Channel Division Multiple Access for MTC under Frequency-Selective Fading

Jinho Choi and Nam Yul Yu

***Abstract*—In this paper, we study random access for machine-type communications (MTC) based on a multiple access scheme that exploits different channel state information (CSI) from active devices under frequency-selective fading, which is called channel division multiple access (ChDMA). In ChDMA, in order to use a low-complexity compressive sensing (CS) algorithm for multiuser detection (MUD) with a high transmission rate, we employ index modulation. Since the CSI is used as a signature for MUD, the CSI estimation becomes crucial, but suffers from the pilot collision. In order to lower the probability of pilot collisions, multiple transmissions of randomly selected pilots are considered.**

***Index Terms*—random access; machine-type communications; compressive sensing**

I. Introduction

There has been a growing interest in machine-type communications (MTC) in order to support a number of devices that are to be connected to a network [1]. The applications of MTC are diverse from health care to wireless sensor networks and the Internet of Things (IoT). For MTC, random access is usually considered due to low signaling/control overhead in supporting a number of devices with a low probability of activity [1]–[3]. For example, in the long term evolution-advanced (LTE-A) system, a random access scheme, called random access (RACH) procedure, is proposed for MTC [4]. In 5G systems, it is also expected to support up to 30,000 devices per cell [5].

As in the RACH procedure, when multiple devices can transmit randomly chosen preambles[1] from a pool of preambles, an access point (AP) needs to detect them simultaneously. If the AP is able to detect all transmitted preambles, those devices can successfully establish connections, which results in throughput improvement. Clearly, this becomes a key advantage of the RACH procedure over single-channel random access schemes (e.g., slotted ALOHA) that have a low throughput. In fact, the RACH procedure can be seen as a multi-channel random access scheme with multiple non-orthogonal channels where each channel is characterized by a different preamble.

For the case that the activity of devices is low, the notion of compressive random access[2] has been proposed in [6]–[9], where the sparse activity is exploited to derive computationally efficient multiple signal detection schemes for a receiver at the AP based on compressive sensing (CS) algorithms [10], [11]. If the sparse activity is not exploited, the AP may use an exhaustive search to detect the set of active devices together with a conventional multiuser detection (MUD) approach for joint signal detection [12]. In this case, MUD with exhaustive search may suffer from a prohibitively high computational complexity for a large number of devices unless multiple channels are orthogonal. It is noteworthy that even if orthogonal sequences are used to create multiple channels (as in code division multiple access (CDMA)), the orthogonality may not be retained due to the distortion by channels as demonstrated in most multicarrier (MC)-CDMA systems [13]. Consequently, under a realistic channel environment (e.g., frequency-selective fading channels), compressive random access would be promising to support massive connections for MTC.

In [14], compressive random access (for sporadic uplink transmissions) is studied under frequency-selective fading channels with the channel estimation, which is different from other approaches in [7], [9] where the channel state information (CSI) is assumed to be perfectly known. Note that in [6], the channel estimation is also considered in compressive random access. In [15], [16], compressive random access is studied for multicarrier systems with the CS-based signal detection in the frequency-domain. Note that in [15], compressive random access is referred to as CS-based MUD in order to emphasize that the notion of CS is exploited in MUD. While CS-based MUD would be a reasonable term, we prefer compressive random access in this paper as in [6] in order to emphasize the nature of unscheduled sporadic transmissions for random access (in MTC) and the sparsity of activity that allows to derive low-complexity CS-based MUD methods.

In this paper, we study a compressive random access scheme over frequency-selective fading channels that can mitigate some problems in conventional compressive random access schemes. As in [6], [7], [14], each device can have a unique spreading code. However, if the number of devices is large, it may not be possible to assign a unique code for each device due to various reasons (e.g., if orthogonal codes are used, the number of devices is limited by the length of codes). In the proposed compressive random access scheme, however, the CSI of active device is used as a signature for MUD. As a result, the proposed compressive random access scheme can support any number of devices as long as there are few active devices and any active device's CSI is different from each other. For a higher transmission rate, we employ index

The authors are with School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST) (*Email: jchoi0114@gist.ac.kr*). This research was supported by GIST Research Fund (for the Project of "Living Energy") in 2017.

[1]Throughout the paper, we assume that preamble and pilot are interchangeable. That is, a pilot sequence is also called a preamble sequence, vice versa.

[2]This term is used in [6] to refer to a random access scheme that allows a receiver to employ low-complexity CS algorithms for MUD.

