# TDD-Sparse Index Multiple Access for Joint Data Detection and Device Identification in MTC

Jinho Choi

***Abstract:*** **In this paper, we study pre-equalization for sparse index multiple access (SIMA), which was proposed to enable one-shot identification and detection in random access, over frequency-selective fading channels in time division duplexing (TDD) mode. Each device is to decide a transmit vector to adapt to its channel for pre-equalization based on the channel reciprocity. Due to pre-equalization, the receiver does not need to estimate the channel state information (CSI) of active devices. We further modify pre-equalization with multiple precoding matrices for a better performance with a lower transmission power. For low signaling overhead in random access, we also consider one-shot identification and detection to perform active device identification and their signal detection simultaneously with a low-complexity algorithm by exploiting the sparsity of device activity as well as unique identification sequences.**

***Index Terms:*** **Compressive sensing, machine type communications, random access.**

## I. INTRODUCTION

MACHINE type communications (MTC) and the Internet of things (IoT) become more important in next generation wireless systems [1]–[3]. Usually, the characteristics of data traffics of MTC and the IoT are different from those in human type communications (HTC). For example, in MTC and the IoT, each device may have a short data packet of a few bytes to transmit to an access point or aggregation point (AP). In addition, there would be a number of devices that can connect to the AP with sparse activity. Thus, conventional coordinated multiple access schemes for HTC may be inefficient for MTC due to their high signaling overhead. On the other hand, random access could be a good choice for MTC as it has low signaling overhead. From this reason, for example, a random access scheme with multiple preambles is proposed in [2] for MTC within the long term evolution-advanced (LTE-A) system.

In random access, since a receiver does not know selected preambles by active devices in advance, it is difficult to employ conventional multiple signal detection or multiuser detection (MUD) approaches that require to know the set of selected preambles unless all possible combinations of selected preambles are considered (which results in a prohibitively high computational complexity in general). However, by exploiting the sparsity of activity, low-complexity MUD algorithms can be de-

rived as in [4]–[8], based on the notion of compressive sensing (CS) [9]–[12], which suit to the AP for MTC where there are a number of devices of sparse activity.

It is noteworthy that unlike the RACH procedure in [2], compressive random access is a "one-shot" transmission scheme that can transmit a short message from an active device without any handshaking process [7], [8], [13]. That is, an active device transmits a packet that includes pilot and data blocks (as one-shot transmission) without any connection request and channel allocation. While one-shot transmission is useful for very short packets (of few bytes) due to low signaling overhead [14], it would be impractical for long packets. Thus, in general, compressive random access would be suitable for MTC where devices transmit short messages of few bytes.

In this paper, we consider a random access scheme based on sparse index multiple access (SIMA), which has been proposed in [15] and [16], as a one-shot scheme in time-division duplexing (TDD) mode. In SIMA, exploiting the sparsity of signature vectors as well as the sparsity of device activity, low-complexity MUD can be performed with device identification. Thus, in SIMA, no explicit transmissions of device identification attributes are required. The proposed random access based on SIMA can enjoy the same advantage and becomes suitable for MTC where devices have short packets of few bytes. In addition, due the the channel reciprocity in TDD, it is possible to employ pre-equalization at each device, which eliminates the need of the channel estimation at the AP.

The rest of the paper is organized as follows. In Section II, we present the system model for random access based on a multicarrier system with unique sparse identification sequences for devices. In Section III, we study pre-equalization to exploit the channel reciprocity so that the AP does not need to estimate channels. For better pre-equalization, we also consider multiple precoding matrices in Section III. In Section IV, we first derive a sparse model with pre-equalization and introduce a low-complexity two-step approach for the identification of active devices using a CS algorithm and a subspace method (for the first and second steps, respectively) to exploit the sparsity of both device activity and sparse identification sequences for devices. We present simulation results in Section V and conclude the paper with some remarks in Section VI.

*Notation*: Matrices and vectors are denoted by upper- and lower-case boldface letters, respectively. The superscripts T and H denote the transpose and complex conjugate, respectively. The $p$-norm of a vector $\mathbf{a}$ is denoted by $||\mathbf{a}||_p$ (If $p = 2$, the norm is denoted by $||\mathbf{a}||$ without the subscript). For a vector $\mathbf{a}$, $\mathrm{diag}(\mathbf{a})$ is the diagonal matrix with the diagonal elements from $\mathbf{a}$. For a matrix $\mathbf{X}$ (a vector $\mathbf{a}$), $[\mathbf{X}]_n$ ($[\mathbf{a}]_n$) represents the $n$th column (element, resp.). If $n$ is a set of indices, $[\mathbf{X}]_n$ is a subma-

trix of $\mathbf{X}$ obtained by taking the corresponding columns. $\mathbb{R}^n$ and $\mathbb{C}^n$ represent the $n$-dimensional real and complex vector spaces, respectively. $\mathbb{E}[\cdot]$ and $\mathrm{Var}(\cdot)$ denote the statistical expectation and variance, respectively. $\mathcal{CN}(\mathbf{a}, \mathbf{R})$ ($\mathcal{N}(\mathbf{a}, \mathbf{R})$) represents the distribution of circularly symmetric complex Gaussian (CSCG) (resp., real-valued Gaussian) random vectors with mean vector $\mathbf{a}$ and covariance matrix $\mathbf{R}$.

## II. SYSTEM MODEL

In this section, we present a multicarrier or orthogonal frequency division multiplexing (OFDM) system for random access and review sparse index multiple access (SIMA) that was proposed in [16].

### A. Random Access in a Multicarrier System

Throughout this paper, we consider a multicarrier system that consists of an AP and $K$ devices with $L$ subcarriers for random access. It is assumed that $K \gg L$, while only a fraction of devices become active to transmit their signals to the AP. For convenience, we denote by $M$ the number of active devices and assume that $M \leq L$. In addition, denote by $\mathcal{K}$ the index set of active devices. Clearly, we have $|\mathcal{K}| = M$. Note that the number of subcarriers is not large enough to assign a unique subcarrier to each device in the system, but is larger than or equal to the number of active devices.

We assume that the AP sends a beacon signal to devices so that devices with data can transmit signals within a slot of $T$ symbols. Let $r_{l;t}$ denote the received signal at the AP through subcarrier $l$ at time $t$ in a slot. Then, it can be shown that

$$\mathbf{r}_t = [r_{0;t}, \cdots, r_{L-1;t}]^{\mathrm{T}}$$
$$= \sum_{k \in \mathcal{K}} \mathbf{H}_k \mathbf{b}_{k;t} + \mathbf{n}_t, \ t = 0, \cdots, T-1, \quad (1)$$

where $\mathbf{H}_k$ and $\mathbf{b}_{k;t}$ are the frequency-domain channel matrix and the $t$th transmitted signal vector or OFDM symbol from device $k \in \mathcal{K}$ over $L$ subcarriers, respectively, and $\mathbf{n}_t = [n_0, \cdots, n_{L-1}]^{\mathrm{T}} \sim \mathcal{CN}(0, N_0 \mathbf{I})$ is the background noise. Note that $\mathbf{r}_t$ in (1) is the received signal from $|\mathcal{K}| = M$ active devices. The frequency-domain channel matrix over a frequency-selective fading is given by

$$\mathbf{H}_k = \mathrm{diag}(H_{k,0}, \cdots, H_{k,L-1}), \quad (2)$$

where $H_{k,l} = \sum_{p=0}^{\nu_k - 1} h_{k,p} e^{-\frac{j 2\pi p l}{L}}$. Here, $\{h_{k,p}\}$ is the channel impulse response (CIR) from device $k$ to the AP, $\nu_k$ is the length of CIR, and $j = \sqrt{-1}$. For convenience, we assume that $\nu_k = \nu$ for all $k$.

