17] and what is called the distortion factor of one-bit ADCs in [12].

The following corollary to Lemma 1 gives the limit of the relative quantization noise variance, which is defined as

$$Q \triangleq \frac{E}{|\rho|^2}. \tag{16}$$

Corollary 1: The relative quantization noise variance in a wideband system approaches

$$Q \to Q' \triangleq \bar{P}_{\text{rx}}\Big(\frac{\pi}{2} - 1\Big), \quad L \to \infty. \tag{17}$$

Note that the relative quantization noise variance $Q \geq Q'$ is always greater than its limit, because Jensen's inequality says that $\rho \leq \sqrt{\frac{2}{\pi P_{\text{rx}}}}$ is smaller than its limit for all L, since the square root is concave. This means that there is less quantization noise in a wideband system, where the number of taps L is large, than in a narrowband system.

To later be able to compare the quantized and unquantized systems, we note that the analyses in the following sections of the paper also can be applied to the unquantized system. If there is no quantization, the variance of the quantization error $E = 0$ and thus the relative quantization noise $Q = Q' = 0$.

IV. Channel Estimation

During the training period, user k transmits an $N = N_{\text{p}}$ symbol long pilot sequence, and the base station performs LMMSE channel estimation. The frequency-domain pilot signal of user k is $\mathsf{x}_k[\nu] \triangleq \frac{1}{\sqrt{N_{\text{p}}}} \sum_{n=0}^{N_{\text{p}}-1} x_k[n] e^{-j2\pi n\nu/N_{\text{p}}}$, which is chosen as

$$\mathsf{x}_k[\nu] = \begin{cases} 0, & (\nu \mod K) + 1 \neq k \\ \sqrt{K} e^{j\theta_k[\nu']}, & (\nu \mod K) + 1 = k \end{cases}, \tag{18}$$

where the integer index $\nu' \triangleq \frac{\nu - k + 1}{K} \in [0, \frac{N_{\text{p}}}{K} - 1]$ and $\theta_k[\nu']$ is a phase that is known to the base station.

We denote the Fourier transform of the quantized signal by:

$$\mathsf{q}_m[\nu] \triangleq \frac{1}{\sqrt{N_{\text{p}}}} \sum_{n=0}^{N_{\text{p}}-1} q_m[n] e^{-j2\pi n\nu/N_{\text{p}}}. \tag{19}$$

Analogously, $\mathsf{y}_m[\nu]$, $\mathsf{e}_m[\nu]$ and $\mathsf{z}_m[\nu]$ denote the Fourier transform of $y_m[n]$, $e_m[n]$ and $z_m[n]$ respectively. Note that $\mathsf{z}_m[n] \sim \mathcal{CN}(0, N_0)$ is IID because the transform is unitary. The Fourier transform of the channel impulse response is scaled differently,

$$\mathsf{h}_{mk}[\nu] \triangleq \sum_{\ell=0}^{L-1} h_{mk}[\ell] e^{-j2\pi \ell\nu/N_{\text{p}}}, \tag{20}$$

so that the received signal in the frequency domain can be written as:

$$\mathsf{y}_m[\nu] = \sum_{k=1}^{K} \sqrt{\beta_k P_k}\, \mathsf{h}_{mk}[\ell] \mathsf{x}_k[\nu] + \mathsf{z}_m[\nu]. \tag{21}$$

