Blind Detection for MIMO Systems With Low-Resolution ADCs Using Supervised Learning

Yo-Seb Jeon*, Song-Nam Hong†, and Namyoon Lee‡
*Future IT Innovation Laboratory, POSTECH, Pohang, Gyeongbuk, Korea
†Department of Electrical and Computer Engineering, Ajou University, Suwon, Gyeonggi, Korea
‡Department of Electrical Engineering, POSTECH, Pohang. Gyeongbuk, Korea
Email: jys900311@postech.ac.kr, snhong@ajou.ac.kr, nylee@postech.ac.kr

***Abstract*—This paper considers a multiple-input-multiple-output (MIMO) system with low-resolution analog-to-digital converters (ADCs). In this system, we propose a novel detection framework that performs data symbol detection without explicitly knowing channel state information at a receiver. The underlying idea of the proposed framework is to exploit supervised learning. Specifically, during channel training, the proposed approach sends a sequence of data symbols as pilots so that the receiver learns a nonlinear function that is determined by both a channel matrix and a quantization function of the ADCs. During data transmission, the receiver uses the learned nonlinear function to detect which data symbols were transmitted. In this context, we propose two blind detection methods to determine the nonlinear function from the training-data set. We also provide an analytical expression for the symbol-vector-error probability of the MIMO systems with one-bit ADCs when employing the proposed framework. Simulations demonstrate the performance improvement of the proposed framework compared to existing detection techniques.**

I. Introduction

Future wireless systems (e.g. mmWave systems) are possible to provide communication links with Gbps data rates by using wide bandwidths. As the bandwidth of the system increases, the sampling rate of an analog-to-digital converter (ADC) should increase linearly with the bandwidth. Unfortunately, the higher sampling rate gives rise to the larger power consumption of the ADC, and it turns out that the power consumption scales linearly with both precision level and sampling rate [1], [2]. Therefore, the use of the ADCs with a few-bit precision has been regarded as a cost-effective solution to reduce power consumption of devices that use high-speed ADCs. When low-resolution ADCs are used in communication systems, the characterization of the channel capacity is challenging, because a received signal is significantly distorted by nonlinear quantization operation at the ADCs. In a simple case in which one-bit ADCs are used in a single-input-single-output (SISO) system, the system capacity has been extensively analyzed [1]–[3]. For instance, it has shown in [3] that quadrature-phase-shift-keying (QPSK) modulation is optimal in an information-theoretical sense.

Recently, multiple-input-multiple-output (MIMO) systems with low-resolution ADCs have received great attentions because the spectral-efficiency limitation due to the quantization at the ADCs can be overcome by the use of multiple antennas [4]–[6]. When low-resolution ADCs are used in MIMO systems, however, the characterization of the channel capacity is much more complicated than in SISO systems, and the exact capacity expression is not known in general [4]. In addition to the capacity analysis, to realize the potential diversity and multiplexing gains of MIMO systems, numerous detection methods have also been proposed for the MIMO systems with the low-resolution ADCs [7]–[11]. For example, linear-detection methods have been developed for high-order-modulation input signals when a least-squares (LS) method is used to estimate the MIMO channel [7], and a near maximum-likelihood (ML) detector that uses a convex-optimization technique has been proposed for the MIMO systems that use one-bit ADCs [11].

Most existing MIMO detection techniques have been developed under the assumption of estimated or perfect channel state information at a receiver (CSIR), to perform coherent detection. In the MIMO system with the low-resolution ADCs, however, accurate CSIR cannot be readily obtained from pilot signals due to the coarse quantization that is applied when receiving signals. To resolve this problem, several channel-estimation methods have been developed to improve the accuracy of CSIR [11]–[13]. The common problem of the methods is that they require a huge amount of the pilot signals. Furthermore, at high SNR, the mean-square error of the LS channel-estimation method degrades significantly [13].

Blind MIMO communications have been studied by some researchers to provide the gains of MIMO systems, even when accurate CSIR is not available [14]–[18]. The capacity of blind MIMO communication systems with perfect (infinite precision) ADCs has been thoroughly analyzed on the basis of the Grassmanian manifold, where block-fading channels are assumed [14]. Data-detection techniques for blind MIMO communications have also been developed in [17], [18], but these techniques are only available for a particular type of modulation, e.g., space-shift-keying modulation [17] and phase-shift-keying (PSK) modulation [18]. The most relevant prior work that considers the use of low-resolution ADCs with the blind MIMO detection is a joint channel-and-data estimation method in [9] developed by using Bayesian inference theory. This method, however, has high implementation complexity, which may not be affordable in practical communication systems.