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

modulation for data transmission. For the CSI estimation, we consider a conventional approach as the RACH procedure [4] with a pilot pool of finite size. In this case, unfortunately, the channel estimation suffers from pilot collision [16]. To mitigate this, we consider multiple transmissions of pilots in the proposed compressive random access scheme. As a result, the proposed scheme can support a large number of devices of sparse activity with a negligible probability of pilot collision.

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). The superscript $\dagger$ denotes the pseudo-inverse. 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 nth column (element, resp.). If n is a set of indices, $[\mathbf{X}]_n$ is a submatrix of $\mathbf{X}$ obtained by taking the corresponding columns. $\mathbb{E}[\cdot]$ 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 for Random Access

Suppose that there are a number of devices associated with an AP. The total number of devices is denoted by K. While K can be very large, we assume that only a fraction of them are active at a time and send short messages to the AP. Thus, random access might be suitable for communications from devices to the AP due to low signaling/control overhead.

Throughout the paper, we assume that each active device transmits a packet within a given frame and a packet consists of $(T + B)$ blocks, where the first B blocks are pilot blocks and the last T blocks are data blocks, which is illustrated in Fig. 1. A pilot block has $N + \bar{P}$ symbols, where $\bar{P}$ is the length of cyclic prefix (CP) and N is the length of a pilot sequence. The structure of a data block is similar to that of a pilot block. Throughout the paper, we assume that the channel is invariant within a frame (i.e., the coherence time is longer than the length of frame).

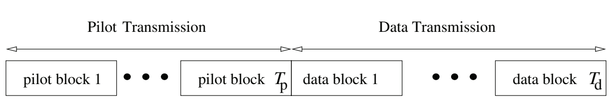


Fig. 1. The structure of a packet consisting of multiple blocks for pilot and data transmissions.

For MTC, we consider sporadic transmissions from devices as in [14]. Thus, in each frame, we may have a different set of active devices. From this, we only consider signal transmissions within one frame. In addition, throughout the paper, we assume that the CSI is to be estimated at the AP and active device is to randomly choose a pilot from a set of pilots (or pilot pool). To see the differences of the assumptions in this paper from those in other existing approaches, we present Table I.

TABLE I
Assumptions of CSI and codes for pilots of the approach in this paper and other existing approaches (from the list of references).

Pilots	Known CSI	CSI to be estimated
Unique codes	[7], [15]	[6], [14]
Randomly selected codes		[16], this paper

A. Pilot Transmissions

For random access as in [4], each active device is to randomly choose one out of L pre-determined pilots and transmit it in order to allow the AP to estimate the CSI of active device. The set of pilot sequences is given by

$$\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_L\}, \tag{1}$$

where $\mathbf{c}_l \in \mathbb{C}^{N \times 1}$ denotes the lth pilot. We assume that all the devices are synchronized and active devices can transmit their randomly selected pilots simultaneously with $B = 1$. Denote by $l(m)$ the index of the pilot chosen by the mth active device and by M the number of active devices among K devices. The received signal after removing CP when the pilot block is transmitted is given by

$$x_q = \sum_{m=1}^{M} \sum_{p=0}^{P-1} h_{k(m),p} c_{l(m),(q-p)_N} + \tilde{n}_q, \ q = 0, \ldots, N-1, \tag{2}$$

where $k(m)$ is the device index of the mth active device, $h_{k,p}$ is the channel impulse response (CIR) from device k to the AP, P is the length of CIR, and $\tilde{n}_q \sim \mathcal{CN}(0, N_0)$ is the background noise. Here, it is assumed that $P \leq \bar{P}$ and $\mathbf{c}_l = [c_{l,0} \ \ldots \ c_{l,N-1}]^{\mathrm{T}}$ represents the lth pilot sequence in $\mathcal{C}$ in (1) and

$$(x)_N = x \ (\mathrm{mod} \ N).$$

For example, suppose that $K = 64$, $L = 10$, and $M = 2$. If the device indices of the first and second active devices are 2 and 54, respectively, we have $k(1) = 2$ and $k(2) = 54$. If the first and second active devices choose $\mathbf{c}_4$ and $\mathbf{c}_{10}$, respectively, we have $l(1) = 4$ and $l(2) = 10$.

