Dynamic Compressive Sensing-Based Multi-User Detection for Uplink Grant-Free NOMA

Bichai Wang, Linglong Dai, Yuan Zhang, Talha Mir, and Jianjun Li

***Abstract*—Non-orthogonal multiple access (NOMA) can support more users than OMA techniques using the same wireless resources, which is expected to support massive connectivity for Internet of Things in 5G. Furthermore, in order to reduce the transmission latency and signaling overhead, grant-free transmission is highly expected in the uplink NOMA systems, where user activity has to be detected. In this letter, by exploiting the temporal correlation of active user sets, we propose a dynamic compressive sensing (DCS)-based multi-user detection (MUD) to realize both user activity and data detection in several continuous time slots. In particular, as the temporal correlation of the active user sets between adjacent time slots exists, we can use the estimated active user set in the current time slot as the prior information to estimate the active user set in the next time slot. Simulation results show that the proposed DCS-based MUD can achieve much better performance than that of the conventional CS-based MUD in NOMA systems.**

***Index Terms*—5G, non-orthogonal multiple access (NOMA), multi-user detection (MUD), dynamic compressive sensing (DCS).**

I. Introduction

MULTIPLE access technology has been regarded as the landmark of each generation of wireless systems [1]. Particularly, orthogonal multiple access (OMA), i.e., orthogonal frequency division multiple access (OFDMA), is used in current 4G systems. In OMA, the number of supportable users is strictly limited by the number of available orthogonal resources, which is difficult to satisfy the demand of massive connectivity in future 5G systems [1]. To address this challenge, non-orthogonal multiple access (NOMA) has been actively investigated [1], [2], which can realize overloading by non-orthogonal resource allocation at the cost of increased receiver complexity.

In current 4G systems, the uplink transmission is scheduled by the base station (BS) in a request-grant procedure, where a large transmission latency and signaling overhead will be caused. This problem becomes worse or even unacceptable when massive connectivity occurs in 5G. Therefore, grant-free transmission is highly expected in uplink NOMA systems, where users can randomly transmit data and the BS cannot know which users are active, and thus user activity has to be detected. Some solutions have been proposed to realize user activity detection based on compressive sensing (CS) [3], [4] by exploiting the user activity sparsity due to sporadic communication in Internet-of-Things (IoT). However, in these CS-based MUD schemes, the signal detection was usually independently realized in different time slots, where the user activity correlation in different time slots was not considered. Recently, joint signal detection in several continuous time slots within a frame was proposed to improve the performance by exploiting the frame-wise sparsity [5], [6], where the user activity was assumed to be unchanged within a whole frame. However, in 5G IoT applications with sporadic communication, where users can randomly access or leave the system, the frame structure transmission is not preferred. Therefore, active user sets may be changed in different time slots.

In this letter, we consider the more practical scenario that the active user sets can be changed in several continuous time slots, and propose a low-complexity dynamic compressive sensing (DCS)-based MUD [7] for NOMA to jointly realize user activity and data detection. Specifically, although users can randomly access or leave the system, some users generally transmit their information in adjacent time slots with a high probability, which leads to the temporal correlation of active user sets in several continuous time slots [7]. By exploiting such temporal correlation, we propose to use the estimated active user set in the current time slot as the initial set to estimate the transmitted signal in the next time slot under the framework of DCS. Simulation results show that the proposed DCS-based MUD can achieve much better performance than that of the conventional CS-based MUD [3], [4].

The rest of this letter is organized as follows. The system model is introduced in Section II. Section III describes the proposed DCS-based MUD scheme, and Section IV presents the performance analysis. Simulation results are provided in Section V. Finally, conclusions are drawn in Section VI.

Notation: Upper-case and lower-case boldface letters denote matrices and vectors, respectively; $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$, $(\cdot)^\dagger$, and $\|\cdot\|_p$ denote the transpose, conjugate transpose, matrix inversion, Moore-Penrose matrix inversion, and l_p norm operation, respectively; $|\Gamma|$ denotes the number of elements in set Γ, and $\Gamma \backslash \hat{\Gamma}$ denotes the set composed of elements in Γ while not in $\hat{\Gamma}$; $\mathbf{h}_\Gamma$ denotes the entries of the vector $\mathbf{h}$ in the set Γ; $\mathbf{H}_\Gamma$ denotes the submatrix comprising the Γ columns of $\mathbf{H}$.

