Compressive Sensing Based Massive Access for IoT Relying on Media Modulation Aided Machine Type Communications

Li Qiao, Jun Zhang, Zhen Gao, Sheng Chen, *Fellow, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—A fundamental challenge of the large-scale Internet-of-Things lies in how to support massive machine-type communications (mMTC). This letter proposes a media modulation based mMTC solution for increasing the throughput, where a massive multi-input multi-output based base station (BS) is used for enhancing the detection performance. For such a mMTC scenario, the reliable active device detection and data decoding pose a serious challenge. By leveraging the sparsity of the uplink access signals of mMTC received at the BS, a compressive sensing based massive access solution is proposed for tackling this challenge. Specifically, we propose a structured orthogonal matching pursuit algorithm for detecting the active devices, whereby the block-sparsity of the uplink access signals exhibited across the successive time slots and the structured sparsity of media-modulated symbols are exploited for enhancing the detection performance. Moreover, a successive interference cancellation based structured subspace pursuit algorithm is conceived for data demodulation of the active devices, whereby the structured sparsity of media modulation based symbols found in each time slot is exploited for improving the detection performance. Finally, our simulation results verify the superiority of the proposed scheme over state-of-the-art solutions.**

***Index Terms*—Internet-of-Things, massive machine type communications, massive access, massive multi-input multi-output, media modulation, compressive sensing.**

I. Introduction

The emerging paradigm of massive machine-type communications (mMTC) is identified as an indispensable component for enabling the massive access of machine-type devices (MTDs) in the emerging Internet-of-Things (IoT) [1]. In stark contrast to conventional human-centric mobile communications, mMTC focuses on uplink-oriented communications serving massive MTDs and exhibits sporadic tele-traffic requiring low-latency and high-reliability massive access [1].

The conventional grant-based access approach relies on complex time and frequency-domain resource allocation before data transmission, which would impose prohibitive signaling overhead and latency on massive mMTC [1]. To support low-power MTDs at low latency, the emerging grant-free approach has attracted significant attention for massive access, since it simplifies the access procedure by directly delivering data without scheduling [2]–[7]. Specifically, by exploiting the block-sparsity of mMTC, the authors of [3] and [4] proposed compressive sensing (CS) solutions for joint active device and data detection, while a maximum a *posteriori* probability based scheme was proposed in [2] for improving the performance attained. Furthermore, MTDs having slowly-varying activity tend to exhibit partial block sparsity, hence a modified orthogonal matching pursuit solution was conceived in [5], while a modified subspace pursuit algorithm was proposed in [6]. It was shown that the previously detected results can be exploited for enhancing the following detection. However, the contributions [2]–[6] only consider single-antenna configurations at both the MTDs and the BS. To achieve higher efficiency and more reliable detection, multi-antenna-aided MTDs using spatial modulation (SM) and massive multi-input multi-output (mMIMO) were considered in [7], [8] at the BS, where a two-level sparse structure based CS (TLSSCS) detector and a structured CS detector was proposed in [7] and [8], respectively. However, increasing the data rate of SM by one bit requires doubling the number of transmit antennas (TAs) [9], [10], which violates the low-cost requirement of MTDs. To improve the uplink (UL) throughput at a low cost and power-consumption, the authors of [11], [12] proposed to employ media modulation at the MTDs, where an iterative interference cancellation detector and a CS detector was employed for multi-user detection in [11] and [12], respectively. However, these authors have not considered active user detection (AUD). To sum up, we provide a brief comparison of the related literature in Table I.

TABLE I
A Brief Comparison of the Related Literature

Contents \ Literature		[2]-[6]	[7]	[8]	[11]	[12]
BS	Single antenna	✓				
	mMIMO		✓	✓	✓	✓
MTDs	Single antenna	✓				
	SM		✓	✓		
	Media modulation				✓	✓
AUD		✓	✓			
Data detection		✓	✓	✓	✓	✓

Against this background, we propose to adopt media modulation at the MTDs for improving the UL throughput and to employ a mMIMO scheme at the BS. Moreover, a CS-based active device and data detection solution is proposed by exploiting both the sporadic traffic and the block-sparsity of mMTC as well as the structured sparsity of media-modulated symbols. Specifically, we first propose a structured orthogonal matching pursuit (StrOMP) algorithm for AUD, where the block-sparsity of UL access signals across the successive time slots and the structured sparsity of media-modulated symbols are exploited. Additionally, a successive interference cancellation

