# Multi-User Detection Using ADMM-Based Compressive Sensing for Uplink Grant-Free NOMA

Ali Cagatay Cirik, *Member, IEEE*, Naveen Mysore Balasubramanya, *Student Member, IEEE*, and Lutz Lampe, *Senior Member, IEEE*

***Abstract*—Non-orthogonal multiple access (NOMA) is considered a primary candidate addressing the challenge of massive connectivity in fifth generation wireless communication systems. In this letter, we propose a low-complexity NOMA mechanism with efficient multi-user detection (MUD) based on the adaptive alternating direction method of multipliers, which is able to jointly detect user activity and transmitted data. The proposed algorithm leverages the transmit symbol estimate and active user set as "prior knowledge," which can be obtained from the previous iterations/time intervals, for improved MUD performance. We demonstrate that our proposed mechanism outperforms the state-of-art MUD NOMA schemes.**

***Index Terms*—Alternative direction method of multipliers (ADMM), compressive sensing, non-orthogonal multiple access (NOMA).**

## I. Introduction

MANY broadband wireless communication systems like the third generation partnership project (3GPP) long term evolution (LTE) have adopted grant-based orthogonal channel access methods. Although this is obviously effective in avoiding inter-user interference and simplifying the receiver design, the number of supported users is limited by the number of available orthogonal resources. Moreover, the grant-based channel access procedure requires multiple messages back and forth, resulting in a large signaling overhead and high latency. Since massive user support is a key requirement of forthcoming Internet-of-Things (IoT) deployments, novel mechanisms for channel access are necessary for next-generation wireless systems. Grant-free access being proposed by the 3GPP for the fifth generation (5G) new radio (NR) is more desirable as it has the following advantages: 1) reduced transmission latency, 2) smaller signalling overhead due to the simplification of the scheduling procedure and 3) improved energy efficiency (battery life) of the user equipments (UEs) with the reduction in signalling and the ON time of the UE.

In grant-free access, due to the lack of UE scheduling on orthogonal time-frequency resources, there is a high probability that different UEs randomly choose the same resource blocks for the uplink transmission resulting in the superposition of data of multiple users (collision). Therefore, non-orthogonal multiple access (NOMA), which is a spectrally efficient and robust against collisions has been proposed to tackle the massive connectivity in 5G by allowing multiple UEs to superpose on the same resources non-orthogonally [1]. The synergy of grant-free transmission and NOMA allows more users to effectively share the same time-frequency resources simultaneously. With the improved robustness towards UE collision as the number of users increases, NOMA reduces the probability for retransmission and the delay associated with the retransmission procedure, which is beneficial especially for the UEs which have a low latency and high reliability requirements. Reducing retransmissions is also beneficial for battery constrained devices.

As with all the NOMA schemes, multi-user detection (MUD) is a key performance enabler for grant-free NOMA systems [2]–[4]. Since the UEs transmit data randomly/autonomously without any uplink grant from the base-station (eNB) in grant-free transmission, the eNB has no prior information of the transmitting UEs, and thus it has to perform blind activity/user detection. The application of MUD to the IoT scenario would require an efficient handling of the case where the number of *active* users is much smaller than the total number of users. This user sparsity enables the application of conventional compressive sensing (CS) algorithms like the orthogonal matching pursuit (OMP) and the iterative support detection (ISD) using truncated basis pursuit (BP) [5] for joint user activity and data detection. These CS algorithms have been utilized by prior works on grant-free NOMA techniques (see [2]–[4]), where MUD is realized independently for each transmission time interval. However, in IoT transmission, signals in consecutive transmission time intervals may be correlated, e.g., the transmit signal can be repeated to enhance coverage or the set of active users may remain unchanged (or change slowly) [4], which necessitates novel algorithms leveraging this correlation for better performance. Recently, a modified version of the original ISD algorithm called the structured ISD (SISD) was proposed in [6], which utilizes both the correlation across consecutive time intervals and the user sparsity aspects for improved performance over OMP and ISD. But the user activity and data detection are realized separately, leading to increased computational complexity.