Let $\mathbf{x} = [x_0 \ \ldots \ x_{N-1}]^{\mathrm{T}}$ and $\tilde{\mathbf{n}} = [\tilde{n}_0 \ \ldots \ \tilde{n}_{N-1}]^{\mathrm{T}}$. In addition, define the discrete Fourier transform (DFT) matrix of size $N \times N$ as $[\mathbf{F}]_{n,q} = e^{-\frac{j2\pi nq}{N}}$, where $n, q \in \{0, \ldots, N-1\}$. The DFT of $\mathbf{x}$ becomes

$$\mathbf{y} = \mathbf{F}\mathbf{x} = \sum_{m=1}^{M} \mathbf{H}_{k(m)} \mathbf{F} \mathbf{c}_{l(m)} + \mathbf{n}, \tag{3}$$

where $\mathbf{n} = \mathbf{F}\tilde{\mathbf{n}}$ and $\mathbf{H}_k$ is the frequency-domain channel matrix from device k to the AP, which is a diagonal matrix and given by $\mathbf{H}_k = \mathrm{diag}(H_{k,0}, \ \ldots H_{k,N-1})$. Here, $H_{k,n} = \sum_{p=0}^{P-1} h_{k,p} e^{-\frac{j2\pi pn}{N}}$.

If there are multiple active devices that choose the same pilot, their pilot signals may not be correctly detected, which result in pilot collision. The pilot collision can be mitigated if each active device can transmit multiple pilots (i.e., $B \geq 1$) as illustrated in Fig. 1. For each pilot transmission, an active device can randomly and independently choose a pilot from $\mathcal{C}$. If $B > 1$, the event of pilot collision can be seen in a higher-dimensional space.

Property 1. *Suppose that $B \geq 1$. In addition, assume that each active device is to randomly choose a pilot from $\mathcal{C}$ with equal probability for each pilot block. Then, the probability of pilot collision, denoted by $P_{\rm pc}$, is given by*

$$P_{\rm pc} \approx 1 - e^{-\frac{M^2}{2L^B}}. \tag{4}$$

Proof: See Appendix A. ■

In order to see the impact of B on the probability of pilot collision, we can consider an example with $M = 10$ and $L = 100$. We have $P_{\rm pc} \approx 0.393$ with $B = 1$, while $P_{\rm pc} \approx 0.005$ with $B = 2$. This demonstrates that the probability of pilot collision can be significantly lowered by multiple pilot transmissions.

For comparison purposes, we can consider a longer pilot block with more pilot sequences in $\mathcal{C}$. For example, suppose that the length of pilot block becomes doubled (which is equivalent to $B = 2$ in terms of the total length of pilot) and the number of pilots in $\mathcal{C}$ is also doubled. Then, the probability of pilot collision becomes $P_{\rm pc} \approx 1 - e^{-\frac{M^2}{4L}}$. If $M = 10$ and $L = 100$, the resulting probability of pilot collision is $P_{\rm pc} \approx 0.221$, which is not a significant improvement compared to that by multiple pilot transmissions.

B. Data Transmissions

After transmitting a randomly chosen pilot, an active device transmits T blocks of data symbols with CP. Denote by $\mathbf{a}_{t;k} = [a_{t;k,0} \ \dots \ a_{t;k,N-1}]^{\rm T}$ the tth data block transmitted by device k, where $a_{t;k,q}$ is the qth element of $\mathbf{a}_{t;k}$. Then, after removing CP, the received signal becomes

$$x_{t;q} = \sum_{m=1}^{M} \sum_{p=0}^{P-1} h_{k(m),p} a_{t;k(m),(q-p)_N} + \tilde{n}_{t;q},$$

where $\tilde{n}_{t;q} \sim \mathcal{CN}(0, N_0)$ is the qth background noise of data block t. Let $\mathbf{x}_t = [x_{t;0} \ \dots \ x_{t;N-1}]^{\rm T}$ and $\tilde{\mathbf{n}}_t = [\tilde{n}_{t;0} \ \dots \ \tilde{n}_{t;N-1}]^{\rm T}$. To perform MUD in the frequency-domain, we can consider the the DFT of $\mathbf{x}_t$, which is

$$\mathbf{y}_t = \mathbf{F}\mathbf{x}_t = \sum_{m=1}^{M} \mathbf{H}_{k(m)} \mathbf{F} \mathbf{a}_{t;k(m)} + \mathbf{n}_t, \tag{5}$$

where $\mathbf{n}_t = \mathbf{F}\tilde{\mathbf{n}}_t$.