In this paper, we propose a novel blind detection frame-

978-1-4673-8999-0/17/$31.00 ©2017 IEEE

work for a MIMO system with low-resolution ADCs when explicit CSIR is not known to a receiver. The key idea of the proposed framework is to interpret a blind MIMO detection problem as a classification problem of supervised learning. The proposed framework consists of two phases: 1) The transmitter sends a sequence of data symbols so that the receiver learns a nonlinear function that describes input-output relations of the system, which is determined by both a channel matrix and a quantization function of the ADCs. 2) Using the learned nonlinear function, the receiver detects the transmitted data symbols. The proposed framework does not require an explicit channel-estimation process, and in the fist phase, pilot signals are substituted with data symbols, so this framework can be interpreted as a new type of a blind detection technique. Specifically, for the first phase, we introduce a channel-training method that learns an empirical conditional probability mass function (PMF) for each symbol vector. For the second phase, using this PMF, we develop two blind detection methods: Minimum-mean-distance detection (MCD) and minimum-center-distance detection (MCD). We also analyze the symbol-vector-error probability for the MIMO systems that use one-bit ADCs and binary-phase-shift-keying (BPSK) modulation. Particularly, by regarding all possible received vectors as codewords of a nonlinear code, we derive the upper bound of the symbol-vector-error probability for a fixed channel matrix, which is characterized in terms of the minimum Hamming distance $d_{\min}$ of the code. We evaluate the performance gains of the proposed blind detection techniques compared to existing detection techniques and show the validation of the analysis, by simulations.

II. System Model

In this section, we present a system model for a MIMO system with low-resolution ADCs.

A transmitter equipped with $N_{\rm t}$ transmit antennas intends to send $N_{\rm t}$ independent data symbols to a receiver equipped with $N_{\rm r}$ receive antennas, where $N_{\rm r} \geq N_{\rm t}$. Let $\mathbf{x}[n] = [x_1[n], x_2[n], \ldots, x_{N_{\rm t}}[n]]^\top \in \mathbb{C}^{N_{\rm t}}$ be the data symbol vector sent by the transmitter at time slot n. The received signal vector $\mathbf{r}[n] \in \mathbb{C}^{N_{\rm r}}$ at time slot n before the ADCs is

$$\mathbf{r}[n] = \mathbf{H}\mathbf{x}[n] + \mathbf{z}[n], \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{N_{\rm r}\times N_{\rm t}}$ denotes a channel matrix, and $\mathbf{z}[n] = [z_1[n], z_2[n], \ldots, z_{N_{\rm r}}[n]]^\top$ is a noise vector in which the element are assumed to be circularly-symmetric complex Gaussian noise with zero mean and variance σ^2, i.e., $z_i[n] \sim \mathcal{CN}(0, \sigma^2)$. Each data symbol $x_i[n]$ satisfies $\mathbb{E}[|x_i[n]|^2] = 1$ and is drawn from a constellation set $\mathcal{X}$ with constellation size $M = |\mathcal{X}|$. For instance, $\mathcal{X} = \{-1, +1\}$ for BPSK modulation. The SNR of the considered system is defined as $\rho = \frac{N_{\rm t}}{\sigma^2}$.

We assume a block-fading channel in which the channel remains constant for T time slots. A transmission frame that consists of T time slots is divided into 1) a channel-training phase and 2) a data-transmission phase. The first $T_{\rm t}$ time slots are allocated for the training phase, and the subsequent $T_{\rm d}$ time slots are allocated for the data-transmission phase, i.e., $T = T_{\rm t} + T_{\rm d}$. In this paper, we assume that the transmitter does not have channel state information and, more importantly, that the receiver does not have explicit knowledge of $\mathbf{H}$. Instead, the receiver learns the effective nonlinear channel function by sending a training sequence during the channel-training phase, as will be explained in Section IV-A.

Each antenna of the receiver is equipped with an RF chain followed by two low-resolution ADCs that are applied to real and imaginary parts of the received signal, respectively. Each ADC performs element-wise B-bit scalar quantization to the input signal. We adopt B-bit uniform scalar quantizer at the ADCs, defined as

$$Q_B(x) = \begin{cases} r_{\rm low} - 0.5\Delta, & x < r_{\rm low}, \\ r_{\rm low} + \left\lfloor \frac{1}{\Delta}(x - r_{\rm low}) \right\rfloor \Delta + 0.5\Delta, & r_{\rm low} \leq x < r_{\rm up}, \\ r_{\rm up} + 0.5\Delta, & r_{\rm up} \leq x, \end{cases} \tag{2}$$

where $r_{\rm low} = (-2^{B-1} + 1)\Delta$, $r_{\rm up} = (2^{B-1} - 1)\Delta$, and Δ is the step size of the uniform quantizer. The set of all possible outputs of this quantizer is

$$\mathcal{Y} = \left\{(-2^{B-1} + 0.5)\Delta, , \ldots, (2^{B-1} - 0.5)\Delta\right\}, \tag{3}$$

where $|\mathcal{Y}| = 2^B$. The total ADC process can be described using a function $Q : \mathbb{C}^{N_{\rm r}} \rightarrow \mathcal{Y}^{2N_{\rm r}}$ defined as

$$Q(\mathbf{x}) = [Q_B(x_{{\rm R},1}), Q_B(x_{{\rm R},2}), \ldots, Q_B(x_{{\rm R},2N_{\rm R}})]^\top, \tag{4}$$

where $x_{{\rm R},i}$ is the ith element of $\mathbf{x}_{\rm R}$, and $\mathbf{x}_{\rm R} = \left[{\rm Re}(\mathbf{x})^\top, {\rm Im}(\mathbf{x})^\top\right]^\top$. Using (4), the real-domain received signal vector $\mathbf{y}[n] \in \mathcal{Y}^{2N_{\rm r}}$ at time slot n after the ADCs is represented simply as $\mathbf{y}[n] = Q(\mathbf{r}[n])$.