### B. SIMA

SIMA is a multicarrier system based multiple access scheme where a unique set of a small number of subcarriers is assigned to each device for channel access, in which by exploiting the sparsity of active subcarriers, it is possible to use CS algorithms for low-complexity receivers [16]. In SIMA, each device has a unique sparse identification vector (SIV), which is denoted by $\mathbf{q}_k \in \mathbb{C}^L$ for device $k$. This is different from the conventional

random access scheme (e.g., [2]), where each device can choose one preamble from a pool of preambles. The main advantages of SIMA over conventional random access are that there are no preamble collisions[1] due to a unique SIV for each device and active device identification can be readily carried out by estimating SIV.

In SIMA, the transmitted signal vector from an active device is modulated by its SIV as follows:

$$\mathbf{b}_{k;t} = \mathbf{q}_k s_{k;t},$$

where $s_{k;t} \in \mathcal{S}$ is the $t$th data symbol. Here, $\mathcal{S}$ is the signal constellation. We assume that $\mathbb{E}[s_{k;t}] = 0$ and $\mathbb{E}[|s_{k;t}|^2] = 1$ for normalization purposes. In addition, the $s_{k;t}$'s are independent and real-valued throughout the paper (e.g., $s_{k;t} \in \mathcal{S} = \{-1, +1\}$). Thus, we have

$$\mathbb{E}[s_{k;t} s_{k';t'}] = \delta_{k,k'} \delta_{t,t'}, \quad (3)$$

where $\delta_{l,m}$ represents the Kronecker delta.

Note that the real-valued version of $\mathbf{q}_k$ can be written as

$$\tilde{\mathbf{q}}_k = \begin{bmatrix} \Re(\mathbf{q}_k) \\ \Im(\mathbf{q}_k) \end{bmatrix}. \quad (4)$$

We assume that the number of non-zero elements of $\tilde{\mathbf{q}}_k$ is $Q$, i.e., $\tilde{\mathbf{q}}_k$ is $Q$-sparse, while $\tilde{\mathbf{q}}_k \in \mathbb{R}^{2L}$. For simplicity, we assume that the elements of $\tilde{\mathbf{q}}_k$ are either 0 or 1. That is, $\tilde{\mathbf{q}}_k$ is assumed to be a binary $Q$-sparse vector of length $2L$. For convenience, denote by $\mathcal{Q}$ the set of all the binary $Q$-sparse vectors of length $2L$. The number of all the possible SIVs or $|\mathcal{Q}|$ becomes $N_I = |\mathcal{Q}| = \binom{2L}{Q}$. Throughout the paper, we assume that $Q \ll L$ and $L$ is sufficiently large so that there could be a number of unique identification sequences available for many devices. For example, if $L = 512$ and $Q = 2$, we have $N_I \approx 5 \times 10^6$.

## III. PRE-EQUALIZATION FOR SIMA IN TDD

In this section, we consider pre-equalization at active devices based on the channel reciprocity in TDD mode for SIMA.

### A. Pre-Equalization in TDD Mode

Pilot symbols need to be transmitted with data symbols to allow the AP to estimate channels as in [7]. Unfortunately, in SIMA, this approach is not suitable as the complexity of the channel estimation can be prohibitively high, because the size of the measurement matrix becomes too big (in this case, the number of columns of the measurement matrix, $[\mathbf{q}_1, \cdots, \mathbf{q}_K]$, is $K$, which can be $N_I$). In order to avoid this problem, in this section, we consider pre-equalization in TDD mode [17]–[19], which can eliminate the need of the channel estimation at the AP by exploiting the channel reciprocity.

We consider two phases for random access in TDD mode. In the first phase, the AP broadcasts pilot symbols (as part of a beacon signal) so that each device can not only synchronize to the AP, but also estimate its (downlink) channel. Due to the

---

[1] In conventional random access where devices do not have unique preambles, preamble collisions happen if multiple devices choose a common preamble simultaneously.

channel reciprocity in TDD, the downlink channel becomes that for uplink transmissions. In the second phase, devices with data become active and transmit signals to the AP. The resulting received signal at the AP is that in (1). Throughout the paper, we assume that each device knows its channel matrix $\mathbf{H}_k$ (i.e., perfect channel state information (CSI) at transmitter[2] is assumed).

With one-tap pre-equalization in the frequency domain, the signal vector to be transmitted from device $k$, i.e., $\mathbf{b}_{k;t}$ in (1), if $k \in \mathcal{K}$ (or device $k$ is active) can be given by

$$\mathbf{b}_{k;t} = \mathbf{A}_k \mathbf{H}_k^{\mathrm{H}} \mathbf{q}_k s_{k;t}, \qquad (5)$$

where $\mathbf{A}_k = \mathrm{diag}(a_{k,0}, \cdots, a_{k,L-1})$ with $a_{k,l} \geq 0$ is the gain matrix that is used to allocate transmission powers over $L$ subcarriers.

Note that in (5), $\mathbf{H}_k^{\mathrm{H}}$ is pre-multiplied to $\mathbf{q}_k s_{k;t}$ for pre-equalization at each active device. Thus, at the AP, the channel estimation is not required, which is a salient feature of TDD-based multiple access schemes [17]–[19]. Substituting (5) into (1), the received signal vector at the AP is written as

$$\mathbf{r}_t = \sum_{k \in \mathcal{K}} \mathbf{H}_k \mathbf{A}_k \mathbf{H}_k^{\mathrm{H}} \mathbf{q}_k s_{k;t} + \mathbf{n}_t. \qquad (6)$$

Suppose that $\mathbf{A}_k$ can be decided to hold the following relation for pre-equalization:

$$\mathbf{H}_k \mathbf{A}_k \mathbf{H}_k^{\mathrm{H}} \mathbf{q}_k = c_k \mathbf{q}_k, \qquad (7)$$

where $c_k > 0$ is constant. From (7), we have

$$\mathbf{r}_t = \sum_{k \in \mathcal{K}} c_k \mathbf{q}_k s_{k;t} + \mathbf{n}_t. \qquad (8)$$

If the AP is able to estimate $\mathbf{q}_k$ and $s_{k;t}$ of the active devices in (8), the AP can not only identify the active devices, but also detect the data symbols transmitted by the active devices, simultaneously.

Note that we may consider one-shot identification/detection using a low-complexity CS algorithm to exploit the sparsity of active devices in (8) as in [7], [8]. Unfortunately, this may not be computationally feasible[3] as the measurement matrix, $[\mathbf{q}_1, \cdots, \mathbf{q}_K]$ is too big.

For convenience, we often use a real-valued representation of (8), which is given by

$$\tilde{\mathbf{r}}_t = \begin{bmatrix} \Re(\mathbf{r}_t) \\ \Im(\mathbf{r}_t) \end{bmatrix} = \sum_{k \in \mathcal{K}} c_k \tilde{\mathbf{q}}_k s_{k;t} + \tilde{\mathbf{n}}_t, \qquad (9)$$

where $\tilde{\mathbf{n}}_t = \begin{bmatrix} \Re(\mathbf{n}_t) \\ \Im(\mathbf{n}_t) \end{bmatrix} \sim \mathcal{N}\left(0, \frac{N_0}{2}\mathbf{I}\right)$.

---

[2]In practice, however, there might be channel estimation errors due to the background noise and other channel impairment. The impact of imperfect CSI on the performance would be studied in the future as a further research topic.

[3]If $K = N_I$, we have $K \approx 5 \times 10^6$ for $N = 512$ and $Q = 2$ as shown earlier. Thus, the complexity can be too high even if a low-complexity CS algorithm, e.g., the orthogonal matching pursuit (OMP) algorithm, is used as its complexity is proportional to $K$.

## B. Pre-Equalization with Multiple Precoding Matrices

For pre-equalization, we have considered $\mathbf{A}_k$ that satisfies (7) with $c_k > 0$. Unfortunately, if some frequency-domain channel coefficients are zeros due to deep fading and the values of $\mathbf{q}_k$ associated with those coefficients are not zeros, we may not be able to find any $\mathbf{A}_k$ satisfying (7). In addition, if some of $\{H_{k,l}\}$ are not zero, but small, then the corresponding diagonal elements of $\mathbf{A}_k$ might be large, which results in a high transmission power. In order to avoid this difficulty, in this subsection, we consider pre-equalization with multiple precoding matrices. For convenience, in this section, we omit the time index $t$.