By (6), the frequency-domain received quantized signal is

$$\begin{aligned} \mathsf{q}_m[\nu] &= \rho \mathsf{y}_m[n] + \mathsf{e}_m[n] & (22)\\ &= \rho\Big(\sum_{k=1}^{K} \sqrt{\beta_k P_k}\, \mathsf{h}_{mk}[\nu] \mathsf{x}_k[\nu] + \mathsf{z}_m[\nu]\Big) + \mathsf{e}_m[\nu] & (23)\\ &= \rho\sqrt{\beta_{k'} P_{k'} K}\, \mathsf{h}_{mk'}[\nu] e^{j\theta_{k'}[\frac{\nu - k' + 1}{K}]} + \rho \mathsf{z}_m[\nu] + \mathsf{e}_m[\nu], & (24) \end{aligned}$$

where $k' \triangleq (\nu \mod K) + 1$ in the last step is the index of the user whose transmit signal is nonzero at tone ν. The sequence $\{\mathsf{q}_m[\nu K + k - 1],\ \nu = 0, \ldots, \frac{N_p}{K} - 1\}$ is thus a phase-rotated and noisy version of the frequency-domain channel of user k, sampled with period K. Because the channel only has L time-domain taps, an observation of the channel tap $h_{mk}[\ell]$ can be made if the sampling period $K \geq \frac{N_{\text{p}}}{L}$ according to the sampling theorem. By performing an inverse transform on the received samples that belong to user k, we obtain

$$\begin{aligned} h'_{mk}[\ell] &\triangleq \sqrt{\frac{K}{N_{\text{p}}}} \sum_{\nu=0}^{\frac{N_{\text{p}}}{K}-1} \mathsf{q}_m[\nu K + k - 1] e^{j2\pi\ell(\nu K + k - 1)/N_{\text{p}}} e^{-j\theta_k[\nu]} & (25)\\ &= \rho\sqrt{\beta_k P_k K} \sqrt{\frac{K}{N_{\text{p}}}} \sum_{\nu=0}^{\frac{N_{\text{p}}}{K}-1} \mathsf{h}_{mk}[\nu K + k - 1] e^{j2\pi\ell(\nu K + k - 1)/N_{\text{p}}} \\ &\quad + \underbrace{\rho \sqrt{\frac{K}{N_{\text{p}}}} \sum_{\nu=0}^{\frac{N_{\text{p}}}{K}-1} \mathsf{z}_m[\nu K + k - 1] e^{j2\pi\ell(\nu K + k - 1)/N_{\text{p}}} e^{-j\theta_k[\nu]}}_{\triangleq z'_{mk}[\ell]} \\ &\quad + \underbrace{\sqrt{\frac{K}{N_{\text{p}}}} \sum_{\nu=0}^{\frac{N_{\text{p}}}{K}-1} \mathsf{e}_m[\nu K + k - 1] e^{j2\pi\ell(\nu K + k - 1)/N_{\text{p}}} e^{-j\theta_k[\nu]}}_{\triangleq e'_{mk}[\ell]} & (26)\\ &= \rho\sqrt{\beta_k P_k N_{\text{p}}}\, h_{mk}[\ell] + \rho z'_{mk}[\ell] + e'_{mk}[\ell]. & (27) \end{aligned}$$

In the first step (25), $\mathsf{q}_m[\nu]$ is replaced by the expression in (24). Then, in (26), we identify the first sum as the inverse transform of the channel spectrum. We note that the Fourier transform is unitary and therefore $z'_{mk}[\ell] \sim \mathcal{CN}(0, N_0)$ is independent across m, k, ℓ and $\mathsf{E}\big[|e'_{mk}[\ell]|^2\big] = \mathsf{E}\big[|\mathsf{e}'_{mk}[\ell]|^2\big] = E$.

The LMMSE estimate of the channel is thus

$$\begin{aligned} \hat{h}_{mk}[\ell] &\triangleq h'_{mk}[\ell] \frac{\mathsf{E}\big[h^*_{mk}[\ell] h'_{mk}[\ell]\big]^*}{\mathsf{E}\big[|h'_{mk}[\ell]|^2\big]} & (28)\\ &= h'_{mk}[\ell] \frac{\rho\sqrt{\beta_k P_k N_{\text{p}}}}{\rho^2 \beta_k P_k N_{\text{p}} + L(\rho^2 N_0 + E)}. & (29) \end{aligned}$$