Unlike the SISD algorithm which relies only on the partial active user set (the locations of non-zero values in the transmit symbol), our approach incorporates the signal value estimate from the previous time interval/iteration along with the partial

Manuscript received July 6, 2017; revised September 5, 2017; accepted September 10, 2017. Date of publication September 14, 2017; date of current version February 16, 2018. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), in part by MITACS, Canada, and in part by Sierra Wireless Inc., Richmond, BC, Canada. The associate editor coordinating the review of this paper and approving it for publication was M. A. Assaad. *(Corresponding author: Ali Cagatay Cirik.)*

The authors are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: cirik@ece.ubc.ca; naveenmb@ece.ubc.ca; lampe@ece.ubc.ca).

Digital Object Identifier 10.1109/LWC.2017.2752165

active user set as "prior knowledge", which could potentially enhance the performance. We propose an alternative direction method of multipliers (ADMM)-based MUD solution in grant-free NOMA systems, which i) scales well to large data sets, ii) has a fast convergence rate to a global optimal solution, and iii) has a low complexity [7]. We show that our approach outperforms the state-of-the-art MUD methods. An ADMM-based MUD algorithm has also been studied in [8]. However, the algorithm in [8] exploits neither the knowledge of partial support set nor the prior information of the transmitted symbols. Moreover, it does not exploit the correlation in different time-slots to enhance the MUD performance.

## II. System Model

We consider an uplink NOMA system with one BS and $K$ users, where all nodes are equipped with a single antenna. For concreteness, we assume that the system uses $N$ orthogonal subbands, where $N < K$, i.e., the overloaded system. Similar to [6], the data symbol is spread and transmitted over $N$ subbands. The received signal at the BS on the $n$-th subcarrier within time interval $t$, $t = 1, \ldots, T$ is expressed as

$$y_n(t) = \sum_{k=1}^{K} g_{nk}(t) s_{nk} x_k(t) + v_n(t), \;\; n = 1, \ldots, N, \tag{1}$$

where $x_k(t)$ is the transmit symbol for active user $k$ at time $t$ taken from a complex-constellation set $\mathcal{X}$, $s_{nk}$ is the $n$-th component of spreading sequence $\mathbf{s}_k$ of length $N$, where $x_k(t)$ is modulated onto. Furthermore, $g_{nk}(t)$ is the channel gain of user $k$ on the $n$-th subcarrier at time $t$ and it is an identically and independent distributed (i.i.d.) complex Gaussian variable with zero mean and unit variance [6]. Finally, $v_n(t)$ is the sample of additive white Gaussian noise with zero mean and variance $\sigma^2$ on subcarrier $n$ and at time $t$.

Stacking the received signals over all $N$ subcarriers, the received vector at time $t$ can be written as

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t, \;\; t = 1, \ldots, T, \tag{2}$$

where $\mathbf{y}_t = [y_1(t), y_2(t), \ldots, y_N(t)]^T$, $\mathbf{x}_t = [x_1(t), x_2(t), \ldots, x_K(t)]^T$, $\mathbf{v}_t = [v_1(t), v_2(t), \ldots, v_N(t)]^T$, and $\mathbf{H}_t \in \mathcal{C}^{N\times K}$ is an equivalent channel matrix, whose $n$-th row and the $k$-th column equals to $g_{nk}(t)s_{nk}$.

Further, we assume that the users are synchronized in a predefined frame structure consisting of multiple time intervals $T$, and that they are active or inactive in the entire frame, enabling us to jointly process $\mathbf{y}_t$, $t = 1, \ldots, T$ for robust MUD performance.[1]

## III. ADMM Algorithm

As mentioned in Section I, the IoT uplink grant-free systems that use a sparse transmit symbol vector $\mathbf{x}_t$, are amenable for CS-based MUD. The goal of CS is to reconstruct sparse signal $\mathbf{x}_t$ from the measurement vector $\mathbf{y}_t$ in (2) by solving the following basis pursuit denoising (BPDN) problem:

$$\min_{\mathbf{x}} \quad \lambda_t \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{y}_t - \mathbf{H}_t\mathbf{x}\|_2^2. \tag{3}$$

[1]Note that although the support set is same within the frame, the transmit signal from the users can vary among different time intervals.