Suppose that there are $J$ pre-determined matrices for precoding, which are given by

$$\mathcal{P} = \{\mathbf{\Psi}_1, \cdots, \mathbf{\Psi}_J\}. \qquad (10)$$

Each device can choose one of the $J$ precoding matrices in (10) to minimize the transmission power for pre-equalization. Suppose that $\mathbf{\Psi}_j$ is chosen. Let $\mathbf{x}_j = \mathbf{\Psi}_j \mathbf{q}_k$, $\mathbf{H} = \mathbf{H}_k$, and $\mathbf{A} = \mathbf{A}_k$ for notational convenience (i.e., we omit the device index $k$). Here, $\mathbf{x}_j$ is the precoded SIV. For convenience, let

$$c_k = 1 \text{ for all } k$$

in the rest of the paper. Then, (7) with precoding becomes

$$\mathbf{HAH}^{\mathrm{H}} \mathbf{x}_j = \mathbf{x}_j. \qquad (11)$$

For convenience, define $\mathbf{a} = [a_0, \cdots, a_{L-1}]^{\mathrm{T}}$, where $a_l$ denotes the $l$th diagonal element of $\mathbf{A}$, i.e., $\mathbf{A} = \mathrm{diag}(a_0, \cdots, a_{L-1})$. For given $\mathbf{A}$, since the transmitted signal becomes $\mathbf{b}_j = \mathbf{AH}^{\mathrm{H}} \mathbf{x}_j s$, the transmission power is given by

$$P_j = \mathbb{E}[||\mathbf{b}_j||^2] = \sum_{l=0}^{L-1} \beta_{j,l} |H_l|^2 a_l^2, \qquad (12)$$

where $\beta_{j,l} = |[\mathbf{x}_j]_l|^2$. In general, the transmission power of device is limited. Thus, to determine $\mathbf{A}$ that satisfies (11) for pre-equalization, we can consider the following constrained minimization of the sum of the squared errors (SSE):

$$\begin{array}{c} \min_{\mathbf{a}} C_j(\mathbf{a}) \\ \text{subject to } P_j \leq P_{\max}, \end{array} \qquad (13)$$

where $C_j(\mathbf{a}) = ||\mathbf{HAH}^{\mathrm{H}} \mathbf{x}_j - \mathbf{x}_j||^2$ and $P_{\max}$ is the maximum transmission power.

**Theorem 1:** The optimal solution to (13) is given by

$$\hat{a}_{j,l}(\lambda) = \frac{1}{|H_l|^2 + \lambda}, \qquad (14)$$

where $\lambda$ is the Lagrange multiplier, which is decided to satisfy the following:

$$P_j = \sum_{l=0}^{L-1} \frac{\beta_{j,l} |H_l|^2}{(|H_l|^2 + \lambda)^2} \leq P_{\max}. \qquad (15)$$

**Proof:** See Appendix A.

Since $P_j = \sum_{l=0}^{L-1} \frac{\beta_{j,l} |H_l|^2}{(|H_l|^2 + \lambda)^2}$ is a decreasing function of $\lambda$, it is possible to find the value of $\lambda$ satisfying (15), denoted by $\lambda_j$,

using a line search method (e.g., the bisection method) [20]. Using $\hat{\mathbf{a}}_j = [\hat{a}_{j,0}(\lambda_j), \cdots, \hat{a}_{j,L-1}(\lambda_j)]^T$, the best precoding matrix from $\mathcal{P}$ is chosen as follows:

$$\hat{j} = \operatorname*{argmin}_{j \in \{1, \cdots, J\}} C_j(\hat{\mathbf{a}}_j). \tag{16}$$

That is, from (16), we can find the precoding matrix that can have the minimum SSE in (13) for (11) with the transmission power constraint, $P_j \leq P_{\max}$. For convenience, the resulting SIMA with pre-equalization in TDD mode is referred to as TDD-SIMA.

Note that the computational complexity to find the optimal precoding matrix depends on the complexity to find $\mathbf{x}_j = \boldsymbol{\Psi}_j \mathbf{q}_k$. Since $\tilde{\mathbf{q}}_k$ is sparse and binary, the complexity to find $\mathbf{x}_j$ is low as no multiplications are required (the real-valued $\mathbf{x}_j$ is a sum of $Q$ column vectors of real-valued $\boldsymbol{\Psi}_j$).

*C. Limitations of TDD-SIMA*

While the TDD-SIMA scheme can be effectively used for MTC with a number of devices due to the availability of a number of unique SIVs, it has some limitations. For example, the channel reciprocity plays a crucial role in reducing the transmission power and pre-equalization. Thus, if no channel reciprocity is assumed, we need to consider the approaches in [7], [8]. As a compressive random access scheme, the sparsity of active devices is important to use low-complexity CS algorithms for MUD. Thus, the activity of devices has to be low. Otherwise, different approaches based on scheduling might be desirable.

## IV. ONE-SHOT IDENTIFICATION AND DETECTION

In this section, by exploiting the sparsity of both device activity and SIVs, we consider a low-complexity two-step approach for one-shot identification/detection in TDD-SIMA.

*A. Sparse Representation*

In this subsection, we present a sparse representation for $\tilde{\mathbf{r}}_t$ (19), which plays a key role in deriving a CS algorithm for active device identification by exploiting the sparsity of both device activity and SIVs.

For convenience, let $i(k)$ denote the index of the optimal precoding matrix that is selected by device $k \in \mathcal{K}$, i.e., $i(k) = \hat{j}$ as in (16). In addition, denote by $\mathcal{K}_i$ the index set of the active devices that choose $\boldsymbol{\Psi}_i$ for notational convenience. Then, we have

$$\cup_{i=1}^J \mathcal{K}_i = \mathcal{K} \text{ and } \mathcal{K}_i \cap \mathcal{K}_{i'} = \emptyset, \ i \neq i'.$$

Suppose that the pre-equalization condition in (11) is satisfied for all active devices. Then, the received signal vector at the AP is given by

$$\mathbf{r}_t = \sum_{k \in \mathcal{K}} \boldsymbol{\Psi}_{i(k)} \mathbf{q}_k s_{k;t} + \mathbf{n}_t = \sum_{i=1}^J \boldsymbol{\Psi}_i \mathbf{z}_{i;t} + \mathbf{n}_t$$
$$= [\boldsymbol{\Psi}_1, \cdots, \boldsymbol{\Psi}_J][\mathbf{z}_{1;t}^T, \cdots, \mathbf{z}_{J;t}^T]^T + \mathbf{n}_t, \tag{17}$$

where

$$\mathbf{z}_{i;t} = \sum_{k \in \mathcal{K}_i} \mathbf{q}_k s_{k;t}. \tag{18}$$

Note that in some cases, (11) may not hold. In this case, the right-hand-side (RHS) term in (17) would still be a good approximation due to the optimal selection of the precoding matrices in (16) with the optimal gain allocation through (13).