The variance of the estimate is

$$\mathsf{E}\big[|\hat{h}_{mk}[\ell]|^2\big] = \phi_k \frac{1}{L}, \tag{30}$$

where the channel quality factor is

$$\phi_k \triangleq \frac{\beta_k P_k N_\text{p}}{\beta_k P_k N_\text{p} + L(N_0 + Q)}. \quad (31)$$

The estimation error $\epsilon_{mk}[\ell] \triangleq h_{mk}[\ell] - \hat{h}_{mk}[\ell]$ is uncorrelated to the channel estimate and has variance

$$\mathsf{E}\big[|\epsilon_{mk}[\ell]|^2\big] = \big(1-\phi_k\big)\frac{1}{L}. \quad (32)$$

Note that we can write the channel in terms of its estimate:

$$h_{mk}[\ell] = \hat{h}_{mk}[\ell] + \epsilon_{mk}[\ell]. \quad (33)$$

When L is large, the difference in channel quality between the quantized and unquantized systems when N_p is the same in both systems is

$$\Delta \triangleq \frac{\phi_k\big|_{Q=0}}{\phi_k\big|_{Q=\bar{P}_\text{rx}(\frac{\pi}{2}-1)}} = 1 + \frac{N_0 + \sum_{k'=1}^{K}\beta_{k'}P_{k'}}{N_0 + \beta_k P_k N_\text{p}/L}\Big(\frac{\pi}{2} - 1\Big). \quad (34)$$

Note that this ratio is decreasing in N_p and that $\Delta \to 1$ as $N_\text{p} \to \infty$. If $N_\text{p} = KL$ and $\beta_k P_k = P$ for all k, for some power P, then

$$\Delta = \frac{\pi}{2} \approx 2\,\text{dB}. \quad (35)$$

Remark 2: The phases $\theta_k[\nu]$ of the pilot symbols are used to distribute the energy of the pilot sequences over time. For example, the choice $\theta_k[\nu] = 0$, $\forall k, \nu$, results in pilot signals whose energy is not evenly distributed:

$$x_k[n] = \begin{cases} \sqrt{\frac{N_\text{p}}{K}}, & \text{if } n = \nu\frac{N_\text{p}}{K} + k - 1, \quad \nu \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}. \quad (36)$$

The same training signal is transmitted every $\frac{N_\text{p}}{K}$-th symbol duration, which conveys little new information for channel estimation. Drawing the phases $\theta_k[\nu]$ from a uniform distribution on the interval $[0, 2\pi)$ works in most cases when $h_{mk}[\ell] \sim \mathcal{CN}(0, \frac{1}{L})$ IID, except when $N_\text{p} = KL$ and, at the same time, either K, $\frac{L}{K}$ are prime numbers or L, $\frac{K}{L}$ are prime numbers. Then the (noise-free) received signal becomes periodic and little new information is gained by multiple copies of the same received signal. In this case, the error variance becomes larger than predicted by (32).

V. Uplink Data Transmission

In this section, the uplink data transmission is studied for one block of $N = N_\text{d}$ symbols. Practical detection with maximum-ratio combining (MRC) based on the estimated channel is presented and applied to the massive MIMO system with one-bit ADCs. The distribution of the symbol estimation error due to quantization and how it affects OFDM is also analyzed. Finally, the performance is evaluated by deriving an achievable rate for the system.

A. Maximum-Ratio Combining

In the maximum-ratio combiner, the received quantized signals are filtered by the filter, whose impulse response is the time-reversed and conjugated channel impulse response:

$$\mathbf{W}[\ell] = \frac{1}{\sqrt{M}}\mathbf{D}_{\boldsymbol{\phi}}^{-\frac{1}{2}}\hat{\mathbf{H}}^\mathsf{H}[-\ell], \quad (37)$$