III. Blind MIMO Detection Techniques

In this section, we propose blind detection techniques for MIMO systems with low-resolution ADCs, which essentially do not require explicit CSIR. The key idea of the proposed detection techniques is to interpret a blind MIMO detection problem as a classification problem of supervised learning. In the classification, to correctly determine the class label of every unlabeled data, a mapping function, called *classifier*, is designed by using a training set which provides a priori information of the relation between class labels and the data. Interestingly, this classifier resembles with a decoding function in wireless communications, which maps a received signal vector to the index of a symbol vector. Motivated by this similarity, we introduce a channel-training method that provides statistical information about the input-output relation of the system, then develop two data-detection methods by exploiting the obtained statistical information.

A. Channel-Training Method

The principal goal of the channel training is to provide the information about the input-output relation of the system, by using $T_{\rm t}$ input-output pairs for a given channel matrix $\mathbf{H}$. To achieve this goal, we consider repeated transmission of all possible symbol vectors during the channel training, so that the receiver can observe the multiple received signals for each symbol vector. Then we use these multiple observations to determine an empirical conditional distribution of the received signals for each symbol vector, which statistically characterizes the input-output relation. Details of the proposed channel-training method is given below.

When the transmitter equipped with $N_{\rm t}$ transmit antennas sends a data symbol vector using a constellation set $\mathcal{X}$, the set of all possible symbol vectors is denoted as $\mathcal{X}^{N_{\rm t}}$, and the number of possible symbol vectors is given by $K = M^{N_{\rm t}}$. Let $\mathbf{x}_k$ be the kth possible symbol vector and L be the number of repetitions of each input vector during training phase. Then the training-sequence matrix for the proposed channel-training method is given by

$$\begin{aligned}\mathbf{X}_{\rm t} &= [\mathbf{x}[1],\ \mathbf{x}[2], \ldots, \mathbf{x}[T_{\rm t}]] \\ &= \Big[\underbrace{\mathbf{x}_1, \ldots, \mathbf{x}_1}_{L \text{ repetition}}, \mathbf{x}_2, \ldots, \mathbf{x}_2, \ldots\ldots, \mathbf{x}_K, \ldots, \mathbf{x}_K\Big]. \quad (5)\end{aligned}$$

This training method requires a total of KL time slots for the period of training, i.e., $T_{\rm t} = KL$. We assume that the receiver has perfect knowledge of both $\mathcal{X}^{N_{\rm t}}$ and L or equivalently $\mathbf{X}_{\rm t}$. Under this premise, the receiver observes the set of the signal vectors associated with the kth symbol vector, denoted by $\mathcal{Y}_{{\rm t},k} = \big\{\mathbf{y}[n] \big| n \in \{(k-1)L+1, \ldots, kL\}\big\}$. Therefore, the set of received signals learned from the channel training is $\mathcal{Y}_{\rm t} = \cup_{k=1}^{K} \mathcal{Y}_{{\rm t},k} \subset \mathcal{Y}^{2N_{\rm r}}$.

In detection theory, a conditional probability function is a sufficient statistic to perform ML detection. Motivated by this, using $\mathcal{Y}_{{\rm t},k}$, the receiver creates an empirical conditional PMF for each symbol vector $\mathbf{x}_k$ as follows:

$$\hat{p}(\tilde{\mathbf{y}}|\mathbf{x}_k) = \frac{1}{L}\sum_{t=1}^{L} \mathbf{1}\left(\tilde{\mathbf{y}} = \mathbf{y}[(k-1)L+t]\right), \quad (6)$$

where $k \in \{1, 2, \ldots, K\}$ and $\mathbf{1}(A)$ is an indicator function that equals 1 if an event A is true and 0 otherwise.

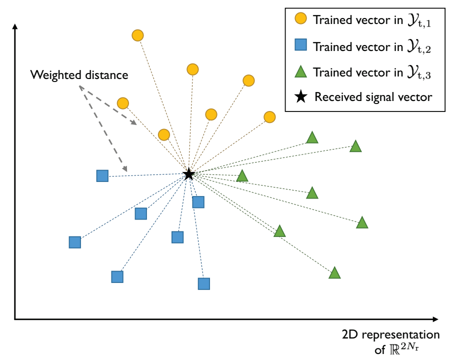
B. Data-Detection Methods

We describe two detection methods to determine a decoding (mapping) function $f : \mathcal{Y}^{2N_{\rm r}} \to \mathcal{K} = \{1, 2, \ldots, K\}$, by exploiting the empirical conditional PMF obtained during the training phase.