The real-valued vector of $\mathbf{r}_t$ can be given by

$$\tilde{\mathbf{r}}_t = \sum_{i=1}^J \tilde{\boldsymbol{\Psi}}_i \tilde{\mathbf{z}}_{i;t} + \tilde{\mathbf{n}}_t, \tag{19}$$

where $\tilde{\boldsymbol{\Psi}}_i = \begin{bmatrix} \Re(\boldsymbol{\Psi}_i) & -\Im(\boldsymbol{\Psi}_i) \\ \Im(\boldsymbol{\Psi}_i) & \Re(\boldsymbol{\Psi}_i) \end{bmatrix}$ and $\tilde{\mathbf{z}}_{i;t} = \begin{bmatrix} \Re(\mathbf{z}_{i;t}) \\ \Im(\mathbf{z}_{i;t}) \end{bmatrix}$. Letting $\tilde{\boldsymbol{\Psi}} = [\tilde{\boldsymbol{\Psi}}_1, \cdots, \tilde{\boldsymbol{\Psi}}_J]$ and $\tilde{\mathbf{z}} = [\tilde{\mathbf{z}}_{1;t}^T, \cdots, \tilde{\mathbf{z}}_{J;t}^T]^T$, we have

$$\tilde{\mathbf{r}}_t = \tilde{\boldsymbol{\Psi}} \tilde{\mathbf{z}}_t + \tilde{\mathbf{n}}_t. \tag{20}$$

The sizes of the matrix $\tilde{\boldsymbol{\Psi}}$ and the vector $\tilde{\mathbf{z}}$ are $2L \times 2JL$ and $2JL \times 1$, respectively. Since the number of non-zero elements of $\tilde{\mathbf{z}}_{i;t}$ is at most $|\mathcal{K}_i|Q$, the number of non-zero elements of $\tilde{\mathbf{z}}_t$ is at most $|\mathcal{K}|Q = MQ$. That is, $\tilde{\mathbf{z}}_t \in \Sigma_{MQ}$. Thus, if $MQ \ll 2JL$, we can see that $\tilde{\mathbf{z}}_t$ is sparse and (20) is a sparse representation with the measurement matrix $\tilde{\boldsymbol{\Psi}}$ of size $2L \times 2JL$.

In general, it is desirable that each precoding matrix is independently and uniformly chosen by devices so that each $\mathbf{z}_{i;t}$ is a superposition of a small number of SIVs for efficient identification of active devices.

*B. A Two-Step Approach for Identification*

Prior to deriving a low-complexity CS based two-step approach for identification, we briefly discuss the maximum likelihood (ML) approach for an optimal performance. Let $f(\{\mathbf{r}_t\} | \mathcal{K}, \{s_{k;t}, k \in \mathcal{K}\})$ denote the likelihood function of $\mathcal{K}$ and $s_{k;t}$, $k \in \mathcal{K}$. Then, from (17), the ML approach to find the set of active devices is given by

$$\mathcal{K}^* = \operatorname*{argmax}_{\mathcal{K}} \max_{s_{k;t} \in \mathcal{S}, k \in \mathcal{K}} \prod_t f(\mathbf{r}_t | \mathcal{K}, \{s_{k;t}, k \in \mathcal{K}\})$$
$$= \operatorname*{argmin}_{\mathcal{K}} \min_{s_{k;t} \in \mathcal{S}, k \in \mathcal{K}} \sum_t \left|\left| \mathbf{r}_t - \sum_{i=1}^J \sum_{k \in \mathcal{K}_i} \boldsymbol{\Psi}_i \mathbf{q}_k s_{k;t} \right|\right|^2. \tag{21}$$

Unfortunately, although the channel estimation is not required, the ML approach is computationally prohibitive as the number of the possible $\mathcal{K}$'s is $\binom{N_I}{M}$ when an exhaustive search is used. To avoid this problem, we can rely on the notion of CS [10], [12], [21] by exploiting the sparsity of both active devices and SIVs.

In the proposed two-step approach, we first estimate the support of $\tilde{\mathbf{z}}_t$ (this is the first[4] step). Then, from the estimated support of $\tilde{\mathbf{z}}_t$, we find the SIVs of active devices, $\mathbf{q}_k$, $k \in \mathcal{K}$ using a subspace method (this is the second step).

In the first step, noting that the support of $\tilde{\mathbf{z}}_{i;t}$ or $\tilde{\mathbf{z}}_t$ is the same for all $t$ in (18), we need to recover the common support. Denote by $\mathcal{J}$ the common support of $\tilde{\mathbf{z}}_t$, i.e.,

$$\mathcal{J} = \operatorname{supp}(\tilde{\mathbf{z}}_t), \ t = 0, \cdots, T-1.$$

---

[4]Note that the first step is not sufficient to estimate all $\mathbf{q}_k$'s of active devices as will be explained later.

Thus, the estimation of $\mathcal{J}$ can be seen as the sparse signal recovery (of sparsity $MQ$) from multiple measurement vectors (MMVs) [22], [23] with the measurement matrix $\tilde{\boldsymbol{\Psi}}$. From $\{\tilde{\mathbf{r}}_t\}$, for example, we can use the OMP algorithm for MMV (OMP-MMV) [22] to estimate $\mathcal{J}$.

For notational convenience in the second step, let $M_i = |\mathcal{K}_i|$ and $\mathcal{K}_i = \{k_i(1), \cdots, k_i(M_i)\}$. That is, $k_i(m)$ denotes the index of the $m$th active device among those choosing the $i$th precoding matrix. In the second-step, we carry out the SIV estimation with $\mathcal{J}$. For convenience, let

$$
\begin{aligned}
\mathcal{J}_i &= \text{supp}(\mathbf{z}_{i;t}) \\
&= \bigcup_{k \in \mathcal{K}_i} \text{supp}(\mathbf{q}_k), \ t = 0, \cdots, T-1,
\end{aligned} \tag{22}
$$

which is the joint support of the SIVs that choose $\boldsymbol{\Psi}_i$. Since

$$
\mathcal{J}_i = \mathcal{J} \cap \{2L(i-1), \cdots, 2Li - 1\} - 2Li, \tag{23}
$$

we can find the $\mathcal{J}_i$'s from $\mathcal{J}$. If there are no active devices that choose $\boldsymbol{\Psi}_i$, $\mathcal{J}_i$ becomes the empty set. If only one active device chooses $\boldsymbol{\Psi}_i$ (i.e., $M_i = |\mathcal{K}_i| = 1$), then $|\mathcal{J}_i| = Q$. In this case, from $\mathcal{J}_i$, the SIV of the active device that chooses $\boldsymbol{\Psi}_i$ can be easily estimated, and the corresponding active device is readily identified. However, if $|\mathcal{J}_i| > Q$, there are multiple active devices that choose $\boldsymbol{\Psi}_i$ and we need to find $M_i$ SIVs from $\mathcal{J}_i$. For example, suppose that $\mathcal{J}_i = \{0, 3, 4, 5\}$ when $2L = 6$ and $Q = 2$. Since $|\mathcal{J}_i| = 4$, there might be two active devices, say $k_i(1) = 1$ and $k_i(2) = 2$, that choose $\boldsymbol{\Psi}_i$ and the supports of their SIVs would be one of the following 6 possible pairs:

$$
\begin{aligned}
(\text{supp}(\tilde{\mathbf{q}}_1), \text{supp}(\tilde{\mathbf{q}}_2)) &= (\{0,3\}, \{4,5\}), (\{0,4\}, \{3,5\}), \\
&\quad (\{0,5\}, \{3,4\}), (\{3,4\}, \{0,5\}), \\
&\quad (\{3,5\}, \{0,5\}), \text{ or } (\{4,5\}, \{0,3\}).
\end{aligned}
$$

Clearly, we need to resolve the ambiguity in this case (i.e., the case of $M_i > 1$), which is the reason to consider the following subspace method for the second step.

**Theorem 2:** Consider the eigenvalue decomposition of the correlation matrix of $\tilde{\mathbf{r}}_t$ as follows:

$$
\mathbf{R}_{\tilde{\mathbf{r}}} = \mathbb{E}\left[\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}_t^{\text{T}}\right] = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^{\text{T}}. \tag{24}
$$

where $\mathbf{E} = [\mathbf{e}_1, \cdots, \mathbf{e}_{2L}]$ and $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \cdots, \lambda_{2L})$. Here, $\lambda_l$ and $\mathbf{e}_l$ denote the $l$th largest eigenvalue of $\mathbf{R}_{\tilde{\mathbf{r}}}$ and its corresponding eigenvector, respectively. Then, for $M < 2L$, we have

$$
||\mathbf{E}_{\text{N}}^{\text{T}} \tilde{\boldsymbol{\Psi}}_{i(k)} \tilde{\mathbf{q}}_k|| = 0, \ k \in \mathcal{K}, \tag{25}
$$

where $\mathbf{E}_{\text{N}} = [\mathbf{e}_{M+1}, \cdots, \mathbf{e}_{2L}]$.
**Proof:** See Appendix B.