II. System Model

We consider a typical uplink NOMA system with one BS and K users, where the single antenna scenario is considered without loss of generality. The transmitted symbol x_k for user k is modulated onto a spreading sequence $\mathbf{s}_k$ of

Manuscript received August 13, 2016; accepted August 19, 2016. Date of publication August 24, 2016; date of current version November 9, 2016. This work was supported by the International Science & Technology Cooperation Program of China (Grant No. 2015DFG12760) and the National Natural Science Foundation of China (Grant No. 61571270). The associate editor coordinating the review of this letter and approving it for publication was A. Lioumpas.

B. Wang, L. Dai, Y. Zhang, and T. Mir are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: wbc15@mails.tsinghua.edu.cn; daill@tsinghua.edu.cn; zhangyuan13@mails.tsinghua.edu.cn; bah15@mails.tsinghua.edu.cn).

J. Li is with the School of Electric and Information Engineering, Zhongyuan University of Technology, Zhengzhou 451191, China (e-mail: jianjun.li@tsinghua.org.cn).

Digital Object Identifier 10.1109/LCOMM.2016.2602264

length N. Particularly, we consider the case of $N < K$, i.e., the overloaded system, where the number of users can be larger than the length of spreading sequences. After that, signals from all active users are superimposed, and then are transmitted over N orthogonal OFDM subcarriers. The received signal on subcarrier n at the BS can be represented as

$$y_n = \sum_{k=1}^{K} g_{nk} s_{nk} x_k + \upsilon_n, \quad n = 1, 2, \cdots, N, \tag{1}$$

where s_{nk} is the nth component of spreading sequence $\mathbf{s}_k$, and υ_n is the Gaussian noise on subcarrier n with zero mean and variance σ^2. g_{nk} is the channel gain of user k on the nth subcarrier, and all of them are independent and identically distributed (i.i.d.) complex Gaussian variables with zero mean and unit variance [3], [8]. We combine the received signals over all N subcarriers, and then the received signal vector $\mathbf{y} = [y_1, y_2, \cdots, y_N]^T$ can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}, \tag{2}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_K]^T$ is the transmitted signal vector for all K users, $\mathbf{H}$ is the measurement matrix of size $N \times K$, whose element h_{nk} in the nth row and the kth column equals to $g_{nk} s_{nk}$, and $\mathbf{v} = [\upsilon_1, \upsilon_2, \cdots, \upsilon_N]^T$ is the noise vector following the distribution $\mathcal{CN}\left(0, \sigma^2 \mathbf{I}_N\right)$.

III. Proposed DCS-Based MUD for NOMA

In practical scenarios, although users can randomly access or leave the system, some users generally transmit their information in adjacent time slots with a high probability, which leads to the temporal correlation of active user sets in several continuous time slots [7]. Particularly, we consider the reconstruction of transmitted signals $\mathbf{x} = \left[\mathbf{x}^{[1]}, \mathbf{x}^{[2]}, \cdots, \mathbf{x}^{[J]}\right]$ from the received signals $\mathbf{y} = \left[\mathbf{y}^{[1]}, \mathbf{y}^{[2]}, \cdots, \mathbf{y}^{[J]}\right]$ in J continuous time slots (e.g., $J = 7$ has been considered in LTE-Advanced standard [6]), and thus we have

$$\mathbf{y}^{[j]} = \mathbf{H}^{[j]} \mathbf{x}^{[j]} + \mathbf{v}^{[j]}, \quad j = 1, 2, \cdots, J, \tag{3}$$

where $\mathbf{H}^{[j]}$ is the equivalent channel matrix in the jth time slot, which can be changed in different time slots, and $\mathbf{v}^{[j]}$ is the Gaussian noise vector in the jth time slot.