To allow MUD, each active device can use the same pilot sequence that was chosen for pilot transmissions as a signature when it transmits data symbols [15], [16]. In this case, the signal vector of the mth active device is given by as

$$\mathbf{a}_{t;k(m)} = \mathbf{c}_{l(m)} d_{t;k(m)}, \tag{6}$$

where $d_{t;k(m)}$ is the data symbol of device $k(m)$ (or the mth active device) in the tth data block. Then, $\mathbf{y}_t$ in (5) becomes

$$\mathbf{y}_t = \sum_{m=1}^{M} \mathbf{w}_m d_{t;k(m)} + \mathbf{n}_t = \mathbf{W}\mathbf{d}_t + \mathbf{n}_t, \tag{7}$$

where $\mathbf{w}_m = \mathbf{H}_{k(m)} \mathbf{c}_{l(m)}$, $\mathbf{W} = [\mathbf{w}_1 \ \dots \ \mathbf{w}_M]$, and $\mathbf{d}_t = [d_{t;k(1)} \ \dots \ d_{t;k(M)}]^{\rm T}$.

III. Compressive Channel Division Multiple Access

In this section, we take into account the fact that each active device has a different CSI and exploit this for the proposed compressive random access scheme where CS-based MUD can be used to detect signals from multiple active devices with index modulation, which can achieve a higher transmission rate than that of conventional approach discussed in Subsection II-B. Throughout this section, we assume that the AP knows $\mathbf{H}_k$ for $k \in \mathcal{A}$, where $\mathcal{A}$ denotes the set of active devices, i.e., $\mathcal{A} = \{k(1), \dots, k(M)\}$.

A. Channel Division Multiple Access with Index Modulation

Noting that each active device can be characterized by its CSI in (5), we do not need to use additional signature (i.e., $\mathbf{c}_{l(m)}$) for MUD. In this subsection, we propose a multiple access based on CSI. While $\mathbf{c}_{l(m)}$ is not used as additional signature, it will be used for index modulation to increase the transmission rate.

Suppose that a device can choose Q positions in a data block and transmit pulses in those positions as a generalization of pulse position modulation (PPM). This becomes index modulation in the time domain and $\mathbf{a}_{t;k(m)}$ becomes a Q-sparse vector, i.e.,

$$\mathbf{a}_{t;k(m)} \in \Sigma_Q,$$

where Σ_Q denotes the set of Q-sparse vectors, i.e., $\Sigma_Q = \{\mathbf{x} \,|\, ||\mathbf{x}||_0 = Q\}$. For convenience, a Q-sparse vector, $\mathbf{a}_{t;k(m)}$, is referred to as an index modulated vector (IMV). In this case, the number of bits per block that can be transmitted by an active device becomes $\lfloor \log_2 \binom{N}{Q} \rfloor$. The received signal during data block t at the AP is given by

$$\mathbf{y}_t = \sum_{m=1}^{M} \mathbf{H}_{k(m)} \mathbf{F} \mathbf{a}_{t;k(m)} + \mathbf{n}_t = \mathbf{\Phi}\mathbf{a}_t + \mathbf{n}_t, \tag{8}$$

where $\mathbf{a}_t = [\mathbf{a}_{t;k(1)}^{\rm T} \ \dots \ \mathbf{a}_{t;k(M)}^{\rm T}]^{\rm T}$ and $\mathbf{\Phi} = [(\mathbf{H}_{k(1)}\mathbf{F}) \ \dots \ (\mathbf{H}_{k(M)}\mathbf{F})] \in \mathbb{C}^{N \times NM}$. The measurement matrix, $\mathbf{\Phi}$, is determined by the CSI of active devices and the channel matrices $\{\mathbf{H}_{k(m)}\}$ are used as signatures of active devices for MUD. From this, the resulting multiple access is referred to as channel division multiple access (ChDMA).

We have few remarks as follows.

- Since $\mathbf{a}_t \in \Sigma_{MQ}$, if $N \gg MQ$, $\mathbf{a}_t$ can be recovered from $\mathbf{y}_t$ in (8) using a CS algorithm in the frequency-domain. We can also exploit the structure of $\mathbf{a}_t$ for a better performance. Each subvector of $\mathbf{a}_t$, $\mathbf{a}_{t;k(m)}$, is Q-sparse. A modified OMP algorithm to take into account this property is presented in [17].