To resolve the ambiguity in the case of $M_i > 1$, we can derive a subspace method based on (25). Define a set of $Q$-sparse binary SIV with the support in $\mathcal{I}_i$ as

$$
\Sigma_{\mathcal{I}_i, Q} = \{\mathbf{q} \,|\, \tilde{\mathbf{q}} \in \mathcal{Q}, \, \text{supp}(\tilde{\mathbf{q}}) \subset \mathcal{I}_i\}.
$$

Then, the SIV of an active device that chooses $\boldsymbol{\Psi}_i$ can be an element of $\Sigma_{\mathcal{I}_i, Q}$. From (25), we can estimate the SIV of an active device that chooses $\boldsymbol{\Psi}_i$ as follows:

$$
\hat{\mathbf{q}} = \underset{\tilde{\mathbf{q}} \in \Sigma_{\mathcal{I}_i, Q}}{\text{argmin}} ||\mathbf{E}_{\text{N}}^{\text{T}} \tilde{\boldsymbol{\Psi}}_i \tilde{\mathbf{q}}||^2. \tag{26}
$$

Using this subspace method, we can decide $M_i$ SIVs of the active devices that choose $\boldsymbol{\Psi}_i$ in the second step. These SIVs are denoted by $\{\hat{\mathbf{q}}_{k_i(1)}, \cdots, \hat{\mathbf{q}}_{k_i(M_i)}\}$.

Note that in (26), in practice, since the eigenvectors in $\mathbf{E}_{\text{N}}$ are not available, their estimates that are the eigenvectors obtained from a sample correlation matrix of the $\tilde{\mathbf{r}}_t$'s, $\frac{1}{T}\sum_t \tilde{\mathbf{r}}_t \tilde{\mathbf{r}}_t^{\text{T}}$, have to be used.

### C. Detection of the Signals from Active Devices

With $\{\hat{\mathbf{q}}_{k_i(1)}, \cdots, \hat{\mathbf{q}}_{k_i(M_i)}\}$, we can detect the signals, $s_{k_i(m);t}$, $m = 1, \cdots, M_i$, from $\mathbf{r}_t$. For convenience, let

$$
\begin{aligned}
\mathbf{D}_i &= [\tilde{\boldsymbol{\Psi}}_i \hat{\mathbf{q}}_{k_i(1)}, \cdots, \tilde{\boldsymbol{\Psi}}_i \hat{\mathbf{q}}_{k_i(M_i)}] \in \mathbb{R}^{2L \times M_i} \\
\mathbf{s}_{i;t} &= [s_{k_i(1);t}, \cdots, s_{k_i(M_i);t}]^{\text{T}} \in \mathbb{R}^{M_i}.
\end{aligned}
$$

If $\hat{\mathbf{q}}_{k_i(m)} = \tilde{\mathbf{q}}_{k_i(m)}$ (assuming that all the estimated SIVs are correct), $\tilde{\mathbf{r}}_t$ can be expressed as

$$
\begin{aligned}
\tilde{\mathbf{r}}_t &= \sum_{i=1}^{J} \mathbf{D}_i \mathbf{s}_{i;t} + \tilde{\mathbf{n}}_t \\
&= [\mathbf{D}_1, \cdots, \mathbf{D}_J][\mathbf{s}_{1;t}^{\text{T}}, \cdots, \mathbf{s}_{J;t}^{\text{T}}]^{\text{T}} + \tilde{\mathbf{n}}_t \\
&= \mathbf{D}\mathbf{s}_t + \tilde{\mathbf{n}}_t,
\end{aligned} \tag{27}
$$

where $\mathbf{D} = [\mathbf{D}_1, \cdots, \mathbf{D}_J] \in \mathbb{R}^{2L \times M}$ and $\mathbf{s}_t = [\mathbf{s}_{1;t}^{\text{T}}, \cdots, \mathbf{s}_{J;t}^{\text{T}}]^{\text{T}} \in \mathbb{R}^M$. There are various MUD methods to estimate $\mathbf{s}_t$ with known $\mathbf{D}$ [24], [25]. Among those, we can use the minimum mean squared error (MMSE) multiuser detector as follows:

$$
\hat{\mathbf{s}}_t = \mathbf{F}_{\text{mmse}} \tilde{\mathbf{r}}_t, \tag{28}
$$

where $\mathbf{F}_{\text{mmse}} = \left(\mathbf{D}^{\text{T}}\mathbf{D} + \frac{N_0}{2}\mathbf{I}\right)^{-1}\mathbf{D}^{\text{T}}$.

## V. SIMULATION RESULTS

In this section, we present simulation results with the precoding matrices that are generated independently. Each element of $\boldsymbol{\Psi}_i$ is an independent CSCG random variable with mean zero and variance $1/L$. In addition, we assume that the length of CIR, $\nu$, is set to 10 (i.e., $\nu = 10$) and $h_{k,p} \sim \mathcal{CN}(0, \frac{1}{\nu})$ is independent (i.e., uniform power delay profile is assumed). Through simulation results, we can see the impact of various parameters (e.g., $L$: The number of subcarriers; $Q$: The sparsity of SIV; $M$: The number of active devices; $J$: The number of precoding matrices for the pre-equalization; and signal-to-noise ratio (SNR)) on the performance. In general, a better detection performance is expected for a smaller $Q$ and a smaller $M$ due to the sparsity constraints in CS. Furthermore, a large $J$ can improve the detection performance due to a better pre-equalization performance. However, since the number of columns of the measurement matrix $\tilde{\boldsymbol{\Psi}}$ increases with $J$, the performance improvement by increasing $J$ might be limited. We can confirm the above behaviors from simulation results in this section.

### A. Simulation Results of Pre-Equalization

In this subsection, we consider the performance of the pre-equalization in (11) with multiple precoding matrices.

The SSE in (13) is used to see the performance of pre-equalization. Fig. 1(a) shows the SSE for various values of $P_{\text{max}}$
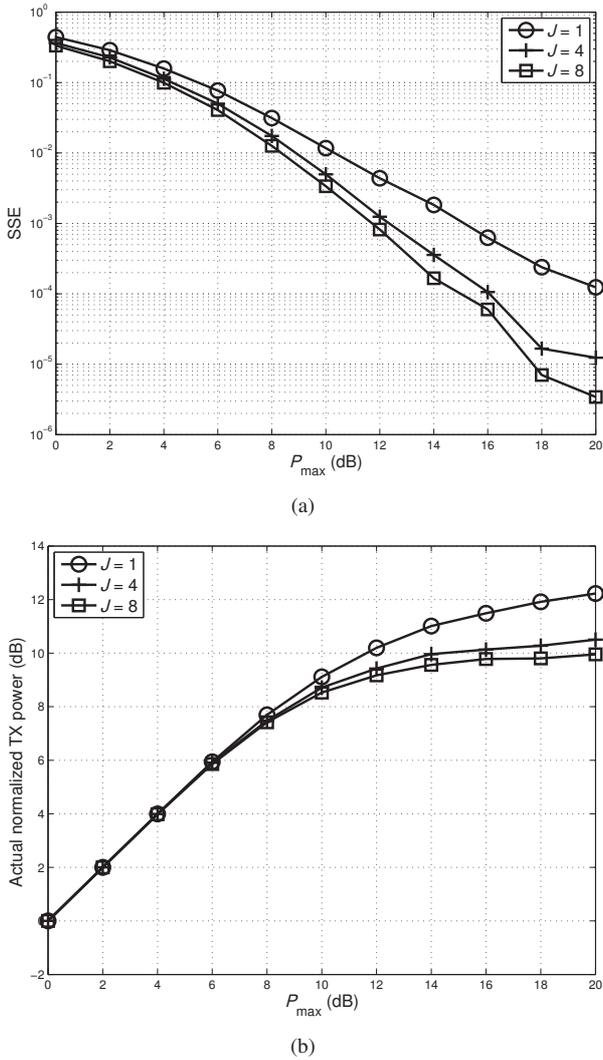
(a)



(b)

Fig. 1. Performances of the pre-equalization when $L = 128$ and $Q = 2$: (a) SSE versus $P_{\max}$ and (b) actual normalized transmission power versus $P_{\max}$.

when $L = 128$ and $Q = 2$. For a larger $P_{\max}$, a better solution to the problem in (13) (or a better pre-equalization) is available. That is, the SSE can decrease with $P_{\max}$, which is clearly shown in Fig. 1(a). In addition, since a better performance of pre-equalization can be achieved with more pre-determined precoding matrices, we can observe that the SSE decreases with $J$ in Fig. 1(a). We also note that the performance gain by increasing $J$ might be saturated once there are a sufficient number of pre-determined precoding matrices.