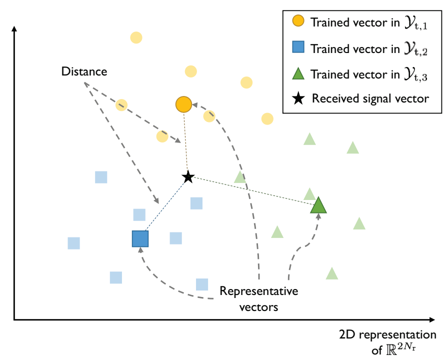
1) Minimum-Mean-Distance Detection (MMD): The key idea of MMD is to compare the mean of the distances from a received signal using the empirical PMFs. As illustrated in Fig. 1(a), the proposed MMD method, $f_{\rm MMD} : \mathcal{Y}^{2N_{\rm r}} \to \mathcal{K}$, selects the index of the symbol vector that yields the conditional minimum mean distance, i.e.,

$$\begin{aligned}f_{\rm MMD}(\mathbf{y}[n]) &= \operatorname*{argmin}_k \mathbb{E}\left[\|\mathbf{y}[n] - \mathbf{y}\|_2 \,|\mathbf{x} = \mathbf{x}_k\right] \\ &\triangleq \operatorname*{argmin}_k \sum_{\mathbf{y}\in\mathcal{Y}_{\rm t}} \|\mathbf{y}[n] - \mathbf{y}\|_2 \,\hat{p}\left(\mathbf{y}|\mathbf{x}_k\right). \quad (7)\end{aligned}$$

The conditional distance calculated from (7) may differ from the true conditional distance, $\mathbb{E}\left[\|\mathbf{y}[n] - \mathbf{y}\|_2 \,|\mathbf{x} = \mathbf{x}_k\right]$, because we use the empirical distribution for the calculation. This empirical distribution, however, converges to the corresponding true distribution when the number of repetitions for the channel training approaches infinity. Therefore, in this case, MMD indeed becomes the minimum-mean-distance criterion.



(a) Minimum-mean-distance detection (MMD)



(b) Minimum-center-distance detection (MCD)

Fig. 1. Illustration of the proposed detection methods when $K = 3$.

2) Minimum-Center-Distance Detection (MCD): One drawback of MMD is that it entails high computational complexity. The MMD method requires to compute all distances among the received signal and the vectors in $\mathcal{Y}_{\rm t}$. Particularly, when the size of $\mathcal{Y}_{\rm t}$ is large (e.g., $L \gg 1$), the complexity of this computation may not be acceptable for use in practical systems. Therefore, to reduce the detection complexity, we present a simple blind detection method, called MCD.

The key idea of MCD is to create a set of K representative vectors at the receiver for decoding, as depicted in Fig. 1(b). Specifically, during the training phase, the receiver has observed L output vectors $\mathcal{Y}_{{\rm t},k} = \{\mathbf{y}[(k-1)L+1], \ldots, \mathbf{y}[kL]\}$ for each symbol vector $\mathbf{x}_k$. We create a representative output vector for the kth input vector $\mathbf{x}_k$ by computing the empirical conditional expectation, i.e.,

$$\mathbb{E}[\mathbf{y}|\mathbf{x}_k] \triangleq \sum_{\mathbf{y}\in\mathcal{Y}_{{\rm t},k}} \mathbf{y}\hat{p}(\mathbf{y}|\mathbf{x}_k) = \frac{1}{L}\sum_{t=1}^{L} \mathbf{y}[(k-1)L+t]. \quad (8)$$

Notice that $\mathbb{E}[\mathbf{y}|\mathbf{x}_k]$ is not necessarily an element of $\mathcal{Y}^{2N_{\rm r}}$. Utilizing $\mathbb{E}[\mathbf{y}|\mathbf{x}_k]$, the proposed MCD method, $f_{\rm MCD} : \mathcal{Y}^{2N_{\rm r}} \to \mathcal{K}$, finds the index that minimizes the distance between $\mathbf{y}[n]$

and $\mathbb{E}[\mathbf{y}|\mathbf{x}_k]$ as follows:

$$f_{\mathrm{MCD}}(\mathbf{y}[n]) = \operatorname*{argmin}_k \|\mathbf{y}[n] - \mathbb{E}[\mathbf{y}|\mathbf{x}_k]\|_2 . \tag{9}$$

Connection to a nearest-centroid classifier (NCC): The principle of MCD is very close to that of NCC. NCC is a simple classifier that assigns the class label of an unlabeled observed vector by using the centroid vectors that represent their classes. Similarly, MCD determines the index of the detected symbol vector by using the conditional-mean vectors that represent the trained signals for each symbol vector. Therefore, MCD can be interpreted as the application of NCC to the detection problem in wireless communications.

IV. Analysis for MIMO systems With One-Bit ADCs

This section provides the analytical characterizations of the detection-error performance for MIMO systems with one-bit ADCs. Particularly, we present the analytical expression for the upper bound of the symbol-vector-error probability (SVEP) when the MCD method is applied. The following theorem is the main result of the analysis.

Theorem 1. *Consider MIMO systems with one-bit ADCs. When MCD is applied, the upper bound of the symbol-vector-error probability (SVEP) for high SNR is*

$$P_{\mathrm{e}}^{\mathrm{vec}} \le C_{N_r,D} \exp\left(-\frac{D\rho g_{\min}^2}{N_{\mathrm{t}}}\right), \tag{10}$$

where $C_{N_r,D} = \sum_{j=D}^{2N_{\mathrm{r}}} \binom{2N_{\mathrm{r}}}{j}$, $D = \left\lfloor \frac{d_{\min}+1}{2} \right\rfloor$, $d_{\min} = \min_{i\neq j} \|Q(\mathbf{Hx}_i) - Q(\mathbf{Hx}_j)\|_0$, $g_{\min} = \min_{(k,l)} |g_{k,l}|$, $g_{k,l}$ *is the* l*th element of* $\mathbf{g}_k$*, and* $\mathbf{g}_k = \left[\mathrm{Re}\left(\mathbf{Hx}_k\right)^\top, \ \mathrm{Im}\left(\mathbf{Hx}_k\right)^\top\right]^\top$.