- We can consider the real-valued representation of $\mathbf{a}_{t;k(m)}$. Then, $\mathbf{a}_{t;k(m)}$ becomes a real-valued vector of $2N$ elements. In this case, the number of bits per block becomes $\lfloor \log_2 \binom{2N}{Q} \rfloor$. That is, more bits[3] can be transmitted using a total of Q sparse pulses in the real and imaginary domains.
- In ChDMA, as shown above, the AP can distinguish and recover the signals from multiple active devices using their different channel matrices, $\{\mathbf{H}_{k(m)}\}$. Thus, the performance depends on the difference between $\mathbf{H}_{k(m)}$'s and a poor performance is expected when P is small (especially, for flat-fading channels). To avoid this difficulty, precoding could be employed. In particular, for each active device, a different precoding matrix, which is associated with the selected pilot, can be used to differentiate its signal from the others. For example, if active device $k(m)$ chooses pilot $l(m)$, the $l(m)$th precoding matrix[4] can be used from a set of pre-determined precoding matrices.

B. Impact of Key Parameters on a Recovery Guarantee

In this subsection, we consider the impact of key parameters (e.g., M, Q, and P) on a recovery guarantee with the coherence in ChDMA.

For convenience, we omit the time index t in this subsection. For a recovery guarantee in most CS algorithms, the properties of the measurement matrix, $\mathbf{\Phi}$, are important. For example, the coherence of $\mathbf{\Phi}$ is often considered, which is defined as

$$\mu(\mathbf{\Phi}) = \max_{l \neq m} \frac{|\boldsymbol{\phi}_l^{\mathrm{H}} \boldsymbol{\phi}_m|}{||\boldsymbol{\phi}_l|| \; ||\boldsymbol{\phi}_m||},$$

where $\boldsymbol{\phi}_l$ is the lth column of $\mathbf{\Phi}$. From [18], for a recovery guarantee (in the absence of noise), a sufficient condition is found as $||\mathbf{s}||_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\mathbf{\Phi})}\right)$. This implies that

$$\mu(\mathbf{\Phi}) < \frac{1}{2MQ - 1}. \tag{9}$$

Thus, for a given $\mathbf{\Phi}$, it is desirable to keep MQ constant for given $\mathbf{\Phi}$. In other words, the sparsity, Q, has to decrease as the number of active devices increases to keep a certain recovery performance.

In order to see the impact of P on a recovery guarantee, consider the following inner product of two columns of $\mathbf{\Phi}$:

$$\nu_{k,k';l,m} = ([\mathbf{H}_k \mathbf{F}]_l)^{\mathrm{H}} [\mathbf{H}_{k'} \mathbf{F}]_m = N \sum_p h_{k,p} h^*_{k',p+l-m}. \tag{10}$$

Since $||[\mathbf{H}_k \mathbf{F}]_l||^2 = \sum_n |H_{k,n} e^{-j\frac{2\pi l n}{N}}|^2 = N \sum_p |h_{k,p}|^2$, from (10), we have

$$\bar{\nu}_{k,k';l,m} = \frac{\sum_p h_{k,p} h^*_{k',p+l-m}}{\sqrt{\sum_p |h_{k,p}|^2} \sqrt{\sum_p |h_{k',p}|^2}}, \tag{11}$$

which is the normalized auto-correlation and cross-correlation functions of $\{h_{k,p}\}$ and $\{h_{k',p}\}$. Note that if $|l - m| \geq P$, $\bar{\nu}_{k,k';l,m}$ becomes 0. From this, since the value of $|\bar{\nu}_{k,k';l,m}|$ tends to be large as $|l-m|$ approaches 0, we are interested in the value of $|\bar{\nu}_{k,k';l,m}|$ when $l = m$ for the coherence of $\mathbf{\Phi}$, because the coherence of $\mathbf{\Phi}$ is the largest value of $|\bar{\nu}_{k,k';l,m}|$ for $(k,l) \neq (k',m)$.

In order to see the properties of $\bar{\nu}_{k,k';l,l}$, we consider the following assumption.

A) The channel coefficients of the CIR are independent and CSCG random variables as follows:

$$h_{k,p} \sim \mathcal{CN}\left(0, \frac{1}{P}\right). \tag{12}$$

That is, the channels are assumed to be Rayleigh multipath channels.

Under **A)**, we can readily have $\mathbb{E}[\bar{\nu}_{k,k';l,l}] = 0, k \neq k'$.