where $\hat{\mathbf{H}}[\ell]$ is the matrix with $\hat{h}_{mk}[\ell]$ on its (m,k)-th position. The scaling by $\mathbf{D}_{\boldsymbol{\phi}} \triangleq \text{diag}(\phi_1, \ldots, \phi_K)$ is done for convenience, such that each row of the combiner matrix has energy $\frac{1}{L}$. The output estimates of the transmit signal are given by

$$\hat{\mathbf{x}}[n] \triangleq \begin{pmatrix} \hat{x}_1[n] \\ \vdots \\ \hat{x}_K[n] \end{pmatrix} = \sum_{\ell=-L+1}^{0} \mathbf{W}[\ell]\mathbf{q}\big[[n-\ell]_{N_\text{d}}\big], \quad (38)$$

where $[x]_{N_\text{d}} \triangleq x \mod N_\text{d}$, and $\mathbf{q}[n] \triangleq (q_1[n], \ldots, q_M[n])^\mathsf{T}$.

B. Estimation Error Due to Quantization

In this section, we assume that the channel taps are uncorrelated to each other to be able to use the Gram-Schmidt process to partition the quantization noise into a sum of terms, in which each term is correlated to a specific channel tap. By doing that, we show that the estimation error due to quantization consists of two parts: one radial distortion and one circularly symmetric. In a wideband system however, the radial distortion is negligible.

If $\{h_{mk}[\ell]\}$ is a set of uncorrelated variables, using the Gram-Schmidt process, the quantization noise can be written as follows:

$$e_m[n] = \sum_{k=1}^{K}\sum_{\ell=0}^{L-1} \frac{\mathsf{E}\big[h_{mk}^*[\ell]e_m[n] \mid \{x_k[n]\}\big]}{\mathsf{E}\big[|h_{mk}[\ell]|^2\big]} h_{mk}[\ell] + d_m[n], \quad (39)$$

where $d_m[n]$ is the residual error that is uncorrelated to all $\{h_{mk}[\ell]\}$ conditioned on $\{x_k[n]\}$. The following lemma gives the coefficients in this sum.

Lemma 2: If $h_{mk}[\ell] \sim \mathcal{CN}(0, \frac{1}{L})$, the normalized conditional correlation

$$\frac{\mathsf{E}\big[h_{mk}^*[\ell]e_m[n] \mid \{x_k[n]\}\big]}{\mathsf{E}\big[|h_{mk}[\ell]|^2\big]} = \sqrt{\frac{2}{\pi}}x_k[n-\ell]\tau[n] \quad (40)$$

$$\xrightarrow{\text{a.s.}} 0, \quad L \to \infty, \quad (41)$$

where

$$\tau[n] \triangleq \frac{\sqrt{P_\text{rx}[n]}}{P_\text{rx}[n]} - \frac{\mathsf{E}\Big[\sqrt{P_\text{rx}[n]}\Big]}{\bar{P}_\text{rx}}. \quad (42)$$

Proof: The proof can be found in [16] and is omitted here for conciseness. ■

The quantization noise now becomes:

$$e_m[n] = \sqrt{\frac{2}{\pi}}\tau[n]\bar{y}_m[n] + d_m[n], \quad (43)$$

where the noise-free received signal is

$$\bar{y}_m[n] \triangleq \sum_{k=1}^{K}\sum_{\ell=0}^{L-1} h_{mk}[\ell]x_k[n-\ell]. \quad (44)$$

By using (6) to write $q_m[n] = \rho y_m[n] + e_m[n]$, the symbol estimate of the receive combiner in (38) can be written as:

$$\hat{x}_k[n] = \sum_{m=1}^{M} \sum_{\ell=-L+1}^{0} w_{km}[\ell]\big(\rho y_m[n-\ell] + e_m[n-\ell]\big). \tag{45}$$