Proof: In this proof, we omit the index n of time slot for ease of exposition. Suppose that the receiver equipped with one-bit ADCs adopts the MCD method. Then the receiver detects the symbol vector as $\hat{\mathbf{x}} = \mathbf{x}_{k^\star}$, where $k^\star = f_{\mathrm{MCD}}(\mathbf{y})$, and $\mathbf{y} = Q(\mathbf{Hx}+\mathbf{z}) \in \{-1,+1\}^{2N_{\mathrm{r}}}$. Let $P_{\mathrm{e},k}^{\mathrm{vec}} = \mathbb{P}\left(\hat{\mathbf{x}} \neq \mathbf{x}_k | \mathbf{x} = \mathbf{x}_k\right)$ be the pair-wise error probability that the detected symbol vector is different from $\mathbf{x}_k$ when the transmitter sends $\mathbf{x}_k$. Then SVEP is defined as

$$P_{\mathrm{e}}^{\mathrm{vec}} = \sum_{k=1}^{M^{N_{\mathrm{t}}}} \mathbb{P}\left(\hat{\mathbf{x}} \neq \mathbf{x}_k, \mathbf{x} = \mathbf{x}_k\right) = \frac{1}{M^{N_{\mathrm{t}}}} \sum_{k=1}^{M^{N_{\mathrm{t}}}} P_{\mathrm{e},k}^{\mathrm{vec}}. \tag{11}$$

Suppose that SNR is sufficiently large to satisfy the following:

$$\bar{\mathbf{y}}_{\mathrm{t},k} = \frac{1}{L}\sum_{t=1}^{L} Q(\mathbf{Hx}_k + \mathbf{z}[(k-1)L+t]) = Q(\mathbf{Hx}_k). \tag{12}$$

Then the detection rule of the MCD method in (9) with the one-bit ADCs is rewritten as

$$k^\star[n] = \operatorname*{argmin}_k \|\mathbf{y} - Q(\mathbf{Hx}_k)\|_2 \tag{13}$$

$$= \operatorname*{argmin}_k \|\mathbf{y} - Q(\mathbf{Hx}_k)\|_0 , \tag{14}$$

where the equality of (14) holds only for the one-bit-ADC case, and $\|\mathbf{a}\|_0$ is the zero norm that denotes the number of nonzero elements in a vector $\mathbf{a}$. From (14), $P_{\mathrm{e},i}^{\mathrm{vec}}$ of the MCD method is upper bounded as

$$P_{\mathrm{e},k}^{\mathrm{vec}} \le \mathbb{P}\Big(\|\mathbf{y}-Q(\mathbf{Hx}_k)\|_0 \ge \min_{j\neq k} \|\mathbf{y}-Q(\mathbf{Hx}_j)\|_0 \Big| \mathbf{x} = \mathbf{x}_k\Big). \tag{15}$$

For further analysis, we define a set $\mathcal{C} = \{Q(\mathbf{Hx}_1), Q(\mathbf{Hx}_2), \ldots, Q(\mathbf{Hx}_K)\}$. We interpret this set as an error-correcting code; each element $Q(\mathbf{Hx}_k)$ can be treated as a codeword vector of $\mathcal{C}$. For any code, one can define the distance between two codes $Q(\mathbf{Hx}_k)$ and $Q(\mathbf{Hx}_i)$ as

$$d_{k,i} = \|Q(\mathbf{Hx}_k) - Q(\mathbf{Hx}_i)\|_0 . \tag{16}$$

Then $\|\mathbf{y} - Q(\mathbf{Hx}_k)\|_0 \ge \left\lfloor \frac{d_{k,i}+1}{2} \right\rfloor$ is the necessary condition for an event that the MCD method outputs $\mathbf{x}_i$ when $\mathbf{x}_k$ was sent. Thus, we obtain an upper bound as

$$\begin{aligned}
P_{\mathrm{e},k}^{\mathrm{vec}} &\le \mathbb{P}\left(\bigcup_{i=1, i\neq k}^{K} \left\{ \|\mathbf{y} - Q(\mathbf{Hx}_k)\|_0 \ge \left\lfloor \frac{d_{k,i}+1}{2} \right\rfloor \right\} \Bigg| \mathbf{x} = \mathbf{x}_k \right) \\
&= \mathbb{P}\left(\|\mathbf{y} - Q(\mathbf{Hx}_k)\|_0 \ge D_k \big| \mathbf{x} = \mathbf{x}_k\right) \\
&= \mathbb{P}\left(\|Q(\mathbf{Hx}_k + \mathbf{z}) - Q(\mathbf{Hx}_k)\|_0 \ge D_k\right) \\
&= \mathbb{P}\left(\sum_{l=1}^{2N_{\mathrm{r}}} \mathbf{1}\left(\mathrm{sign}(g_{k,l} + z_{\mathrm{R},l}) \neq \mathrm{sign}(g_{k,l})\right) \ge D_k\right),
\end{aligned} \tag{17}$$