The search for a solution that achieves exact reconstruction using fewer measurements led to the modified-CS idea by exploiting the partial known support set $\mathcal{T}$, which denotes the set of location of nonzero values in the signal $\mathbf{x}_t$. To that end, the sparse recovery problem becomes the one that tries to find the vector $\mathbf{x}_t$ that is sparsest outside the set $\mathcal{T}$, that is

$$\min_{\mathbf{x}} \quad \lambda_t \|\mathbf{x}\|_{1,\mathbf{w}} + \frac{1}{2}\|\mathbf{y}_t - \mathbf{H}_t\mathbf{x}\|_2^2, \tag{4}$$

where $\|\mathbf{x}\|_{1,\mathbf{w}} = \sum_{i=1}^{K} w_i|x_i|$ with $w_i = 0$, $i \in \mathcal{T}$, $w_i = 1$, $i \notin \mathcal{T}$. Note that the first term in (4) does not penalize the nonzero terms whose locations are known, which differs from conventional model where all terms are penalized. However, the problem (4) puts no cost on $\mathbf{x}$ except the cost imposed by the data term. Thus, when very few measurements are available or when the noise is large, $\mathbf{x}$ can become larger than required. To address this, when reliable prior signal value knowledge is available, Lu and Vaswani [9] exploit this knowledge to design improved solutions and solve the "regularized modified-BPDN" problem formulated as:

$$\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{y}_t - \mathbf{H}_t\mathbf{x}\|_2^2 + \frac{\alpha_t}{2}\left\|\mathbf{x} - \hat{\mathbf{x}}_t\right\|_2^2 + \lambda_t\|\mathbf{x}\|_{1,\mathbf{w}}, \tag{5}$$

where $\hat{\mathbf{x}}_t$ denotes an erroneous estimate (prior information) of the signal values $\mathbf{x}_t$ obtained in the previous iteration/time-slot and is used to achieve better reconstruction quality. Here, $\alpha_t$ and $\lambda_t$ are the regularization parameters.

We adopt the regularized modified BPDN algorithm (5) as our CS algorithm to solve the MUD problem in uplink grant-free systems, which exploits both the partial support and the transmitted signal estimate as prior information. However, Lu and Vaswani [9] have not proposed any method to solve regularized BPDN problem (5). To that end, in the following, we will propose an ADMM-based algorithm, which has been shown to be simple, effective, and fast convergent [7], [10]. To be able to apply the ADMM algorithm, the problem (5) can be equivalently written as

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z}} \quad & \underbrace{\frac{1}{2}\|\mathbf{y}_t - \mathbf{H}_t\mathbf{x}\|_2^2 + \frac{\alpha_t}{2}\left\|\mathbf{x} - \hat{\mathbf{x}}_t\right\|_2^2}_{f(\mathbf{x})} + \underbrace{\lambda_t\|\mathbf{z}\|_{1,w}}_{g(\mathbf{z})} \\ \text{s.t.} \quad & \mathbf{x} = \mathbf{z}. \end{aligned} \tag{6}$$

Introducing a dual variable $\mathbf{t}_t$ and a penalty parameter $\rho_t$, the augmented Lagrangian is expressed as

$$\mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{t}_t) = f(\mathbf{x}) + g(\mathbf{z}) + \mathbf{t}_t^T(\mathbf{x} - \mathbf{z}) + \frac{\rho_t}{2}\|\mathbf{x} - \mathbf{z}\|_2^2.$$

The ADMM algorithm decomposes the main optimization problem in (6) into easier and smaller local subproblems such that the optimization variables $\mathbf{x}$ and $\mathbf{z}$ can be computed separately in an alternating fashion. The scaled ADMM form of the problem (6) consists of three iterations:

$$\mathbf{x}_t^{[k+1]} = \arg\min_{\mathbf{x}}\left(f(\mathbf{x}) + \frac{\rho_t}{2}\left\|\mathbf{x} - \mathbf{z}_t^{[k]} + \mathbf{u}_t^{[k]}\right\|_2^2\right), \tag{7}$$

$$\mathbf{z}_t^{[k+1]} = \arg\min_{\mathbf{z}}\left(g(\mathbf{z}) + \frac{\rho_t}{2}\left\|\mathbf{x}_t^{[k+1]} - \mathbf{z} + \mathbf{u}_t^{[k]}\right\|_2^2\right), \tag{8}$$

$$\mathbf{u}_t^{[k+1]} = \mathbf{u}_t^{[k]} + \mathbf{x}_t^{[k+1]} - \mathbf{z}_t^{[k+1]}, \tag{9}$$