Thus, we define the error due to quantization as

$$e'_k[n] \triangleq \sum_{m=1}^{M} \sum_{\ell=-L+1}^{0} w_{km}[\ell] e_m[n-\ell] \tag{46}$$

$$= \sqrt{\frac{2}{\pi}} \sum_{m=1}^{M} \sum_{\ell=-L+1}^{0} w_{km}[\ell]\tau[n-\ell]\bar{y}_m[n-\ell]$$
$$+ \sum_{m=1}^{M} \sum_{\ell=-L+1}^{0} w_{km}[\ell] d_m[n-\ell] \tag{47}$$

Because the first term in (47) contains the noise-free received signal $\bar{y}_m[n]$, it results in a radial distortion, i.e., error that contains a term that is proportional to the transmit signal $x_k[n]$ or the negative transmit signal $-x_k[n]$ (depending on the sign of $\tau[n]$). When the number of channel taps goes to infinity, two things happen. Firstly, the radial distortion that contains $\tau[n]$ vanishes because $\tau[n] \to 0$ as $L \to \infty$ according to Lemma 2. Secondly, because the number of terms in the second sum in (47) grows with L, the sum converges in distribution to a Gaussian random variable according to the central limit theorem, the variance of which is

$$\mathsf{E}\big[|e'_k[n]|^2\big] \to \mathsf{E}\big[|d_m[n]|^2\big] = E, \quad L \to \infty. \tag{48}$$

Thus,

$$e'_k[n] \xrightarrow{\text{dist.}} \mathcal{CN}(0, E), \quad L \to \infty. \tag{49}$$

C. Achievable Rate

An achievable rate for the uplink of the massive MIMO system with one-bit ADCs is given. The limit of the achievable rate, when the number of channel taps L grows large, is then derived in closed form. In a numerical study, it is seen that this limit closely approximates the achievable rate of a wideband system also with practical values of L.

Using the orthogonality principle, the estimate $\hat{x}_k[\nu]$ can be written as a sum of two terms

$$\hat{x}_k[n] = a x_k[n] + \zeta_k[n], \tag{50}$$

where $\zeta_k[n]$ is the residual error. By choosing the factor $a \triangleq \mathsf{E}\big[x_k^*[n]\hat{x}_k[n]\big]$, the variance of the error $\zeta_k[n]$ is minimized and the error becomes uncorrelated to the transmit signal $x_k[n]$. The variance of the error term is then

$$\mathsf{E}\big[|\zeta_k[n]|^2\big] = \mathsf{E}\big[|\hat{x}_k[n]|^2\big] - \big|\mathsf{E}\big[x_k^*[n]\hat{x}_k[n]\big]\big|^2. \tag{51}$$

In [18], it was shown that uncorrelated noise with a Gaussian distribution minimizes the mutual information $I(x_k[n]; \hat{x}_k[n])$ among all distributions. By assuming that the noise term is Gaussian, we obtain the achievable rate (in bits)

$$R_k = \log_2\left(1 + \frac{\big|\mathsf{E}\big[x_k^*[n]\hat{x}_k[n]\big]\big|^2}{\mathsf{E}\big[|\hat{x}_k[n]|^2\big] - \big|\mathsf{E}\big[x_k^*[n]\hat{x}_k[n]\big]\big|^2}\right) \tag{52}$$

for user k.

The following theorem gives the limit of the achievable rate R_k for MRC as the number of channel taps goes to infinity.

Theorem 1: When the small-scale fading coefficients are IID and $h_{mk}[\ell] \sim \mathcal{CN}\big(0, \frac{1}{L}\big)$, the achievable rate R_k in (52) for MRC approaches

$$R_k \to R'_k, \quad L \to \infty, \tag{53}$$

where

$$R'_k \triangleq \log_2\Big(1 + \frac{2}{\pi}\frac{\phi_k \beta_k P_k M}{N_0 + \sum_{k'=1}^{K} \beta_{k'} P_{k'}}\Big). \tag{54}$$

Proof: See the Appendix. ■

In the same way as in Theorem 1, an achievable rate for the unquantized MRC system can be obtained:

$$R_0 = \log_2\left(1 + \frac{\phi_k \beta_k P_k M}{N_0 + \sum_{k'=1}^{K} \beta_{k'} P_{k'}}\right) \tag{55}$$

Note that the rate of the unquantized system is achievable for all L.