where $D_k = \min_{i\neq k} \left\lfloor \frac{d_{k,i}+1}{2} \right\rfloor$, $\mathrm{sign}(\cdot)$ is the signum function, $g_{k,l}$ is the lth element of $\mathbf{g}_k$, $z_{\mathrm{R},l}$ is the lth element of $\mathbf{z}_{\mathrm{R}}$,

$$\mathbf{g}_k = \begin{bmatrix} \mathrm{Re}\left(\mathbf{Hx}_k\right) \\ \mathrm{Im}\left(\mathbf{Hx}_k\right) \end{bmatrix}, \text{ and } \mathbf{z}_{\mathrm{R}} = \begin{bmatrix} \mathrm{Re}(\mathbf{z}) \\ \mathrm{Im}(\mathbf{z}) \end{bmatrix}. \tag{18}$$

From the fact that $z_{\mathrm{R},l}$ is i.i.d. as $\mathcal{N}(0, \frac{\sigma^2}{2})$ for all l, we have

$$\begin{aligned}
&\mathbb{P}\left(\mathrm{sign}(g_{k,l} + z_{\mathrm{R},l}) \neq \mathrm{sign}(g_{k,l})\right) \\
&= \left\{1 - \Phi\left(\sqrt{\frac{2\rho|g_{k,l}|^2}{N_{\mathrm{t}}}}\right)\right\} \triangleq P_{k,l}^{\mathrm{pair}},
\end{aligned} \tag{19}$$

where $\Phi(\cdot)$ is the cumulative distribution of the standard normal random variable. Let $\mathcal{S}_i \subset \{1,2,\ldots,2N_{\mathrm{r}}\}$ be the ith possible subset of $\{1,2,\ldots,2N_{\mathrm{r}}\}$ consists of D elements. By the above definitions, (17) can be rewritten as

$$\begin{aligned}
&\mathbb{P}\left(\sum_{l=1}^{2N_{\mathrm{r}}} \mathbf{1}\left(\mathrm{sign}(g_{k,l} + z_{\mathrm{R},l}) \neq \mathrm{sign}(g_{k,l})\right) \ge D_k\right) \\
&= \sum_{j=D_k}^{2N_{\mathrm{r}}} \sum_{i=1}^{\binom{2N_{\mathrm{r}}}{j}} \left\{ \prod_{l\in\mathcal{S}_i} P_{k,l}^{\mathrm{pair}} \prod_{l'\notin\mathcal{S}_i} (1 - P_{k,l'}^{\mathrm{pair}}) \right\}.
\end{aligned} \tag{20}$$

Aggregating the results in (11), (17), and (20) yields

$$P_{\mathrm{e}}^{\mathrm{vec}} \le \frac{1}{M^{N_{\mathrm{t}}}} \sum_{k=1}^{M^{N_{\mathrm{t}}}} \sum_{j=D_k}^{2N_{\mathrm{r}}} \sum_{i=1}^{\binom{2N_{\mathrm{r}}}{j}} \left\{ \prod_{l\in\mathcal{S}_i} P_{k,l}^{\mathrm{pair}} \prod_{l'\notin\mathcal{S}_i} (1 - P_{k,l'}^{\mathrm{pair}}) \right\}. \tag{21}$$

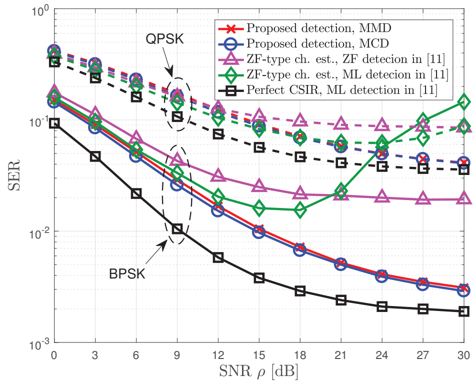


Fig. 2. SER vs. SNR of the proposed and conventional detection techniques for a MIMO system with one-bit ADCs, when $(N_\mathrm{t}, N_\mathrm{r}) = (2, 4)$. We set $(T, T_\mathrm{t}) = (500, 12)$ for BPSK and $(T, T_\mathrm{t}) = (500, 48)$ for QPSK. Channels are modeled as Rayleigh fading. This setting corresponds to the $L = 3$ case for the proposed channel-training method. We assume that the SNR during the training phase is set to be 3-dB higher than the SNR operating for the data transmission.