Manuscript received November 14, 2019; revised May 21, 2020; accepted June 28, 2020. Date of publication July 1, 2020; date of current version October 13, 2020. This work was supported in part by NSFC under Grants 61827901 and 61701027, in part by the Beijing Municipal Natural Science Foundation under Grants 4182055 and L182024, in part by the Young Elite Scientists Sponsorship Program by CAST, in part by the Talent Innovation Project of BIT, and in part by the Open Research Fund of the Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Ministry of Industry and Information Technology, Nanjing University of Aeronautics and Astronautics (NUAA), under the Grant KF20202103. The work of Lajos Hanzo was supported in part by the Engineering and Physical Sciences Research Council projects EP/N004558/1, EP/P034284/1, EP/P034284/1, EP/P003990/1 (COALESCE), of the Royal Society's Global Challenges Research Fund Grant and in part by the European Research Council's Advanced Fellow Grant QuantCom. The review of this article was coordinated by Dr. H. Zhu. *(Corresponding author: Jun Zhang.)*

Li Qiao and Jun Zhang are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: qiaoli@bit.edu.cn; buaazhangjun@vip.sina.com).

Zhen Gao is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China, with the Advanced Research Institute of Multidisciplinary Science (ARIMS), BIT, Beijing 100081, China, and also with the Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Ministry of Industry and Information Technology, NUAA, Nanjing 210016, China (e-mail: gaozhen16@bit.edu.cn).

Sheng Chen is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K., and also with the King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: sqc@ecs.soton.ac.uk).

Lajos Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/TVT.2020.3006318

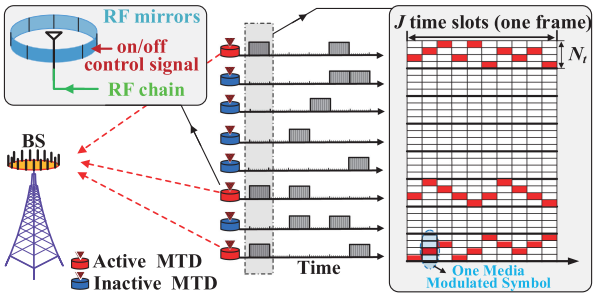

Fig. 1. Proposed media modulation based mMTC scheme, where the UL access signal exhibits block-sparsity in a frame and structured sparsity in each time slot.

based structured subspace pursuit (SIC-SSP) algorithm is proposed for demodulating the detected active MTDs, where the structured sparsity of media-modulated symbols in each time slot is exploited for enhancing the decoding performance.

Note that the proposed StrOMP and SIC-SSP algorithms belong to the family of greedy algorithms. As a benefit of their near-optimal performance attained at a low complexity, greedy algorithms have been popularly used in mMTC scenarios [2]–[7], [13]. Finally, our simulation results verify the superiority of the proposed scheme over cutting-edge benchmarks.

Notation: Boldface lower and upper-case symbols denote column vectors and matrices, respectively. For a matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$, $\|\mathbf{A}\|_F$, $\mathbf{A}_{[m,n]}$ denote the transpose, Hermitian transpose, pseudo-inverse, Frobenius norm, the m-th row and n-th column element of $\mathbf{A}$, respectively. $\mathbf{A}_{[\Omega,:]}$ ($\mathbf{A}_{[:,\Omega]}$) is the sub-matrix containing the rows (columns) of $\mathbf{A}$ indexed in the ordered set Ω. $\mathbf{A}_{[\Omega,m]}$ is the m-th column of $\mathbf{A}_{[\Omega,:]}$. For a vector $\mathbf{x}$, $\|\mathbf{x}\|_p$, $[\mathbf{x}]_m$, $[\mathbf{x}]_{m:n}$ and $[\mathbf{x}]_\Omega$ are the l_p norm, m-th element, m-th to n-th elements, and entries indexed in the ordered set Ω of $\mathbf{x}$, respectively. For an ordered set Γ and its subset Ω, $|\Gamma|_c$, $\Gamma[m]$, and $\Gamma \setminus \Omega$ are the cardinality of Γ, m-th element of Γ, and complement of subset Ω in Γ, respectively. $[K]$ is the set $\{1, 2, \ldots, K\}$.

II. System Model

We first introduce the proposed media modulation based mMTC scheme and then focus on our massive access technique relying on joint active device and data detection at the BS.