Remark 3: When the pilot length $N_\text{p} = KL$ and $\beta_k P_k = P$ is the same for all k, the SINR in (55) of the unquantized MRC system is

$$\frac{\pi}{2} \frac{\phi_k\big|_{Q=0}}{\phi_k\big|_{Q=\bar{P}_\text{rx}(\frac{\pi}{2}-1)}} = \frac{\pi^2}{4} \approx 4\,\text{dB} \tag{56}$$

higher than the SINR in (54) of the quantized system, independently of transmit power P_k. With perfect channel state information $\phi_k\big|_{Q=\bar{P}_\text{rx}(\frac{\pi}{2}-1)} = \phi_k\big|_{Q=0} = 1$, the difference is $\frac{\pi}{2} \approx 2\,\text{dB}$. This performance degradation coincides with earlier results that show that the capacity of a SISO channel [1] and a MIMO channel [2] decreases by a factor $\frac{2}{\pi}$ at low SNR with one-bit ADCs and perfect channel state information. Quantization thus leads to a twofold SINR degradation: the MRC symbol estimates suffer from a $\frac{\pi}{2}$ degradation from quantization noise and an additional degradation due to the $\frac{\pi}{2}$ lower channel quality ϕ_k observed in (35).

To evaluate the derived achievable rate, we perform Monte-Carlo simulations to obtain R_k numerically for Gaussian transmit signals $x_k[n] \sim \mathcal{CN}(0, 1)$. The result is shown in Figure 2, where it can be seen that the rate R_k is close to its limit R'_k when the number of channel taps $L \geq 5$, which corresponds to a moderately frequency-selective channel. When the number of users is large, the convergence is immediate and R_k is a good approximation of R_k for all L. This was observed in Remark 1.

VI. Conclusion

We have showed that it is possible to use low-complexity linear receivers and channel estimators in a massive MIMO system even with one-bit ADCs when the channel is frequency selective. In a wideband system, where the number of channel taps is large, and where the received power from each user is the same, the effective SINR of MRC decreases by $\frac{\pi^2}{4} \approx 4\,\text{dB}$ as compared to the equivalent unquantized case. This SINR degradation is independent of the number of users and the transmit power.

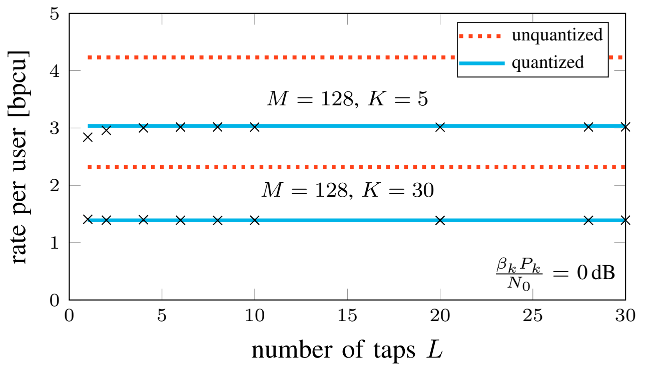

Fig. 2. The achievable rate R_k marked × and its limit R_k' drawn with a solid line for a system with 128 antennas that serves 5 and 30 users over an L-tap channel with Rayleigh fading taps. The dotted line shows the rate of the unquantized system. The channel is estimated with $N_\text{p} = KL$ pilot symbols. The rate R_k was computed for 500 channel realizations, during each of which 10000 Gaussian distributed symbols per user were sent.