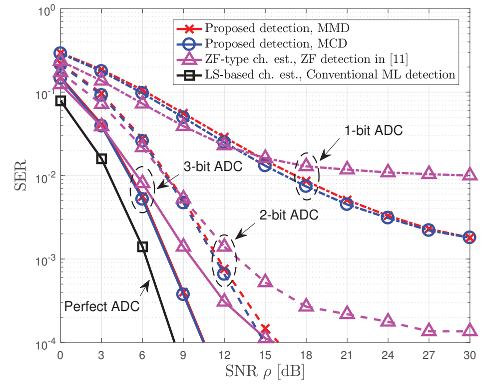


Fig. 3. SER vs. SNR of the proposed and conventional detection techniques for a MIMO system with various numbers of ADC bits and QPSK modulation, when $(N_\mathrm{t}, N_\mathrm{r}) = (2, 8)$. We set $T = 500$ and $T_\mathrm{t} = 48$. Channels are modeled as Rayleigh fading. We adopt the uniform quantizer defined in (2) with the step size of $\Delta = 0.5$. This setting corresponds to the $L = 3$ case for the proposed channel-training method. We assume that the SNR during the training phase is set to be 3-dB higher than the SNR operating for the data transmission.

Because $P_{k,l}^{\mathrm{pair}}$ is the decreasing function of $|g_{k,l}|$, (21) is further upper bounded as

$$P_\mathrm{e}^{\mathrm{vec}} \le \frac{1}{M^{N_\mathrm{t}}} \sum_{k=1}^{M^{N_\mathrm{t}}} C_{N_r, D_k} \prod_{l=l_{k,1}^\star}^{l_{k,D_k}^\star} P_{k,l}^{\mathrm{pair}} \tag{22}$$

$$\le C_{N_r,D} \left\{ 1 - \Phi\left(\sqrt{\frac{2\rho g_{\min}^2}{N_\mathrm{t}}} \right) \right\}^D \tag{23}$$

$$\le C_{N_r,D} \exp\left(-\frac{D\rho g_{\min}^2}{N_\mathrm{t}} \right), \tag{24}$$

where $D = \min_k D_k$, $C_{N_r,D} = \sum_{j=D}^{2N_\mathrm{r}} \binom{2N_\mathrm{r}}{j}$, $l_{k,i}^\star$ is the index of the element of $\mathbf{g}_k$ that has the ith-minimum absolute value, and $g_{\min} = \min_{(k,l)} |g_{k,l}|$. ∎

Proposition 1 demonstrates that the upper bound of the symbol-error probability decreases exponentially with SNR, ρ, the minimum channel gain, $g_{\min}^2$, the minimum distance, $D = \left\lfloor \frac{d_{\min}+1}{2} \right\rfloor$, and the inverse of N_t. The most interesting parameter here is $d_{\min}$ which is determined by the channel matrix $\mathbf{H}$ and the quantization function $Q(\cdot)$. Because small $d_{\min}$ significantly increases the detection-error probability, it is important to design the communication system with large enough minimum distance of the code $\mathcal{C} = \{Q(\mathbf{H}\mathbf{x}_1), Q(\mathbf{H}\mathbf{x}_2), \ldots, Q(\mathbf{H}\mathbf{x}_K)\}$. One simple way to enlarge the minimum distance is to increase the number of receive antennas N_r. For example, if $N_\mathrm{t} = 2$, we have four possible input vectors $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\} \in \{-1, +1\}^2$ and a code $\mathcal{C} = \{Q(\mathbf{H}\mathbf{x}_1), Q(\mathbf{H}\mathbf{x}_2), Q(\mathbf{H}\mathbf{x}_3), Q(\mathbf{H}\mathbf{x}_4)\} \in \{-1, +1\}^{2N_\mathrm{r}}$ for the communications. Clearly, the minimum distance $d_{\min}$ increases with N_r, because each codeword can be mapped into a higher-dimensional space. This characteristic can be interpreted as a *receive diversity gain* in the MIMO system with one-bit ADCs.

V. Numerical Results

In this section, we provide simulation results to evaluate the performance of the proposed blind detection techniques compared to existing detection techniques, for the MIMO system with low-resolution ADCs. We also validate the analysis given in Section IV, by simulations.

Fig. 2 shows the symbol-error-rates (SERs) of the proposed and conventional detection techniques for the MIMO systems with the one-bit ADCs. As SNR increases, the SERs of all proposed detection techniques converge to the SER lower bound achieved by the ML detection with perfect CSIR. Whereas, two conventional detection methods that are based on zero-forcing (ZF) type channel estimation show severe SER degradation in most cases. This is because the accuracy of CSIR obtained by the channel-estimation method is inaccurate when the pilot length is insufficient and the quantization error at the ADC is high. Particularly, the proposed techniques outperform the ML detection with the ZF-type channel estimation, except in the limited region of high SER (e.g., SER > 0.1).

Fig. 3 shows the SERs of the proposed and conventional detection techniques for MIMO systems with various numbers of ADC bits and QPSK modulation. It is observed that the proposed detection techniques provide a significant SER reduction as the number of ADC bits increases. Particularly, the ADC with 3-bit resolution is sufficient for the proposed techniques to achieve the SER near to the perfect-ADC case with LS channel estimation, whereas the ZF detection with the ZF-type channel estimation still shows an error floor.