A. Proposed Media Modulation Based mMTC Scheme

As illustrated in Fig. 1, we propose that all K MTDs adopt media modulation for enhanced UL throughput and the BS employs mMIMO using N_r receive antenna elements for reliable massive access. In the UL, each symbol consists of the conventional modulated symbol and of the media-modulated symbol, and each MTD relies on a single conventional antenna and M_r extra radio frequency (RF) mirrors [11], [12], [14]–[16]. By adjusting the binary on/off status of the M_r RF mirrors, we have $N_t = 2^{M_r}$ mirror activation patterns (MAPs), and the media-modulated symbol is obtained by mapping $\log_2(N_t) = M_r$ bits to one of the N_t MAPs. Therefore, if the conventional M-QAM symbol is adopted, the overall UL throughput of an MTD is $\eta = M_r + \log_2 M$ bit per channel use (bpcu). By contrast, to convey extra bits, SM relying on a single RF chain and multiple TAs will activate one of the TAs for UL transmission [9], [10]. To elaborate a little further, to achieve the same extra throughput, media modulation only requires a single UL TA and a linearly increasing number of RF mirrors, while SM requires an exponentially increasing number of TAs [9]–[12], [14]–[16]. Clearly, media modulation is more attractive for mMTC owing to its increased UL throughput, which is achieved at a negligible power consumption and hardware cost [14]–[16]. Moreover, using a mMIMO UL receiver is the most compelling technique. By leveraging the substantial diversity gain gleaned from hundreds of antennas, the mMIMO BS is expected to achieve high-reliability UL multi-user detection, in the context of mMTC. By integrating the complementary benefits of media modulation at the MTDs and mMIMO reception at the BS into mMTC, we arrive at an excellent solution.

B. Massive Access of the Proposed mMTC Scheme

As shown in Fig. 1, we assume that the activity patterns of the K MTDs remain unchanged in a frame, which consists of J successive time slots. Hence we only focus our attention on the massive access for a given frame. Specifically, the signal received at the BS in the j-th ($\forall j \in [J]$) time slot, denoted by $\mathbf{y}^j \in \mathbb{C}^{N_r \times 1}$, can be expressed as

$$\mathbf{y}^j = \sum_{k=1}^{K} a_k g_k^j \mathbf{H}_k \mathbf{d}_k^j + \mathbf{w}^j = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{x}_k^j + \mathbf{w}^j = \mathbf{H}\widetilde{\mathbf{x}}^j + \mathbf{w}^j, \tag{1}$$

where the activity indicator a_k is set to one (zero) if the k-th MTD is active (inactive), while $g_k^j \in \mathbb{C}$, $\mathbf{d}_k^j \in \mathbb{C}^{N_t \times 1}$, and $\mathbf{x}_k^j = a_k g_k^j \mathbf{d}_k^j \in \mathbb{C}^{N_t \times 1}$ are the conventional modulated symbol, media-modulated symbol, and equivalent UL access symbol of the k-th MTD in the j-th time slot, respectively. Furthermore, $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is the multi-input multi-output (MIMO) channel matrix associated with the k-th MTD, $\mathbf{w}^j \in \mathbb{C}^{N_r \times 1}$ is the noise with elements obeying the independent and identically distributed (i.i.d.) complex Gaussian distribution $\mathcal{CN}(0, \sigma_w^2)$, while $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_K] \in \mathbb{C}^{N_r \times (KN_t)}$ and $\widetilde{\mathbf{x}}^j = [(\mathbf{x}_1^j)^T, (\mathbf{x}_2^j)^T, \ldots, (\mathbf{x}_K^j)^T]^T \in \mathbb{C}^{(KN_t) \times 1}$ are the aggregate MIMO channel matrix and UL access signal in the j-th time slot, respectively.

Note that for any $\mathbf{d}_k^j$ given $\forall j \in [J]$ and $\forall k \in [K]$, only one of its entries is one and the others are all zeros, i.e.,

$$\mathrm{supp}\{\mathbf{d}_k^j\} \in [N_t], \ \| \mathbf{d}_k^j \|_0 = 1, \ \| \mathbf{d}_k^j \|_2 = 1, \tag{2}$$

where $\mathrm{supp}\{\cdot\}$ is the support set of its argument. Furthermore, we consider the Rayleigh MIMO channel model, hence the elements in $\mathbf{H}_k$ for $\forall k \in [K]$ follow the i.i.d. complex Gaussian distribution $\mathcal{CN}(0, 1)$. We assume that the channels remain time-invariant for a relatively long period in typical IoT scenarios, hence $\{\mathbf{H}_k\}_{k=1}^K$ can be accurately estimated at the BS via periodic updates.