We assume that the sparsity level of $\mathbf{x}^{[j]}$, i.e., the maximum number of nonzero elements of $\mathbf{x}^{[j]}$, is S, and the support of $\mathbf{x}^{[j]}$ is defined as

$$\Gamma^{[j]} = \left\{k : k \in \{1, 2, \cdots, K\}, x_k^{[j]} \neq 0\right\}, \tag{4}$$

which denotes the index set of nonzero elements in $\mathbf{x}^{[j]}$. Different from the previous work in [6], where the active user sets were assumed to be unchanged in several continuous time slots within a whole frame, in this letter we consider the more general scenario that the active user sets can be changed in several continuous time slots, which is more practical in 5G IoT applications with sporadic communication.

According to the statistics of the mobile traffic [8], the number of active users is usually much smaller than the number of all potential users even in busy hours. Thus, the MUD problem is inherently a sparse signal recovery problem [5], where $\mathbf{x}^{[1]}, \mathbf{x}^{[2]}, \cdots, \mathbf{x}^{[J]}$ are sparse transmitted vectors.

Algorithm 1 The Proposed DCS-Based MUD

Input:
Received signals: $\mathbf{y}^{[1]}, \mathbf{y}^{[2]}, \cdots, \mathbf{y}^{[J]}$;
Equivalent channel matrices: $\mathbf{H}^{[1]}, \mathbf{H}^{[2]}, \cdots, \mathbf{H}^{[J]}$;
Sparsity level: S.
Output:
Reconstructed sparse signals: $\hat{\mathbf{x}}^{[1]}, \hat{\mathbf{x}}^{[2]}, \cdots, \hat{\mathbf{x}}^{[J]}$.

```
$\hat{\Gamma}_c = \emptyset$.
for $j = 1$ to $J$ do
  $i = 1$, $\hat{\Gamma}^{[j](i)} = \hat{\Gamma}_c$, $\hat{\mathbf{x}}^{[j](i)} = \left(\mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}}\right)^{\dagger} \mathbf{y}^{[j]}$,
  $\mathbf{r}^{[j](i)} = \mathbf{y}^{[j]} - \mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}} \hat{\mathbf{x}}^{[j](i)}$.
  while $i = 1$ or $\left\|\mathbf{r}^{[j](i)}\right\|_2 < \left\|\mathbf{r}^{[j](i-1)}\right\|_2$ do
    $i = i + 1$.
    $\hat{\Gamma}^{[j](i)} = \hat{\Gamma}^{[j](i-1)} \cup \arg\max_k \left|\left(\mathbf{h}_k^{[j]}\right)^H \mathbf{r}^{[j](i-1)}\right|^2$,
    where $\mathbf{h}_k^{[j]} = \mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}}(:, k)$.
    $\hat{\mathbf{x}}^{[j](i)} = \left(\mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}}\right)^{\dagger} \mathbf{y}^{[j]}$.
    if $\left|\hat{\Gamma}^{[j](i)}\right| \geq S$
      $\hat{\Gamma}^{[j](i)} = \{S$ indices with largest magnitude in $\hat{\mathbf{x}}^{[j](i)}\}$.
    end if
    $\mathbf{r}^{[j](i)} = \mathbf{y}^{[j]} - \mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}} \left(\mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}}\right)^{\dagger} \mathbf{y}^{[j]}$.
  end while
  $\hat{\Gamma}^{[j]} = \hat{\Gamma}^{[j](i-1)}$, $\hat{\Gamma}_c = \hat{\Gamma}^{[j]}$.
  $\hat{\mathbf{x}}^{[j]} = \left(\mathbf{H}^{[j]}_{\Gamma^{[j]}}\right)^{\dagger} \mathbf{y}^{[j]}$.
end for
return $\hat{\mathbf{x}} = \left[\hat{\mathbf{x}}^{[1]}, \hat{\mathbf{x}}^{[2]}, \cdots, \hat{\mathbf{x}}^{[J]}\right]$.
```

Furthermore, as mentioned before, some users generally transmit their information in adjacent time slots with a high probability, which leads to the temporal correlation of active user sets in several continuous time slots [7]. By exploiting such temporal correlation, the estimated support of $\mathbf{x}^{[j]}$ in the jth time slot can be used as the initial support to estimate $\mathbf{x}^{[j+1]}$ in the $(j+1)$th time slot, where $j = 1, 2, \cdots, J-1$. Accordingly, we propose the DCS-based MUD to realize joint user activity and data detection in several continuous time slots.