In Fig. 1(b), we show the actual normalized transmission power for various values of $P_{\max}$. Since the actual transmission power is $\min_j P_j$, it can decrease with $J$, while it increases with $P_{\max}$. It is noteworthy that the actual transmission power is saturated as $P_{\max}$ increases once a sufficiently low SSE is achieved. In particular, the actual transmission power increases slowly once $P_{\max}$ is greater than 10 dB and becomes slower for a larger $J$, because a lower SSE can be achieved with a larger $J$.

In order to see the impact of $Q$ on the performance of pre-equalization, we consider different values of $Q$ when $L = 128$, $J = 4$, and $P_{\max} = 10$ dB and show the simulation results in
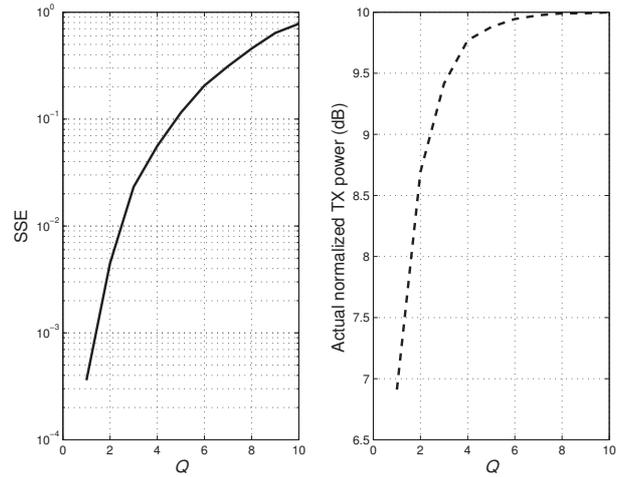


Fig. 2. Performances of the pre-equalization for various values of $Q$ when $L = 128$, $J = 4$, and $P_{\max} = 10$ dB.

Fig. 2. As $Q$ increases, the sparsity increases, which results in a poor CS recovery performance [12]. Thus, both the SSE and actual transmission power increase with $Q$. Consequently, although the number of SIVs, $N_I$, increases with $Q$, we may need to have a small $Q$ for a good performance of pre-equalization.

### B. Simulation Results of Identification

In this subsection, we mainly focus on the probability of successful identification, which is the probability that the SIVs of *all*[5] $M$ active devices are successfully identified. In order to estimate $\mathcal{J}$, we use the OMPMMV algorithm in [22].
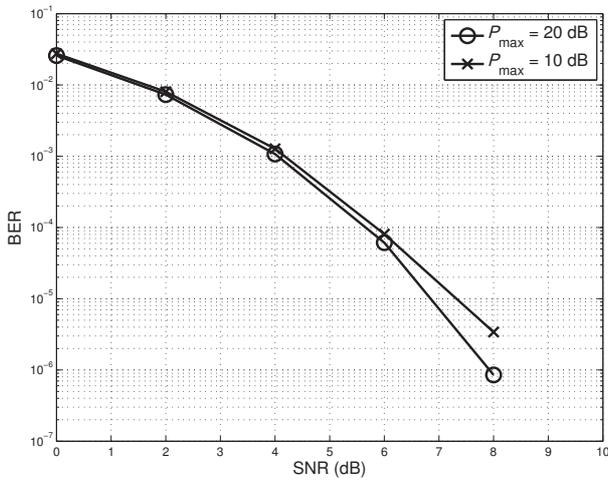
Fig. 3(a) shows the probability of successful identification for different values of the SNR, which is given by $\mathrm{SNR} = \mathbb{E}[|s_{k;t}|^2]/N_0 = 1/N_0$, when $L = 128$, $J = 4$, $Q = 2$, $T = 50$, and $M = 5$. We consider two values of $P_{\max}$: 10 and 20 dB. Although a larger $P_{\max}$ has a lower SSE as shown in Fig. 1, we can see that $P_{\max} = 10$ dB would be sufficiently large to have a low SSE and, as a result, there is no significant difference between the probabilities of successful identification with $P_{\max} = 10$ and 20 dB as shown in Fig. 3(a). Furthermore, we can see that once the SNR is greater than 6 dB, there is no improvement of the probabilities of successful identification by increasing the SNR. This demonstrates that the performance of active device identification could depend on other parameters, e.g., $M$, once the SNR is sufficiently high. This behavior results from the performance limitation of CS algorithms due to sparsity constraints. In other words, for a high probability of successful identification, we need not only a high SNR, but also low sparsity (the performance in terms of $M$ will be shown in Fig. 4).

In Fig. 3(b), we also show the bit error rate (BER) for various values of SNRs. This BER is obtained only when the identification of all $M$ active devices is successful using the MMSE multiuser detector in Section IV-C. We can see that once the SNR is sufficiently high (i.e., greater than 8 dB), the BER becomes

---

[5]That is, we assume that the identification is not successful even if the AP fails to identify just few active devices. In practice, active devices that are not successfully identified are back-logged and re-try in the next slot.
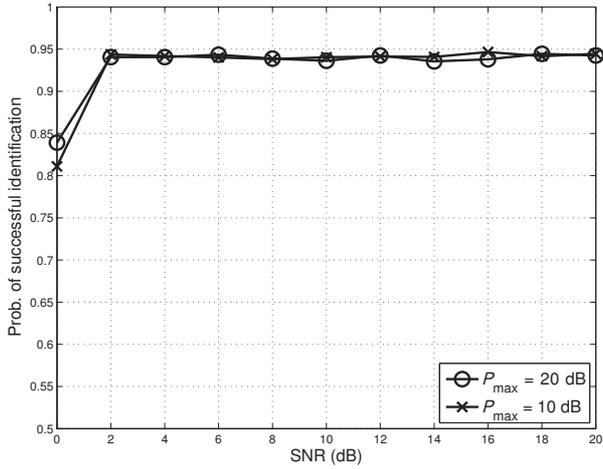
(a)



(b)

Fig. 3.   Performances of active device identification when $L = 128$, $J = 4$, $Q = 2$, $T = 50$, and $M = 5$: (a) Probability of successful identification and (b) BER.



Fig. 4.   Probability of successful identification for various values of $M$ when $J = 4$, $Q = 2$, $T = 50$, and SNR $= 10$ dB (with $P_{\max} = 20$ dB).



Fig. 5.   Probability of successful identification for various values of $J$ when $Q = 2$, $T = 50$, and SNR $= 10$ dB (with $P_{\max} = 20$ dB).

virtually zero. Thus, in the TDD-SIMA scheme, the identification of active devices is more critical than the detection of data symbols from active devices [7], [8].

Fig. 4 shows the probability of successful identification for various values of $M$ when $J = 4$, $Q = 2$, $T = 50$, and SNR $= 10$ dB (with $P_{\max} = 20$ dB). We can see that the probability of successful identification decreases with $M$ and it can be improved by increasing $L$.

Fig. 5 shows the probability of successful identification for various values of $J$ when $Q = 2$, $T = 50$, and SNR $= 10$ dB (with $P_{\max} = 20$ dB). We also consider different combinations for $(L, M)$ in Fig. 5. In general, we can see that the probability of successful identification increases with $J$. Thus, the more precoding matrices, the better identification of active devices. In addition, as shown in Fig. 1(a), the actual transmission power also decreases with $J$. Consequently, we can see that in TDD-SIMA, it would be crucial to have multiple precoding matrices.