III. Proposed CS-Based Massive Access Solution

In typical IoT scenarios, the MTDs generate sporadic tele-traffic [2]–[7], which indicates that $\mathbf{a} = [a_1, a_2, \ldots, a_K]^T \in \mathbb{C}^{K \times 1}$ is a sparse vector and $K_a = \| \mathbf{a} \|_0 \ll K$. Moreover, this activity pattern exhibits the block-sparsity, since $\mathbf{a}$ typically remains unchanged in J successive time slots within a frame [2]–[4], [7]. Furthermore, $\mathbf{x}_k^j = a_k g_k^j \mathbf{d}_k^j$ for $\forall j \in [J]$ exhibits the structured sparsity [11], [12], due to the sparse nature of media-modulated symbols' feature as illustrated in (2). The block-sparsity and structured sparsity of the UL signals inspire us to invoke CS theory to detect the active devices and demodulate the data at the BS.

To exploit the block-sparsity of active MTD patterns, we first rewrite the received signals within a frame as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}, \tag{3}$$

where we have $\mathbf{Y} = [\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^J] \in \mathbb{C}^{N_r \times J}$, $\mathbf{H} \in \mathbb{C}^{N_r \times (KN_t)}$, $\mathbf{X} = [\widetilde{\mathbf{x}}^1, \widetilde{\mathbf{x}}^2, \ldots, \widetilde{\mathbf{x}}^J] \in \mathbb{C}^{(KN_t) \times J}$, and $\mathbf{W} = [\mathbf{w}^1, \mathbf{w}^2, \ldots, \mathbf{w}^J] \in \mathbb{C}^{N_r \times J}$. Thus the massive access problem can be formulated as the following

Algorithm 1: Proposed StrOMP Algorithm.

Input: $\mathbf{Y} \in \mathbb{C}^{N_r \times J}$, $\mathbf{H} \in \mathbb{C}^{N_r \times (KN_t)}$, and threshold P_{th}.
Output: The index set of estimated active MTDs $\Gamma \subseteq [K]$, $\widehat{K_a} = |\Gamma|_c$.
1 **Initialization**: The iterative index i=1, the residual matrix $\mathbf{R}^{(0)}$=$\mathbf{Y}$, $\Gamma^{(0)}$=$\emptyset$. We define $\mathbf{m} \in \mathbb{C}^{K\times 1}$ as an intermediate block correlation variable. For possible active MTDs given their temporary index set Λ, their MAP's index set is $\widetilde{\Lambda}$=$\{\widetilde{\Lambda}_n\}_{n=1}^{|\Lambda|_c}$, where $\widetilde{\Lambda}_n$=$\{N_t(\Lambda[n]-1)+u\}_{u=1}^{N_t}$ is the MAP's index set of the n-th MTD in Λ for $n \in [|\Lambda|_c]$;
2 **while** 1 **do**
3 $\quad [\mathbf{m}]_k = \sum_{l=(k-1)N_t+1}^{kN_t} \sum_{j=1}^{J} |(\mathbf{H}_{[:,l]})^H \mathbf{R}^{(i-1)}_{[:,j]}|^2$, for $k \in [K]$;
4 $\quad k^\star = \arg\max_{\widehat{k}\in[K]} [\mathbf{m}]_{\widehat{k}}$;
5 $\quad \Lambda = \Gamma^{(i-1)} \cup k^\star$; {Possible support estimate}
6 $\quad \mathbf{B}_{[\widetilde{\Lambda},:]} = (\mathbf{H}_{[:,\widetilde{\Lambda}]})^\dagger \mathbf{Y}, \mathbf{B}_{[[KN_t]\setminus\widetilde{\Lambda},:]} = 0$;{Coarse signal estimate via LS}
7 $\quad \eta^\star_{n,j} = \arg\max_{\widehat{\eta}_{n,j}\in\widetilde{\Lambda}_n} |\mathbf{B}_{[\widehat{\eta}_{n,j},j]}|^2$, for $n \in [|\Lambda|_c]$, $j \in [J]$;
8 $\quad \Omega^{(j)} = \{\eta^\star_{n,j}\}_{n=1}^{|\Lambda|_c}$, for $j \in [J]$;
9 $\quad \mathbf{A}_{[\Omega^{(j)},j]} = (\mathbf{H}_{[:,\Omega^{(j)}]})^\dagger \mathbf{Y}_{[:,j]}, \mathbf{A}_{[[KN_t]\setminus\Omega^{(j)},j]} = 0$, for $j \in [J]$; {Fine signal estimate via LS}
10 $\quad \mathbf{R}^{(i)} = \mathbf{Y} - \mathbf{H}\mathbf{A}$; {Residue Update}
11 $\quad$ **if** $\left\|\mathbf{R}^{(i-1)}\right\|_F - \left\|\mathbf{R}^{(i)}\right\|_F < P_{\text{th}}$ **then**
12 $\quad\quad$ **break**; {Terminates the while-loop}
13 $\quad$ **else**
14 $\quad\quad \Gamma^{(i)} = \Lambda$; {Support estimate update}
15 $\quad\quad i = i + 1$;
16 $\quad$ **end**
17 **end**
Result: $\Gamma = \Gamma^{(i-1)}$, $\widehat{K_a} = |\Gamma|_c$.