Particularly, to realize accurate signal recovery by exploiting the temporal correlation of the active user sets, an accurate estimation of the active user set in the initial time slot is required, which can be realized by increasing the transmit power in the first time slot or using more reliable recovery algorithms with high computational complexity, e.g., basis pursuit denoising (BPDN) [7]. Note that this procedure just increases the system overhead or complexity at the BS in the first time slot, which is still superior to conventional uplink scheduling mechanisms, where the complicated uplink request and downlink scheduling will result in a large transmission latency and signaling overhead. The details of the proposed DCS-based MUD are shown in **Algorithm 1**, which is based on the classical orthogonal matching pursuit (OMP) algorithm [3]. Specifically, the main procedure of the proposed DCS-based MUD can be presented as follows.

1) (Step 3) Initialization: In the first time slot, the initial support is set as $\hat{\Gamma}_c = \emptyset$. In the jth time slot, where $j = 2, \cdots, J$, the estimated support $\hat{\Gamma}^{[j-1]}$ in the $(j-1)$th time slot can be used as the initial support in the jth time slot.
2) (Step 6) Update the estimated support: In each iteration, the user that correlates best with the residue signal will be included in the active user set at first. For example, in the ith iteration, the active user set in the jth time slot can be updated by

$$\hat{\Gamma}^{[j](i)} = \hat{\Gamma}^{[j](i-1)} \cup \arg\max_k \left| \left(\mathbf{h}_k^{[j]}\right)^H \mathbf{r}^{[j](i-1)} \right|^2, \quad (5)$$

where $\mathbf{h}_k^{[j]}$ is the kth column of $\mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}}$, and $\mathbf{r}^{[j](i-1)}$ is the residue signal, which can be calculated by Step 11 in the $(i-1)$th iteration.
3) (Step 7) Update the estimated signal: The sparse transmitted signal vector in the jth time slot can be updated by the least square (LS) algorithm:

$$\hat{\mathbf{x}}^{[j](i)} = \left(\mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}}\right)^{\dagger} \mathbf{y}^{[j]}. \quad (6)$$

4) (Step 9) Revise the estimated support: If the number of elements in the estimated support exceeds the sparsity lever S, S users with the largest magnitude in the reconstructed signal will be regarded as active users. The estimated support is revised by

$$\hat{\Gamma}^{[j](i)} = \{S \text{ indices with largest magnitude in } \hat{\mathbf{x}}^{[j](i)}\}. \quad (7)$$

5) (Step 11) Update the residue signal:

$$\mathbf{r}^{[j](i)} = \mathbf{y}^{[j]} - \mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}} \left(\mathbf{H}^{[j]}_{\hat{\Gamma}^{[j](i)}}\right)^{\dagger} \mathbf{y}^{[j]}. \quad (8)$$

The iteration will be terminated when the energy of the residue error signal is not reduced. After the estimated active user set in each time slot has been obtained, the sparse transmitted signal vector in the jth time slot can be recovered by the LS algorithm (see step 14).

Different from the conventional CS-based MUD [3], [4], the proposed DCS-based MUD exploits the temporal correlation within the active user sets to realize joint user activity and data detection in several continuous time slots. Thus, the active user sets can be detected with a high reliability by exploiting the partially known support. Meanwhile, unlike frame-wise MUD [5], [6], which assumes the active user sets remain unchanged in several continuous time slots within a whole frame, the proposed DCS-based MUD allows the active user sets in different time slots to be different, which makes it more practical in uplink grant-free NOMA systems.

In addition, the proposed DCS-based MUD scheme has the same order of computational complexity in each iteration, e.g., $O\left(NS\left(K+S^2\right)\right)$ in OMP, as the conventional CS-based MUD [3]. However, as the number of iterations in DCS-based MUD is lower than that of the CS-based MUD due to the partially known support, the proposed scheme has lower computational complexity.

IV. Performance Analysis

To investigate the performance guarantee of the proposed DCS-based MUD, we provide the condition of successful recovery and the upper bound of the signal detection error in this section.