For performance comparisons, we consider a simple approach that is based on orthogonal frequency division multiple access (OFDMA). In the OFDMA bas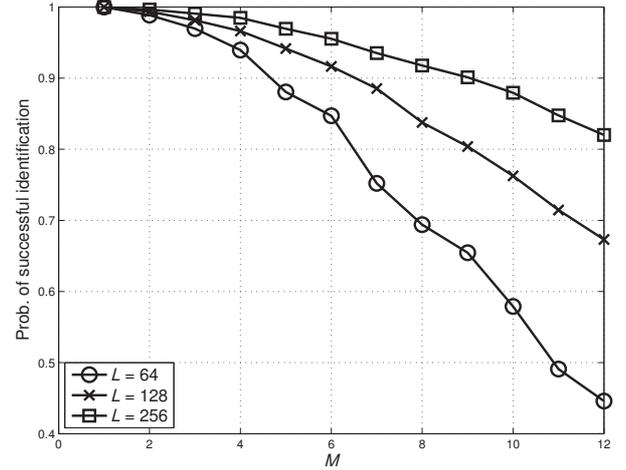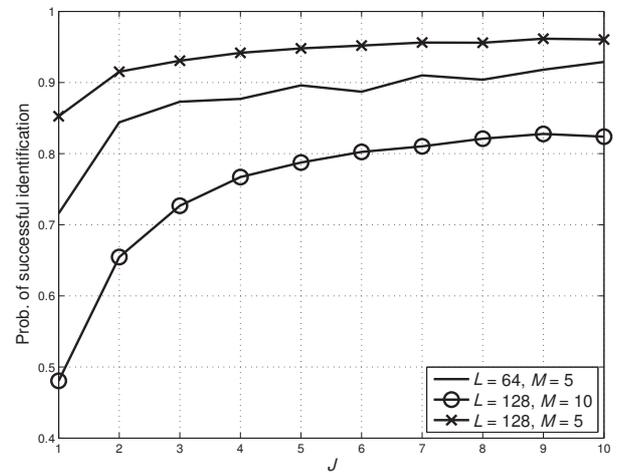ed approach, the subcarriers are divided into multiple clusters where each cluster consists of $G$ subcarriers. A device is to randomly choose a cluster to transmit signals for random access. Therefore, if multiple active devices choose the same cluster, there might be collisions. In this case, when $M$ devices are active, the throughput[6] becomes

$$\eta(M) = M \left( 1 - \frac{G}{L} \right)^{M-1}.$$

Note that a cluster needs to have multiple subcarriers to achieve a better detection performance by improving the diversity gain, while the throughput decreases with $G$ as the probability of collision increases. Thus, there is a trade-off between the detection performance and the probability of collision through $G$. For simplicity, the perfect CSI estimation is assumed for the OFDMA based approach. In Fig. 6, we show the simulation results of the OFDMA based approach and the proposed approach,

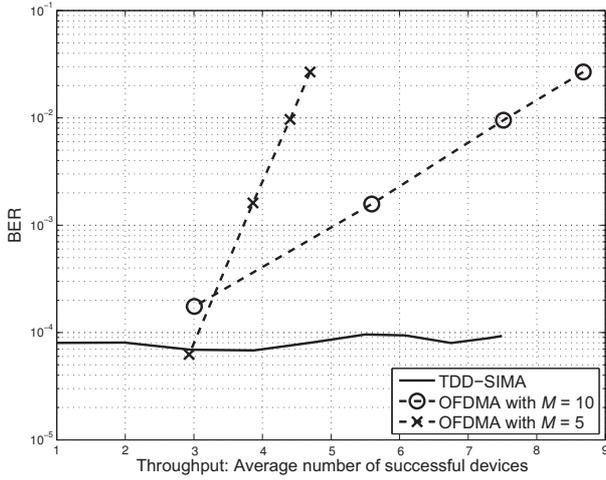[6]The throughput is the average number of devices that can successfully access a channel without any collision.

Fig. 6. Throughput versus BER when $L = 128$ and SNR = 6 dB (with $P_{\max} = 10$ dB, $Q = 2$, and $J = 4$ for the proposed approach).

i.e., the TDD-SIMA scheme, when $L = 128$ and SNR = 6 dB. For the OFDMA based approach, we consider $G \in \{2, 4, 8, 16\}$. As $G$ increases, the number of clusters decreases as $L$ is fixed. We can see that the throughput decreases with $G$, while a better BER is achieved for a larger $G$. This confirms the trade-off between the throughput and detection performance. For the TDD-SIMA scheme, we assume $Q = 2$ and $J = 4$, while $P_{\max} = 10$ dB for the pre-equalization. In the TDD-SIMA scheme, the throughput becomes the product of the number of active devices and the probability of successful identification. From Fig. 6, we can observe that the TDD-SIMA scheme can provide a low BER for a wide range of the throughput, which is important for one-shot identification/detection as a successfully identified active device needs to transmit data symbols reliably so that no re-trials are required with a high probability. We can also see that the BER of the OFDMA based approach is higher than that of the TDD-SIMA scheme. Furthermore, the BER of the OFDMA based approach increases with the throughput. Thus, although the throughput is high, re-trial or re-transmission may be needed as data packets cannot be decoded due to a high BER, which eventually degrades the throughput (once re-transmissions are taken into account).

## VI. CONCLUDING REMARKS

We studied a random access scheme, called TDD-SIMA, which employs SIVs as unique identification sequences for devices. TDD-SIMA has exploited the channel reciprocity in TDD mode for pre-equalization with multiple precoding matrices, which can eliminate the need of the channel estimation at the AP and lower the transmission power at each device. Due to the sparsity of SIVs as well as the sparsity of active devices, in TDD-SIMA, one-shot identification/detection became possible using a low-complexity two-step method at the AP, which can significantly reduce signaling overhead as active device identification can be performed with data detection. In the first step, a low-complexity CS algorithm was used to estimate a sparse signal which is the superposition of multiple sparse signals from

active devices. To identify the unique SIVs of active devices from the estimated sparse signal in the first step, a subspace method was proposed. Simulation results confirmed that the derived two-step method can successfully identify multiple active devices in TDD-SIMA.

## APPENDICES
### I. PROOF OF THEOREM 1

Since $\mathbf{H}$ and $\mathbf{A}$ are diagonal, the SSE becomes

$$C_j(\mathbf{a}) = \|\mathbf{HAH}^{\mathrm{H}}\mathbf{x}_j - \mathbf{x}_j\|^2$$
$$= \sum_l \beta_{j,l}(|H_l|^2 a_l - 1)^2. \qquad (29)$$

For convenience, let $\omega_{j,l} = \beta_{j,l}|H_l|^2$. Then, (13) can be rewritten as

$$\min_{\mathbf{a}} C_j(\mathbf{a})$$
$$\text{subject to } \sum_l \omega_{j,l} a_l^2 \le P_{\max}. \qquad (30)$$

Since the cost function, $C_j(\mathbf{a})$, is quadratic in $\mathbf{a}$ and the constraint set is convex, the method of Lagrange multipliers can be used to find the optimal $\mathbf{a}$ that is given by

$$\hat{a}_{j,l}(\lambda) = \frac{|H_l|^2 \beta_{j,l}}{|H_l|^4 \beta_{j,l} + \lambda \omega_{j,l}}, \qquad (31)$$

which becomes (14) after some manipulations.

Substituting (31) into (12), we have $P_j = \sum_{l=0}^{L-1} \omega_{j,l} \hat{a}_{j,l}^2(\lambda)$, which becomes (15). This completes the proof.