Property 2. *Let $X = |\bar{\nu}_{k,k';l,l}|^2$, which a normalized inner product of any two distinct columns of $\mathbf{\Phi}$. Note that X is independent of l (i.e., the value of X is the same for all l). For a sufficiently large P, the probability that X is less than $\frac{1}{(2MQ-1)^2}$ is given by*

$$\Pr\left(X \leq \frac{1}{(2MQ-1)^2}\right) = 1 - \exp\left(-\frac{P}{(2MQ-1)^2}\right). \tag{13}$$

Proof: See Appendix B. ∎

From (9), to keep a certain recovery performance, it is desirable to have

$$\frac{P}{(2MQ-1)^2} \approx \text{Const.} \tag{14}$$

Thus, for a larger P, we can have more active devices with a fixed Q or a large Q (i.e., a higher data rate) with a fixed M.

IV. Simulation Results

In this section, we present simulation results when the CIR of each active device is independently generated under Assumption **A)**. For CS-based MUD, the OMP algorithm is used, while the block OMP algorithm [19] is used for the channel estimation to exploit the block sparsity. For a pool of pilots, we use orthogonal Alltop sequences [20] with $N = L$.

A. Simulation Results of Channel Estimation

As mentioned earlier, the main advantage of using a common pool of pilots in compressive random access (over using a unique pilot for each device) is to support any number of devices at the cost of pilot collision. To compare the performance of the channel estimation based on a common pool of pilots with that based on unique pilots, we consider the following two schemes:

- CRA-1: in this scheme, a common pilot pool is used with Alltop sequences of length $N = 47$;
- CRA-2: each device has a unique random pilot sequence (each element is one of $\{(\pm 1 \pm j)/\sqrt{2N}\}$) and the

[3] In addition, pulse amplitude modulation (PAM) can be used for non-zero elements of $\mathbf{a}_{t;k}$ to send more bits.

[4] This implies that the number of precoding matrices should be the same as the number of pilots in $\mathcal{C}$, L.

number of devices is up to $K = \{25N, 50N\}$, where $N = 94$ in this scheme (i.e., the scheme in [6], [14]).

In CRA-1, if $B = 2$, the total length of pilot sequence becomes $47 \times 2 = 94$, which is the same as that of CRA-2 where only one pilot sequence (per active device) is transmitted. In Fig. 2, the successful recovery rates of CRA-1 and CRA-2 are shown for various values of M when $P = 4$ and signal-to-noise ratio (SNR) = 20 dB. Here, the SNR is defined as $\text{SNR} = \frac{||\mathbf{c}_l||^2}{N_0} = \frac{1}{N_0}$. Note that in CRA-2, the block sparsity is equal to M. On the other hand, in CRA-1, the block sparsity is less than or equal to M due to pilot collision. From Fig. 2, we can see that the performance of CRA-1 can be similar to that of CRA-2 although its N is a half of that of CRA-2 due to good pilot sequences (Alltop sequences) and an effectively smaller block sparsity (than M). In spite the fact that the performances of CRA-1 and CRA-2 are similar to each other when M is small (up to $M = 7$), we note that the successful recovery rate of CRA-1 decreases quickly and becomes much lower than that of CRA-2 for a large M ($M \geq 8$). However, fortunately, since the region of a high successful recovery rate is usually of interest for compressive random access, we can claim that the performance of CRA-1 is comparable to that of CRA-2, while it can support any number of devices.

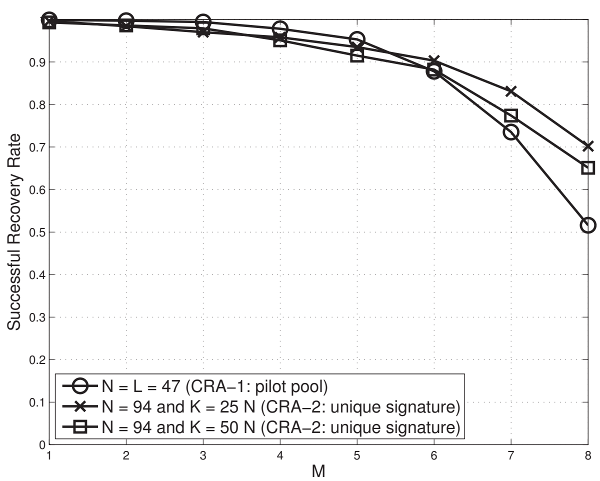
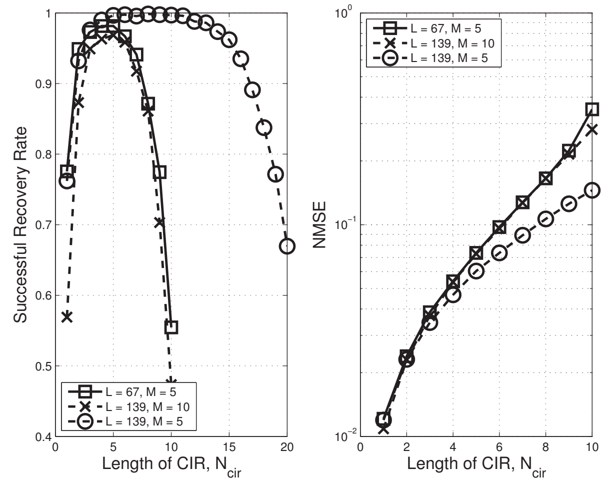


Fig. 2. Successful recovery rate versus M when $P = 4$ and SNR = 20 dB.