In both figures 2 and 3, two proposed detection methods show very similar SER performance for all antenna configurations and SNRs. Surprisingly, MCD outperforms MMD even if it has the lowest detection complexity. This superiority is possible because the empirical conditional PMFs are unreliable when the repetition number at the channel training is insufficient ($L = 3$). Therefore, MCD is a more attractive detection

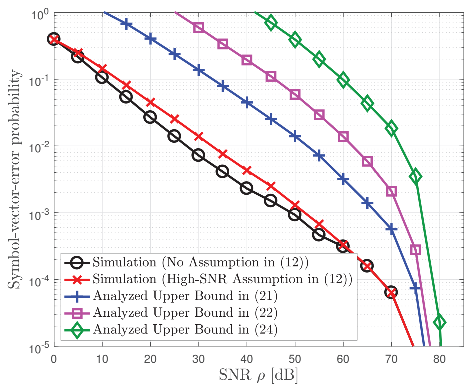


Fig. 4. Comparison among the analyzed upper bounds in Section IV and the symbol-vector-error probability of MCD when BPSK modulation is applied with $(N_{\rm t}, N_{\rm r}) = (2, 2)$, $D = 1$, and one-bit ADCs. We set $T_{\rm t} = 10$ and $T = 3 \times 10^4$. Simulation results are obtained by averaging out the error performance of 1000 random channels with Rayleigh fading, that have $D = 1$.

method than MMD, for use in practical MIMO systems with low-resolution ADCs.

Fig. 4 compares the SVEP performance of MCD with the analyzed upper bounds derived in Section IV, by simulations. Simulation results show that all SVEP curves obtained from the simulations are lower than the upper bounds derived in Section IV. The differences between the SVEP curves of the analysis and simulation decrease as SNR increases; this result implies that the analysis results well characterizes the high-SNR performance of the MCD method. Two simulated SVEP curves with and without the a high-SNR assumption in (12) only show a small difference. Therefore, in the considered scenario, the analysis results derived with the high-SNR assumption can also be used to characterize SVEP performance for all SNR values.

VI. Conclusion

In this paper, we have presented novel blind detection techniques for MIMO systems with low-resolution ADCs, inspired by supervised learning. The proposed techniques do not require explicit CSIR; instead they learn a nonlinear function that characterizes the input-output relation of the system, including the effects of the channel matrix and the quantization at the ADCs. For the one-bit ADCs and BPSK modulation, we have analyzed the symbol-vector-error probability of the proposed method, and have shown that the upper bound of this probability decreases exponentially with the minimum distance that can increase with the number of receive antennas. Simulations show that the proposed approach improves the SER and is also effective for use with the ADCs that have an arbitrary number of precision levels.

Acknowledgement

This work was supported by the research program with the Intel Labs, CA, USA. This work was also supported by the ICT R&D program (IITP-2017-B0717-17-0002) and the ICT consilience creative program (IITP-R0346-16-1007) of MSIP/IITP, Korea.

References

[1] B. Murmann, "ADC performance survey 1997-2016," [Online]. Available: http://web.stanford.edu/∼murmann/adcsurvey.html

[2] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.,* vol. 17, no. 4, pp. 539–550, Apr. 1999.

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

[4] J. Mo and R. W. Heath, Jr., "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.,* vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[5] N. Liang and W. Zhang, "Mixed-ADC massive MIMO," *IEEE J. Sel. Areas Commun.,* vol. 34, no. 4, pp. 983–997, Apr. 2016.

[6] C. Mollen, 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. (ICASSP),* Mar. 2016.

[7] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "One-bit massive MIMO: Channel estimation and high-order modulations," in *Proc. IEEE Int. Conf. Commun. (ICC),* June 2015.

[8] 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.

[9] C.-K. Wen, C.-J. Wang, S. Jin, K.-K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.,* vol. 64, no. 10, pp. 2541–2556, May 2016.

[10] S.-N. Hong, Y.-S. Jeon, and N. Lee, "MIMO systems with low-resolution ADCs: Linear coding approach," *submitted to IEEE Trans. Wireless Commun.,* 2016.

[11] 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," *IEEE Trans. Commun.,* vol. 64, no. 5, pp. 2005–2018, May 2016.

[12] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, Jr., "Channel estimation in millimeter wave MIMO systems with one-bit quantization," in *Proc. IEEE Asilomar Conf. Signals Syst. Comput.,* Nov. 2014, pp. 957–961.

[13] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," arXiv:1609.07427 [cs.IT], Sep. 2016. [Online]. Available: http://arxiv.org/abs/1609.07427

[14] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory,* vol. 48, no.2, pp. 359–383, Feb. 2002.

[15] M. Abuthinien, S. Chen, and L. Hanzo, "Semi-blind joint maximum likelihood channel estimation and data detection for MIMO Systems," *IEEE Signal Process. Lett.,* vol. 15, pp. 202–205, 2008.

[16] S. Chen, S. Sugiura, and L. Hanzo, "Semi-blind joint maximum likelihood channel estimation and data detection for MIMO Systems," *IEEE Signal Process. Lett.,* vol. 17, no. 12, pp. 993–996, Dec. 2010.

[17] H.-W. Liang, R. Y. Chang, W.-H. Chung, and S.-Y. Kuo, "A reduced-complexity blind detector for MIMO system using K-means clustering algorithm," in *Proc. IEEE Veh. Tech. Conf. (VTC) Spring,* June 2013.

[18] A. Schenk and R. F. H. Fischer, "Noncoherent detection in massive MIMO systems," in *Proc. Int. ITG Workshop Smart Antennas (WSA),* Mar. 2013.