Appendix
Outline of Proof of Theorem 1

To evaluate the achievable rate in (52), the MRC estimate $\hat{x}_k[n]$ is partitioned into components that are uncorrelated to the transmit signal $x_k[n]$. By using (33), the received signal can be rewritten as follows

$$y_m[n] = \sum_{k=1}^{K} \sqrt{\phi_k \beta_k P_k} \bar{y}_{mk}[n] + u_m[n] + z_m[n], \tag{57}$$

where

$$\bar{y}_{mk}[n] \triangleq \sum_{\ell=0}^{L-1} \frac{1}{\sqrt{\phi_k}} \hat{h}_{mk}[\ell] x_k[n-\ell], \tag{58}$$

$$u_m[n] \triangleq \sum_{k=1}^{K} \sum_{\ell=0}^{L-1} \sqrt{\beta_k P_k} \epsilon_{mk}[\ell] x_k[n-\ell]. \tag{59}$$

Using (57), the symbol estimate in (45) becomes:

$$\hat{x}_k[n] = \sum_{m=1}^{M} \sum_{\ell=-L+1}^{0} w_{km}[\ell] \Bigg(\rho \sum_{k'=1}^{K} \sqrt{\phi_{k'} \beta_{k'} P_{k'}} \bar{y}_{mk'}[n-\ell] + \rho u_m[n-\ell] + \rho z_m[n-\ell] + e_m[n-\ell] \Bigg) \tag{60}$$

$$= \rho \sum_{k'=1}^{K} \sqrt{\phi_{k'} \beta_{k'} P_{k'}} \hat{x}'_{kk'}[n] + \rho u'_k[n] + \rho z'_k[n] + e'_k[n], \tag{61}$$

where

$$\hat{x}'_{kk'}[n] \triangleq \sum_{m=1}^{M} \sum_{\ell=-L+1}^{0} w_{km}[\ell] \bar{y}_{mk'}[n-\ell] \tag{62}$$

$$u'_k[n] \triangleq \sum_{m=1}^{M} \sum_{\ell=-L+1}^{0} w_{km}[\ell] u_m[n-\ell]; \tag{63}$$

the terms $z'_k[n]$ and $e'_k[n]$ are defined analogously to $u'_k[n]$.

The terms $\hat{x}'_{kk'}[n]$ can further be split up in a part that is correlated to the transmit signal and a part that is not:

$$\hat{x}'_{kk'}[n] = \alpha_{kk'} x_k[n] + i_{kk'}[n], \tag{64}$$

where $\alpha_{kk'} = \mathsf{E}\big[x_k^*[n] \hat{x}'_{kk'}[n]\big]$ and $i_{kk'}[n]$ is the interference that is uncorrelated to $x_k[n]$. It is seen that $\alpha_{kk'} = 0$ for all $k' \neq k$, i.e., only the term $\hat{x}'_{kk}[n]$ is correlated to the transmit signal $x_k[n]$. We denote the gain $G_k \triangleq |\alpha_{kk}|^2$ and the interference variance $I_{kk'} \triangleq \mathsf{E}\big[|i_{kk'}[n]|^2\big]$. To derive G_k and $I_{kk'}$ the second moments of the channel coefficients have to be evaluated. It was done in [19], [20] for an IID Rayleigh fading channel $h_{mk}[\ell] \sim \mathcal{CN}(0, \frac{1}{L})$:

$$G_k = M, \qquad I_{kk'} = 1. \tag{65}$$

The estimated signal can thus be written as the sum of the following terms:

$$\hat{x}_k[n] = \rho \sum_{k'=1}^{K} \sqrt{\phi_{k'} \beta_{k'} P_{k'}} \big(\alpha_{kk'} x_k[n] + i_{kk'}[n]\big) + \rho u'_k[n] + \rho z'_k[n] + e'_k[n]. \tag{66}$$

It can be shown that each term in this sum is uncorrelated to the other terms. Most correlations are easy to show, except the correlation between the error due to quantization $e'_k[n]$ and the transmit signal $x_k[n]$. This step can be found in [16] and is omitted for conciseness.