In order to analyze the detection reliability, we firstly introduce the restricted isometry property (RIP) for signal reconstruction performance analysis in CS theory. Specifically, the measurement matrix $\mathbf{H}$ satisfies the RIP of order S if there exists a constant $\delta \in (0,1)$ such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\mathbf{H}\mathbf{x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \quad (9)$$

holds true for all S-sparse vectors $\mathbf{x}$, which means the maximum number of nonzero elements in $\mathbf{x}$ is S. Particularly, the Toeplitz matrix based on pseudo random noise can satisfy RIP with a high probability [9].

We also introduce the residue-orthogonality [9] of the projection operator used in CS theory: For an arbitrary vector $\mathbf{y} \in \mathbb{C}^N$ and a sampling matrix $\mathbf{H}_I \in \mathbb{C}^{N\times|I|}$ of full column rank, we have

$$\begin{cases} \mathbf{H}_I^H \operatorname{resid}(\mathbf{y}, \mathbf{H}_I) = 0, \\ \|\operatorname{resid}(\mathbf{y}, \mathbf{H}_I)\|_2 \le \|\mathbf{y}\|_2, \end{cases} \quad (10)$$

where $\operatorname{resid}(\mathbf{y}, \mathbf{H}_I) = \mathbf{y} - \mathbf{H}_I\mathbf{H}_I^{\dagger}\mathbf{y}$ denotes the residue signal.

For simplicity, we consider AWGN channels, then the measurement matrix $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_K]$ is equal to the spreading matrix $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_K]$, and $\mathbf{H}^{[1]} = \mathbf{H}^{[2]} = \cdots = \mathbf{H}^{[J]} \triangleq \mathbf{H}$. The spreading matrix is designed as a Toeplitz matrix based on pseudo random noise, and thus $\mathbf{H}$ satisfies $\mathbf{h}_i^H\mathbf{h}_i = 1$ and $\mathbf{h}_i^H\mathbf{h}_k \approx 0, \forall i \neq k$.

Thus, by applying RIP condition and residue-orthogonality (we omit the time slot superscript j for notational simplicity), $\left\|\mathbf{r}^{(i)}\right\|_2$ of the proposed DCS-based MUD can be derived as [9]

$$\begin{aligned} \left\|\mathbf{r}^{(i)}\right\|_2 &= \left\|\operatorname{resid}\left(\mathbf{y}, \mathbf{H}_{\hat{\Gamma}^{(i)}}\right)\right\|_2 \\ &\le \left\|\operatorname{resid}|\left(\mathbf{H}_{\Gamma\setminus\hat{\Gamma}^{(i)}}\mathbf{x}_{\Gamma\setminus\hat{\Gamma}^{(i)}}, \mathbf{H}_{\hat{\Gamma}^{(i)}}\right)\right\|_2 \\ &\quad + \left\|\operatorname{resid}\left(\mathbf{v}, \mathbf{H}_{\hat{\Gamma}^{(i)}}\right)\right\|_2 \\ &\le \left\|\mathbf{H}_{\Gamma\setminus\hat{\Gamma}^{(i)}}\mathbf{x}_{\Gamma\setminus\hat{\Gamma}^{(i)}}\right\|_2 + \|\mathbf{v}\|_2 \\ &\le \sqrt{1+\delta_{2S+1}}\left\|\mathbf{x}_{\Gamma\setminus\hat{\Gamma}^{(i)}}\right\|_2 + \|\mathbf{v}\|_2. \end{aligned} \quad (11)$$

Similarly, $\left\|\mathbf{r}^{(i-1)}\right\|_2$ can be also derived as

$$\left\|\mathbf{r}^{(i-1)}\right\|_2 \ge \frac{1-2\delta_{2S+1}}{\sqrt{1-\delta_{2S+1}}}\left\|\mathbf{x}_{\Gamma\setminus\hat{\Gamma}^{(i-1)}}\right\|_2 - \|\mathbf{v}\|_2. \quad (12)$$