Fig. 3 shows the performances of the channel estimation for different values of P when SNR = 20 dB. We can observe that the successful recovery rate increases with P and then decreases. In the channel estimation, the signal to be recovered is the CIR. Thus, under Assumption **A)**, the variation of the signal power (i.e., the squared norm of $\mathbf{h}_k$) becomes smaller as P increases. Thus, as P increases, the recovery of CIR can be more successful. However, if P becomes too large, the block OMP algorithm suffers from a large block sparsity and cannot recover the signal well unless $N = L$ is large. The normalized mean squared error (NMSE) of the channel estimate, which is given by

$$\text{NMSE} = \frac{\mathbb{E}[||\mathbf{g}_l - \hat{\mathbf{g}}_l||^2]}{\mathbb{E}[||\mathbf{g}_l||^2]},$$

where $\hat{\mathbf{g}}_l$ is the estimated $\mathbf{g}_l$ if $\mathbf{g}_l$ is a non-zero vector, is also shown in Fig. 3. We can see that the NMSE increases with P (as there are more channel coefficients to be estimated when P increases).



(a) (b)

Fig. 3. Performance of channel estimation for different values of P when SNR = 20 dB: (a) Successful recovery rate; (b) NMSE (only for the signals that can be recovered by the block OMP algorithm).

B. Simulation Results of MUD in ChDMA

In this subsection, we show the symbol error rate (SER) of ChDMA when index modulation is used as a generalization of PPM. Since the real-valued version of $\mathbf{a}_{t;k}$ is Q-sparse and its length is $2N$, we assume that each IMV can transmit $\log_2 \binom{2N}{Q}$ bits. Thus, the bit energy is given by $E_\text{b} = \frac{A^2}{\log_2 \binom{2N}{Q}}$, where A is the amplitude of non-zero elements of the real-valued version of $\mathbf{a}_{t;k}$.

Fig. 4 shows the SER for various values of $\frac{E_\text{b}}{N_0}$ when $N = 139$, $Q = 2$, $M = 4$, and $P = 10$. We can observe the error floor (i.e., the SER cannot be lowered although $\frac{E_\text{b}}{N_0}$ increases once $\frac{E_\text{b}}{N_0}$ is sufficiently high (in this case, $\frac{E_\text{b}}{N_0} \geq -5$ dB)). This error floor is due to the recovery performance of the CS-based MUD method (i.e., the OMP algorithm in this section). As discussed in Subsection III-B, a better performance can be achieved if P becomes larger or MQ becomes smaller.

According to (13), in ChDMA, the performance of signal recovery depends on the product of M and Q. In addition, the performance becomes better as P increases for a fixed MQ. This prediction from (13) can be confirmed by Fig. 5, where the SER for various values of P with different combinations of M and Q is shown when $L = 139$ and $\frac{E_\text{b}}{N_0} = 10$ dB.

V. Concluding Remarks

We have proposed a compressive random access scheme with a relatively small number of good pilot sequences that is suitable for MTC to support any number of devices over frequency-selective fading channels. For a sufficiently low probability of pilot collision, we also proposed to employ

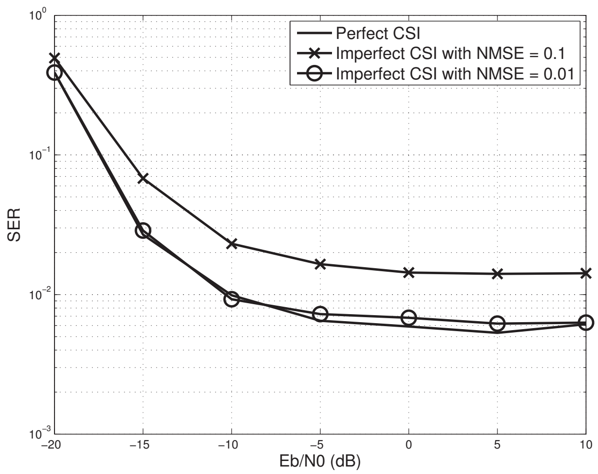


Fig. 4. Symbol error rate for various values of $\frac{E_{\rm b}}{N_0}$ when $L = 139$, $Q = 2$, $M = 4$, and $P = 10$.