**Algorithm 1** Proposed ADMM Algorithm

1: Initialize $\mathcal{T}^{[0]} = \emptyset$, $k = 1$, $\alpha_t^{[0]} = 0$, $\hat{\mathbf{x}}_t^{[0]} = \mathbf{0}$, $\mathbf{r}_t^{[0]} = \infty$, $\mathbf{s}_t^{[0]} = \infty$, $\mathbf{z}_t^{[0]} = \mathbf{0}$, $\mathbf{u}_t^{[0]} = \mathbf{0}$, and $\mathbf{w}^{[0]}$ is a unit vector.
2: **repeat**
3: **for** $t = 1 : T$ **do**
4: Compute $\mathbf{x}_t^{[k]}$ using (10) and $\mathbf{z}_t^{[k]}$ using (11).
5: Compute the scaled dual variable $\mathbf{u}_t^{[k]}$ using (9).
6: **end for**
7: $\mathbf{z}^{[k]} = \sum_{t=1}^{T} \left|\mathbf{z}_t^{[k]}\right|$.
8: Update the support threshold $\beta^{[k]}$ using the "first significant jump" [5]. In particular, first sort $|\mathbf{z}^{[k]}|$ in the increasing order, i.e., $|z_{(1)}^{[k]}| \leq |z_{(2)}^{[k]}| \leq \ldots \leq |z_{(K)}^{[k]}|$ ($|z_{(i)}^{[k]}|$ denotes the $i$-th largest component of $\mathbf{z}^{[k]}$ by magnitude.). The first significant jump is the smallest $i$ such that

$$|z_{(i+1)}^{[k]}| - |z_{(i)}^{[k]}| > \frac{7}{N}\frac{\|\mathbf{z}^{[k]}\|_\infty}{\min(k, 6)}. \tag{12}$$

Then, we have $\beta^{[k]} = |z_{(i)}^{[k]}|$.
9: Update the support set $\mathcal{T}^{[k]} = \left\{i : |z_i^{[k]}| > \beta^{[k]}\right\}$, and the signal value $\hat{\mathbf{x}}_t^{[k]} = \mathbf{z}_t^{[k]}$. Set the weights $w_i^{[k]} = 0,\ i \in \mathcal{T}^{[k]}$, $w_i^{[k]} = 1,\ i \notin \mathcal{T}^{[k]}$.
10: Compute the primal residual $\mathbf{r}_t^{[k]} = \mathbf{x}_t^{[k]} - \mathbf{z}_t^{[k]}$, and the dual residual $\mathbf{s}_t^{[k]} = -\rho_t\left(\mathbf{z}_t^{[k]} - \mathbf{z}_t^{[k-1]}\right)$.
11: Update the regularization parameter according to:

$$\alpha_t^{[k]} = (KC_\alpha)/\left(N\left(\|\mathbf{r}_t^{[k]}\|_2 + \|\mathbf{s}_t^{[k]}\|_2\right)\right). \tag{13}$$

12: $k \leftarrow k + 1$.
13: **until** convergence or maximum number of iterations is reached.

where $\mathbf{u}_t = \frac{1}{\rho_t}\mathbf{t}_t$ is the scaled dual variable. The optimal $\mathbf{x}_t$ can be computed from (7), and it is given as

$$\mathbf{x}_t^{[k+1]} = \left(\mathbf{H}_t^H\mathbf{H}_t + (\rho_t + \alpha_t)\mathbf{I}_K\right)^{-1} \times \left(\mathbf{H}_t^H\mathbf{y}_t + \alpha_t\hat{\mathbf{x}}_t + \rho_t\left(\mathbf{z}_t^{[k]} - \mathbf{u}_t^{[k]}\right)\right). \tag{10}$$

Even though the problem (8) is not differentiable, the closed-form solution of (8) can be computed through sub-differential calculus, i.e., using the soft thresholding operator, and the optimal $\mathbf{z}_t^{[k+1]}$ is expressed as:

$$z_i^{[k+1]}(t) = \max\left\{x_i^{[k+1]}(t) + u_i^{[k]}(t) - \frac{\lambda_t w_i}{\rho_t}, 0\right\} - \max\left\{-x_i^{[k+1]}(t) - u_i^{[k]}(t) - \frac{\lambda_t w_i}{\rho_t}, 0\right\}. \tag{11}$$

The proposed ADMM algorithm is given in Algorithm 1 and the important steps of it are explained in detail below:

- In Step 7, the fact that same users are active (same support set) in $T$ consecutive time-slots is exploited. In particular, the MUD is improved by adding all the estimated signal vectors in each time interval and simultaneously updating the same support set in $T$ continuous time intervals. This leads to a better performance than the conventional algorithms in [2]–[4], which calculate the support set in each time interval independently.
- The support set detection scheme is based on thresholding used in Step 8 and the threshold is chosen by locating the "first significant jump" in the increasingly sorted sequence $\mathbf{z}$ obtained in Step 7. This amounts to sweeping the increasing sequence $\mathbf{z}$ and looking for the first jump larger than a threshold. Intuitively, the "first significant jump" scheme works well because the true nonzero entries are large in magnitude and small in number, while the false ones are large in number and small in magnitude. Therefore, the first huge difference between adjacent points appear on the boundary of the true supports and false supports [5].
- In Step 9, the support set $\mathcal{T}$ is updated using the threshold found in Step 8 and weights $w_i$ are updated based on $\mathcal{T}$.
- In Step 10, the primal and residual variables are computed to check the convergence of the ADMM algorithm [7].
- In Step 11, the regularization parameter is updated. Particularly, when $N/K$ is small (large), low (high)-dimensional measurement vector might bring less (more) information than the estimated signal $\hat{\mathbf{x}}_t$, and thus the objective function should be penalized with a larger (smaller) weight $\alpha_t$ [10].

**Algorithm 2** Dynamic ADMM Algorithm

**Require:** At time $t = 0$, solve the following problem to compute the signal value $\hat{\mathbf{x}}_0$ and support set $\mathcal{T}_0$ estimations.

$$\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{y}_0 - \mathbf{H}_0\mathbf{x}\|_2^2 + \lambda_0\|\mathbf{x}\|_1.$$

1: **for** $t > 0$ **do**
2: Set $\mathcal{T} = \mathcal{T}_{t-1}$
3: Solve the following problem using the proposed ADMM method given in Algoritm 1 for $T = 1$ to compute $\hat{\mathbf{x}}_t$ and $\mathcal{T}_t$.

$$\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{y}_t - \mathbf{H}_t\mathbf{x}\|_2^2 + \frac{\alpha_t}{2}\left\|\mathbf{x} - \hat{\mathbf{x}}_{t-1}\right\|_2^2 + \lambda_t\|\mathbf{x}\|_{1,\mathbf{w}}.$$

4: **end for**

## A. Dynamic Recovery

We now relax our system assumptions and allow users to access or leave the system within a frame. In this case, the set of active users changes slowly, which results in the temporal correlation of support sets in continuous time intervals. If the initial signal and/or support set can be recovered accurately in the current time interval, then, by exploiting such temporal correlation, the dynamic version of the proposed ADMM algorithm can be applied recursively to compute the signal and support set in the next time interval as shown in Algorithm 2. Note that unlike the shared support set discussed in Algorithm 1, where $T$ time intervals are jointly processed, here each time interval is processed independently.

# IV. Simulation Results

In this section, we examine the MUD performance of the proposed ADMM algorithm. The number of users and orthogonal subcarriers are $K = 108$ and $N = 72$, respectively. We assume that 12 users out of 108 are active and choose the number of time intervals $T = 7$, similar to [6].[2] Spreading sequences are generated based on the pseudo random noise

[2]The MUD performance deteriorates as the number of active users increases.

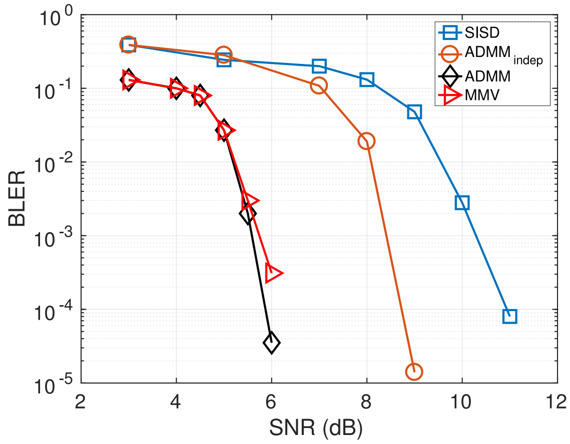


Fig. 1. BLER comparison of different algorithms.

sequence. In-line with the current LTE systems, we use convolution turbo code (CTC) of rate $1/3$ and quaternary phase shift keying (QPSK) as the modulation scheme. The constant $C_\alpha$ in (13), maximum number of iterations and Lagrangian parameter $\rho_t$ are set to 0.1, 100 and 1, respectively.