### II. PROOF OF THEOREM 2

Under (11), from (18), $\tilde{\mathbf{z}}_{i;t}$ becomes

$$\tilde{\mathbf{z}}_{i;t} = [\tilde{\mathbf{q}}_{k_i(1)}, \cdots, \tilde{\mathbf{q}}_{k_i(M_i)}] \begin{bmatrix} s_{k_i(1);t} \\ \vdots \\ s_{k_i(M_i);t} \end{bmatrix} = \mathbf{B}_{i;t}\mathbf{u}_i, \qquad (32)$$

where $\mathbf{B}_{i;t} = [s_{k_i(1);t}\mathbf{I}, \cdots, s_{k_i(M_i);t}\mathbf{I}]$ and $\mathbf{u}_i = [\tilde{\mathbf{q}}_{k_i(1)}^{\mathrm{T}}, \cdots, \tilde{\mathbf{q}}_{k_i(M_i)}^{\mathrm{T}}]^{\mathrm{T}}$. Note that $\mathbf{B}_{i;t} \in \mathbb{R}^{2L \times 2LM_i}$ and $\mathbf{u}_i \in \mathbb{R}^{2LM_i}$. In addition, since $\mathbf{u}_i$ is an $M_iQ$-sparse vector, $\mathbf{B}_{i;t}\mathbf{u}_i$ is a $2L \times 1$ vector of $M_iQ$ sparsity. It can be shown that

$$\tilde{\mathbf{z}}_t = \tilde{\mathbf{z}}_t(\{\mathbf{u}_i\}, \mathbf{B}_t) = \begin{bmatrix} \mathbf{B}_{1;t}\mathbf{u}_1 \\ \vdots \\ \mathbf{B}_{J;t}\mathbf{u}_J \end{bmatrix}, \qquad (33)$$

where $\mathbf{B}_t = \{\mathbf{B}_{1;t}, \cdots, \mathbf{B}_{J;t}\}$. Let

$$\mathbf{v}_{k_i(m)} = \tilde{\mathbf{\Psi}}_i \tilde{\mathbf{q}}_{k_i(m)} \in \mathbb{R}^{2L \times 1}. \qquad (34)$$

Then, from $\mathbf{v}_{k_i(m)}$, we can define

$$\mathbf{V}_i = [\mathbf{v}_{k_i(1)}, \cdots, \mathbf{v}_{k_i(M_i)}] \in \mathbb{R}^{2L \times M_i} \qquad (35)$$

and $\mathbf{V} = [\mathbf{V}_1, \cdots, \mathbf{V}_J] \in \mathbb{R}^{2L \times M}$.

From (18), it can be shown that

$$\tilde{\mathbf{\Psi}}\tilde{\mathbf{z}}_t = \sum_{i=1}^{J} \tilde{\mathbf{\Psi}}_i \left( \sum_{k \in \mathcal{K}_i} \tilde{\mathbf{q}}_k s_{k;t} \right) = \sum_{k \in \mathcal{K}} \tilde{\mathbf{\Psi}}_{i(k)} \tilde{\mathbf{q}}_k s_{k;t}. \tag{36}$$

Then, from (3), we can further show that

$$\mathbf{R}_{\tilde{\mathbf{r}}} = \mathbb{E}\left[ \tilde{\mathbf{\Psi}}\tilde{\mathbf{z}}_t \tilde{\mathbf{z}}_t^{\mathrm{T}} \tilde{\mathbf{\Psi}}^{\mathrm{T}} \right] + \frac{N_0}{2}\mathbf{I} = \mathbf{V}\mathbf{V}^{\mathrm{T}} + \frac{N_0}{2}\mathbf{I}. \tag{37}$$

From (37), we have $\lambda_1 \geq \cdots \geq \lambda_M > \lambda_{M+1} = \lambda_{2L} = \frac{N_0}{2}$. Since $\mathrm{Span}(\mathbf{e}_{M+1}, \cdots, \mathbf{e}_{2L})$ is orthogonal to $\mathrm{Range}(\mathbf{V})$, from (36) and (37), we can have the following important relationship:

$$\mathrm{Range}(\mathbf{E}_{\mathrm{N}}) \perp \tilde{\mathbf{\Psi}}_{i(k)}\tilde{\mathbf{q}}_k, \ k \in \mathcal{K}, \tag{38}$$

which implies (25). This completes the proof.

## REFERENCES

[1] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, pp. 86–93, June 2013.

[2] 3GPP TR 37.868 V11.0, *Study on RAN improvments for machine-type communications*, Oct. 2011.

[3] ITU-T, *Y.2060: Overview of the Internet of things*, 2012.

[4] H. Zhu and G. Giannakis, "Exploiting sparse user activity in multiuser detection," *IEEE Trans. Commun.*, vol. 59, pp. 454–465, Feb. 2011.

[5] L. Applebaum, W. U. Bajwa, M. F. Duarte, and R. Calderbank, "Asynchronous code-division random access using convex optimization," *Physical Commun.*, vol. 5, no. 2, pp. 129–147, 2012.

[6] F. Fazel, M. Fazel, and M. Stojanovic, "Random access compressed sensing over fading and noisy communication channels," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 2114–2125, May 2013.

[7] H. F. Schepker, C. Bockelmann, and A. Dekorsy, "Exploiting sparsity in channel and data estimation for sporadic multi-user communication," in *Proc. ISWCS*, Aug. 2013, pp. 1–5.

[8] G. Wunder, P. Jung, and C. Wang, "Compressive random access for post-LTE systems," in *Proc. IEEE ICC*, June 2014, pp. 539–544.

[9] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, pp. 4203–4215, Dec. 2005.

[10] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[11] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[12] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[13] J. Choi, "On the sparsity for random access in machine type communications under frequency-selective fading," in *Proc. IEEE ICC*, May 2016, pp. 1–5.

[14] H. S. Dhillon, H. Huang, H. Viswanathan, and R. A. Valenzuela, "Fundamentals of throughput maximization with random arrivals for M2M communications," *IEEE Trans. Commun.*, vol. 62, pp. 4094–4109, Nov. 2014.

[15] J. Choi, "Sparse index multiple access," in *Proc. IEEE GlobalSIP*, Dec. 2015, pp. 324–327.

[16] J. Choi, "Sparse index multiple access for multi-carrier systems with precoding," *J. Commun. and Neworks*, vol. 18, pp. 3226–3237, June 2016.

[17] D. Mottier and D. Castelain, "SINR-based channel pre-equalization for uplink multi-carrier CDMA systems," in *Proc. PIMRC*, vol. 4, Sept. 2002, pp. 1488-1492.

[18] S. Nobilet, J. F. Helard, N. Jolivet, and R. Gillard, "Performance of pre-distortion techniques for uplink MC-CDMA systems with TDD and FDD modes," in *Proc. WPMC*, Oct. 2002, pp. 655–662.

[19] J. Choi, "Interference mitigation using transmitter filters in CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 51, pp. 657–666, July 2002.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[21] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, Feb. 2006.

[22] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Processing*, vol. 54, pp. 4634–4643, Dec. 2006.

[23] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Trans. Inf. Theory*, vol. 58, pp. 1135–1146, Feb. 2012.

[24] S. Verdu, *Multiuser Detection*. Cambridge University Press, 1998.

[25] J. Choi, *Optimal Combining and Detection*. Cambridge University Press, 2010.

**Jinho Choi** (SM'02) was born in Seoul, Korea. He received B.E. (magna cum laude) degree in electronics engineering in 1989 from Sogang University, Seoul, and M.S.E. and Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon in 1991 and 1994, respectively. He is with Gwangju Institute of Science and Technology (GIST) as a Professor. Prior to joining GIST in 2013, he was with the College of Engineering, Swansea University, United Kingdom, as a Professor/Chair in Wireless. His research interests include wireless communications and array/statistical signal processing. He authored two books published by Cambridge University Press in 2006 and 2010. He received the 1999 Best Paper Award for Signal Processing from EURASIP, 2009 Best Paper Award from WPMC (Conference), and is Senior Member of IEEE. Currently, he is an Editor of IEEE Trans. Communications and had served as an Associate Editor or Editor of other journals including IEEE Communications Letters, Journal of Communications and Networks (JCN), IEEE Transactions on Vehicular Technology, and ETRI journal.