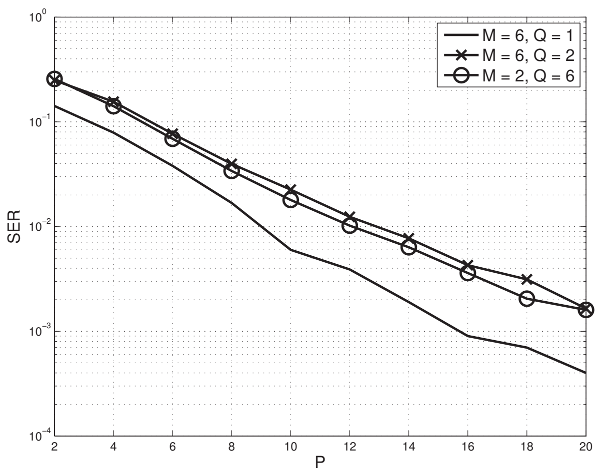


Fig. 5. Symbol error rate for various values of P when $L = 139$ and $\frac{E_{\rm b}}{N_0} = 10$ dB.

multiple pilot transmissions. Since the unique CSI of each active device can be exploited to allow the AP to differentiate data symbols of one active device from those of other active devices, there is no need to assign unique signature sequences to devices. Consequently, the proposed compressive random access was able to support a large number of devices with sparse activity for sporadic transmissions in MTC.

Appendix A
Proof of Property 1

A pilot chosen by an active device for each pilot block can be seen as a random variable uniformly distributed over $\{1, \ldots, L\}$. For $B \geq 1$, the set of the B pilots chosen by an active device is also seen as a random vector uniformly distributed over $\{1, \ldots, L\}^B$. Thus, we have $P_{\rm pc} = 1 - \frac{\frac{(L^B)!}{(L^B - M)!}}{L^{BM}} \approx 1 - e^{-\frac{M^2}{2L^B}}$ for $L^B \gg M$. This completes the proof.

Appendix B
Proof of Property 2

For a sufficiently large P, under **A)**, from [21], we have $\bar{\nu}_{k,k';l,l} \sim \mathcal{CN}\left(0, \frac{1}{P}\right)$. From this, it follows that $X \sim {\rm Exp}\,(P) = Pe^{-Px}$, $k \neq k'$, where ${\rm Exp}(\lambda)$ represents the exponential distribution with parameter λ (the mean is $\frac{1}{\lambda}$). Since $F_X(x) = 1 - \exp(-Px)$, (13) can be obtained.

References

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

[2] K.-D. Lee, S. Kim, and B. Yi, "Throughput comparison of random access methods for M2M service over LTE networks," in *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pp. 373–377, Dec 2011.

[3] D. Niyato, P. Wang, and D. I. Kim, "Performance modeling and analysis of heterogeneous machine type communications," *IEEE Trans. Wireless Communications*, vol. 13, pp. 2836–2849, May 2014.

[4] 3GPP TR 37.868 V11.0, *Study on RAN improvments for machine-type communications*, October 2011.

[5] M. Condoluci, M. Dohler, G. Araniti, A. Molinaro, and K. Zheng, "Toward 5G densenets: architectural advances for effective machine-type communications over femtocells," *IEEE Communications Magazine*, vol. 53, pp. 134–141, January 2015.

[6] G. Wunder, . Stefanovi, P. Popovski, and L. Thiele, "Compressive coded random access for massive MTC traffic in 5G systems," in *2015 49th Asilomar Conference*, pp. 13–17, Nov 2015.

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

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

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

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

[11] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, April 2006.

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

[13] K. Fazel and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems*. John Wiley & Sons, 2003.

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

[15] F. Monsees, M. Woltering, C. Bockelmann, and A. Dekorsy, "Compressive sensing multi-user detection for multicarrier systems in sporadic machine type communication," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1–5, May 2015.

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

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

[18] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inform. Theory*, vol. 52, pp. 6–18, Jan 2006.

[19] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Processing*, vol. 58, pp. 3042–3054, June 2010.

[20] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Springer, 2013.

[21] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forcasting*. Springer, second ed., 2002.