The variances of $u'_k[n]$ and $z'_k[n]$ are given by

$$\mathsf{E}\big[|u'_k[n]|^2\big] = \mathsf{E}\big[|u_m[n]|^2\big] = \sum_{k'=1}^{K} \beta_{k'} P_{k'} \big(1 - \phi_{k'}\big), \tag{67}$$

$$\mathsf{E}\big[|z'_k[n]|^2\big] = \mathsf{E}\big[|z_m[n]|^2\big] = N_0, \tag{68}$$

By evaluating the expectations in the rate expression in (52), we obtain

$$\big|\mathsf{E}\big[x_k^*[n] x_k[n]\big]\big|^2 \to \rho^2 \phi_k \beta_k P_k G_k, \tag{69}$$

$$\mathsf{E}\big[|\hat{x}_k[n]|^2\big] \to \rho^2 \Big(\phi_k \beta_k P_k G_k + \sum_{k'=1}^{K} \big(\phi_k \beta_k P_k I_{kk'} + \beta_{k'} P_{k'} (1 - \phi_{k'}) \big) + N_0 + Q' \Big), \tag{70}$$

as $L \to \infty$. Here we used Corollary 1. Letting the number of channel taps $L \to \infty$ thus results in the rate

$$R'_k = \log_2(1 + \mathsf{SINR}_k), \tag{71}$$

where

$$\mathsf{SINR}_k = \frac{\phi_k \beta_k P_k G_k}{\sum_{k'=1}^{K} \big(\phi_k \beta_k P_k I_{kk'} + \beta_{k'} P_{k'} (1 - \phi_{k'}) \big) + N_0 + Q'}. \tag{72}$$

By letting $I_{kk'} = 1$ and $G_k = M$ as in (65) and using Corollary 1 and (10), we obtain (54).

References

[1] A. J. Viterbi and J. K. Omura, *Principle of Digital Communication and Coding*. McGraw-Hill, Inc., 1979.

[2] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *Proc. IEEE Int. Symp. Inform. Theory*, Jun. 2007.

[3] J. Mo and R. W. Heath, Jr., "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, Oct. 2015.

[4] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC," *ArXiv E-Print*, Apr. 2014, arXiv:1404.7736 [cs.IT].

[5] J. Choi, J. Mo, and R. W. Heath, Jr., "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *ArXiv E-Print*, Jul. 2015, arXiv:1507.04452 [cs.IT].

[6] S. Wang, Y. Li, and J. Wang, "Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2156–2168, Apr. 2015.

[7] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *ArXiv E-Print*, Sep. 2015, arxiv:1509.07928 [cs.IT].

[8] J. Mo, P. Schniter, N. González-Prelcic, and R. W. Heath, Jr., "Channel estimation in millimeter wave MIMO systems with one-bit quantization," in *Proc. Asilomar Conf. Signals, Syst. and Comput.*, Nov. 2014.

[9] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Oct. 2010.

[10] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath, Jr., "One-bit ADCs in wideband massive MIMO systems with OFDM transmission," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Mar. 2016.

[11] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[12] Q. Bai and J. A. Nossek, "Energy efficiency maximization for 5G multi-antenna receivers," *Trans. Emerging Telecommun. Technologies*, vol. 26, no. 1, pp. 3–14, Jan. 2015.

[13] *3GPP TS36.141 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) Conformance Testing (Release 10)*, 3GPP Std., Rev. V10.2.0, 2011.

[14] T. Koch and A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5421–5445, Sep. 2013.

[15] R. Narasimha, M. Lu, N. R. Shanbhag, and A. C. Singer, "BER-optimal analog-to-digital converters for communication links," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3683–3691, Jul. 2012.

[16] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath, Jr., "Linear receivers and channel estimators for wideband massive MIMO with one-bit ADCs," 2016, in preparation.

[17] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.

[18] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[19] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[20] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "On the optimality of single-carrier transmission in large-scale antenna systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 276–279, Apr. 2012.