Fig. 1 shows the block error rate (BLER) performance of SISD [6], multiple measurement vectors (MMV) [11],[3] and a version of the proposed algorithm, $\text{ADMM}_{\text{indep}}$, which processes each time interval independently. We notice that our ADMM scheme outperforms SISD as it utilizes both the partial support set and the signal estimate. Moreover, we observe that the ADMM algorithm slightly outperforms the MMV algorithm. Finally, comparing ADMM with $\text{ADMM}_{\text{indep}}$, it is seen that the structured processing of multiple time intervals brings gains over the independent time interval processing.

Table I presents the average central processing unit (CPU) time that is consumed only by the CS algorithms (excluding the LTE and CTC processing) in recovering the transmit signals per block for different measurement rate $N/K$ with fixed $N = 72$ at received signal-to-noise ratio (SNR) of 5 dB. It is shown that the proposed ADMM algorithm has far less time complexity than SISD algorithm. The ADMM algorithm has a lower complexity than SISD because while the user activity and data detection are separately implemented in an iterative fashion in the SISD algorithm, the user activity detection is included into the data detection as one single step in the ADMM algorithm and thus reduces the computational complexity [10]. As analyzed in [12], MMV and ADMM algorithms have linear- and cubic-time computational complexities per iteration with respect to the number of users $K$. However, our simulations suggest that the proposed ADMM requires around 10 iterations to convergence while MMV requires hundreds of iterations. Therefore, when $K$ is low, ADMM has a lower CPU time due to the fast-convergence rate. On the other hand, when $K$ increases the cubic-time complexity of the ADMM algorithm dominates the CPU time, and MMV runs faster. Therefore, we propose to employ our ADMM-based MUD algorithm when $N/K$ is larger than 0.5. However, the MMV algorithm requires the channel and the active user set to remain the same for all time intervals, which is difficult to observe in high mobility scenarios, while the proposed ADMM algorithm is suitable with both varying channels and user sets in each time interval.

[3]The MATLAB codes for SISD and MMV were downloaded from http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html and http://dsp.ucsd.edu/~zhilin/MFOCUSS.m, respectively.

TABLE I
AVERAGE CPU TIME PER BLOCK (IN SECONDS)

| $N/K$ | 0.3 | 0.4 | 0.6 | 0.8 |
|---|---|---|---|---|
| ADMM | 36.5 | 18.4 | 7.4 | 4.2 |
| SISD | 162.3 | 148.4 | 141.6 | 52.3 |
| MMV | 22.8 | 19.6 | 17.4 | 13.2 |

## V. CONCLUSION

In this letter, we have considered the MUD performance for uplink grant-free NOMA systems and proposed an efficient ADMM-based algorithm to jointly realize user activity and data detection. Simulation results show that the proposed ADMM algorithm, which exploits both transmit symbol estimate and active user set as prior information, provides improved MUD performance along with faster convergence when compared with the state-of-the-art SISD algorithm. Moreover, it has a similar MUD performance as the MMV algorithm. Considering its complexity, we conclude that it is beneficial to employ the proposed ADMM algorithm when the number of users in the system is relatively low.

## REFERENCES


[1] L. Dai *et al.*, "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] A. Bayesteh, E. Yi, H. Nikopour, and H. Baligh, "Blind detection of SCMA for uplink grant-free multiple-access," in *Proc. ISWCS*, Barcelona, Spain, Aug. 2014, pp. 853–857.

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

[4] 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.

[5] Y. Wang and W. Yin, "Sparse signal reconstruction via iterative support detection," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 462–491, 2010.

[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] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[8] Q. He, T. Q. S. Quek, Z. Chen, and S. Li, "Compressive channel estimation and multi-user detection in C-RAN," in *Proc. ICC*, Paris, France, May 2017, pp. 1–6.

[9] W. Lu and N. Vaswani, "Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 182–196, Jan. 2012.

[10] H. Liu, B. Song, H. Qin, and Z. Qiu, "An adaptive-ADMM algorithm with support and signal value detection for compressed sensing," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 315–318, Apr. 2013.

[11] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[12] Z. Chen, R. Molina, and A. K. Katsaggelos, "Robust recovery of temporally smooth signals from under-determined multiple measurements," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1779–1791, Apr. 2015.