In order to compare $\left\|\mathbf{r}^{(i)}\right\|_2$ and $\left\|\mathbf{r}^{(i-1)}\right\|_2$, we need to obtain the relationship between $\left\|\mathbf{x}_{\Gamma\setminus\hat{\Gamma}^{(i)}}\right\|_2$ and $\left\|\mathbf{x}_{\Gamma\setminus\hat{\Gamma}^{(i-1)}}\right\|_2$. From the set $\tilde{\Gamma}^{(i)} \triangleq \hat{\Gamma}^{(i-1)} \cup \arg\max_k \left|(\mathbf{h}_k)^H\mathbf{r}^{(i-1)}\right|^2$ in the process of updating the estimated active user set (see step 6), we can derive the following two inequalities [9]:

$$\begin{aligned} \frac{\left\|\mathbf{x}_{\Gamma\setminus\tilde{\Gamma}^{(i)}}\right\|_2}{\sqrt{S}} &\le \frac{2\delta_{2S+1}}{(1-\delta_{2S+1})^2}\left\|\mathbf{x}_{\Gamma\setminus\hat{\Gamma}^{(i-1)}}\right\|_2 \\ &\quad + \frac{2(1+\delta_{2S+1})}{1-\delta_{2S+1}}\|\mathbf{v}\|_2, \end{aligned} \quad (13)$$

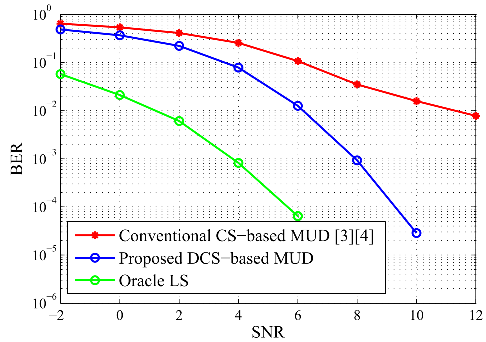

Fig. 1. BER performance comparison against SNR in uplink grant-free NOMA systems, where the overloading factor is 200%.

and

$$\left\|\mathbf{x}_{\Gamma\backslash\hat{\Gamma}^{(i)}}\right\|_2 \le \frac{1+\delta_{2S+1}}{1-\delta_{2S+1}}\left\|\mathbf{x}_{\Gamma\backslash\tilde{\Gamma}^{(i)}}\right\|_2 + \frac{2}{1-\delta_{2S+1}}\|\mathbf{v}\|_2. \quad (14)$$

By substituting (13) and (14) into (11) and (12), the termination condition $\left\|\mathbf{r}^{(i)}\right\|_2 \ge \left\|\mathbf{r}^{(i-1)}\right\|_2$ can lead to the result as

$$\|\mathbf{v}\|_2 \ge \frac{\delta_{2S+1}}{\sqrt{S}+2}\left\|\mathbf{x}_{\Gamma\backslash\hat{\Gamma}^{(i-1)}}\right\|_2 = \frac{\delta_{2S+1}}{\sqrt{S}+2}\left\|\mathbf{x}_{\Gamma\backslash\hat{\Gamma}}\right\|_2. \quad (15)$$

We can see from (15) that if the noise power is larger than that of the residue active user signals multiplied with a coefficient, the iteration process will be terminated, and the final estimated active user set can be obtained as $\hat{\Gamma} = \hat{\Gamma}^{(i-1)}$.

Then, we analyze the relationship between the reconstruction distortion $\|\hat{\mathbf{x}}-\mathbf{x}\|_2$ and $\left\|\mathbf{x}_{\Gamma\backslash\hat{\Gamma}}\right\|_2$ as

$$\begin{aligned}\|\hat{\mathbf{x}}-\mathbf{x}\|_2 &\le \left\|\mathbf{x}_{\hat{\Gamma}} - \mathbf{H}_{\hat{\Gamma}}^{\dagger}\mathbf{y}\right\|_2 + \left\|\mathbf{x}_{\Gamma\backslash\hat{\Gamma}}\right\|_2 \\ &\le \left\|\mathbf{x}_{\hat{\Gamma}} - \mathbf{H}_{\hat{\Gamma}}^{\dagger}\left(\mathbf{H}_{\Gamma}\mathbf{x}_{\Gamma}\right)\right\|_2 + \left\|\mathbf{H}_{\hat{\Gamma}}^{\dagger}\mathbf{v}\right\|_2 + \left\|\mathbf{x}_{\Gamma\backslash\hat{\Gamma}}\right\|_2 \\ &\le \frac{1}{1-\delta_{2S+1}}\left\|\mathbf{x}_{\Gamma\backslash\hat{\Gamma}}\right\|_2 + \frac{1+\delta_{2S+1}}{1-\delta_{2S+1}}\|\mathbf{v}\|_2. \end{aligned} \quad (16)$$

By substituting (16) into (15), we have

$$\|\hat{\mathbf{x}}-\mathbf{x}\|_2 \le \frac{2+\sqrt{S}+\delta_{2S+1}+\delta_{2S+1}^2}{\delta_{2S+1}(1-\delta_{2S+1})}\|\mathbf{v}\|_2, \quad (17)$$

which means that the signal reconstruction error is upper bounded by the noise power $\|\mathbf{v}\|_2$, the sparsity level S, and the RIP of the measurement matrix.

V. Simulation Results

In this section, we compare the BER performance of the proposed DCS-based MUD, the conventional CS-based MUD, i.e., the classical OMP algorithm considered in [3], and the oracle LS algorithm, where the active users are assumed to be exactly known at the BS, with QPSK modulation. Particularly, the number of users is $K = 200$, the number of orthogonal subcarriers is $N = 100$, and thus the overloading factor is 200%. The spreading matrix is designed as a Toeplitz matrix based on pseudo random noise, and thus RIP can be satisfied with a high probability [9].

Fig. 1 shows the BER performance comparison of the proposed DCS-based MUD, the conventional CS-based MUD, and the oracle LS algorithm. Particularly, the number of continuous time slots is set to $J = 7$ according to the LTE-Advanced protocol [6], and the number of active users is $S = 20$. Simulation results show that under the setting considered in this letter, the DCS-based MUD can achieve about 5 dB SNR gains over CS-based MUD when BER is 10^{-2}, which benefits from the exploiting of temporal correlation of the active user sets within several continuous time slots. Furthermore, as analyzed in Section IV, both the residue active user signals and signal reconstruction error decrease with the decrease of noise power. Therefore, the proposed DCS-based MUD has superior performance especially when SNR is high as shown in Fig. 1. Furthermore, compared to the oracle LS algorithm, the DCS-based MUD suffers from about 3 dB performance loss when BER is 10^{-4}, which is caused by the inaccurate support estimation of the greedy recovery algorithms in CS with low SNR.

VI. Conclusions

In this letter, by exploiting the temporal correlation of the active user sets, we have proposed the DCS-based MUD to realize both user activity detection and data detection for uplink NOMA systems in several continuous time slots. Unlike conventional CS-based MUD schemes that independently detect multi-user signal in each time slot, the proposed DCS-based MUD can realize simultaneous multi-user signal detection in several continuous time slots. This is achieved by using the estimated active user set in the current time slot as the initial set to estimate the transmitted signal in the next time slot under the framework of DCS. The performance analysis provides the condition of successful recovery and the upper bound of the recovery error. Simulation results show that the proposed scheme can achieve better BER performance in uplink grant-free NOMA systems.

References

[1] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[2] Z. Ding, L. Dai, and H. V. Poor, "MIMO-NOMA design for small packet transmission in the Internet of Things," *IEEE Access*, vol. 4, pp. 1393–1405, Apr. 2016.

[3] B. Shim and B. Song, "Multiuser detection via compressive sensing," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 972–974, Jul. 2012.

[4] F. Monsees, M. Woltering, C. Bockelmann, and A. Dekorsy, "Compressive sensing multi-user detection for multicarrier systems in sporadic machine type communication," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–5.

[5] A. T. Abebe and C. G. Kang, "Iterative order recursive least square estimation for exploiting frame-wise sparsity in compressive sensing-based MTC," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1018–1021, May 2016.

[6] B. Wang, L. Dai, T. Mir, and Z. Wang, "Joint user activity and data detection based on structured compressive sensing for NOMA," *IEEE Commun. Lett.*, vol. 20, no. 7, pp. 1473–1476, Jul. 2016.

[7] N. Vaswani and J. Zhan, "Recursive recovery of sparse signal sequences from compressive measurements: A review," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3523–3549, Jul. 2016.

[8] J. P. Hong, W. Choi, and B. D. Rao, "Sparsity controlled random multiple access with compressed sensing," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 998–1010, Feb. 2015.

[9] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.