optimization problem

$$\min_{\mathbf{X}} \| \mathbf{Y} - \mathbf{H}\mathbf{X} \|_F^2 = \min_{\{\widetilde{\mathbf{x}}^j\}_{j=1}^J} \sum_{j=1}^{J} \| \mathbf{y}^j - \mathbf{H}\widetilde{\mathbf{x}}^j \|_2^2$$

$$= \min_{\{a_k, \mathbf{d}_k^j, g_k^j\}_{j=1,k=1}^{J,K}} \sum_{j=1}^{J} \| \mathbf{y}^j - \sum_{k=1}^{K} a_k g_k^j \mathbf{H}_k \mathbf{d}_k^j \|_2^2$$

$$\text{s.t. (2) and } \|\mathbf{a}\|_0 \ll K. \tag{4}$$

In the following subsections, we will first utilize the proposed StrOMP algorithm to determine the indices of active devices. On that basis, the associated data is further detected based on the proposed SIC-SSP algorithm. Finally, we will discuss the computational complexity of the proposed algorithms.

A. Proposed StrOMP Algorithm for AUD

The proposed StrOMP procedure listed in Algorithm 1 was evolved from the orthogonal matching pursuit (OMP) algorithm of [17]. Specifically, line 3 calculates the sum correlation $\mathbf{m}$ associated with all N_t MAPs in J time slots for each MTD; line 5 combines $k^\star$ (i.e., the most likely active MTD) with $\Gamma^{(i-1)}$ to update the possible support set Λ; in line 6, the coarse signal estimate is obtained by the least squares (LS) algorithm; lines 7∼9 exploit the structured sparsity of media-modulated symbols to estimate the possible MAPs based on the coarsely estimated signal $\mathbf{B}$, and then the fine signal estimate is obtained in line 9 for enhancing the robustness to noise; line 10 updates the residual by using the finely estimated signal $\mathbf{A}$. In line 11, if the energy difference of the residual in adjacent iterations $\|\mathbf{R}^{(i-1)}\|_F - \|\mathbf{R}^{(i)}\|_F$ falls below a predefined threshold, the loop is curtailed, otherwise the iteration continues.

The classical OMP algorithm requires the sparsity level K_a, whereas the proposed StrOMP algorithm adaptively acquires the number of active MTDs without knowing K_a. Compared to the OMP algorithm, the proposed StrOMP achieves an improved detection performance by exploiting the block-sparsity (line 3) and the structured sparsity (lines 7∼9) of the UL signals.

Algorithm 2: Proposed SIC-SSP Algorithm.

Input: $\mathbf{Y} = [\mathbf{y}^1, \mathbf{y}^2, ..., \mathbf{y}^J] \in \mathbb{C}^{N_r \times J}$, $\mathbf{H} \in \mathbb{C}^{N_r \times (KN_t)}$, and the output of Algorithm 1: Γ, $\widehat{K_a}$.
Output: Reconstructed UL access signal $\mathbf{X} = [\widetilde{\mathbf{x}}^1, \widetilde{\mathbf{x}}^2, ..., \widetilde{\mathbf{x}}^J]$.
1 **for** $j = 1 : J$ **do**
2 $\quad$ **for** $s = 1 : \widehat{K_a}$ **do**
3 $\quad\quad$ **if** $s = 1$ **then**
4 $\quad\quad\quad \mathbf{v} = \mathbf{y}^j$, $\Lambda = \Gamma$, where $\mathbf{v}$ is the measurement vector and Λ is the remaining set of MTDs to be decoded, and the definitions of $\widetilde{\Lambda}$ and $\widetilde{\Lambda}_n$ are the same as those in Algorithm 1;{Initialization}
5 $\quad\quad$ **end**
6 $\quad\quad i = 1$, $\Psi^{(0)} = \emptyset$, $\mathbf{r}^{(0)} = \mathbf{v}$; {Initialization}
7 $\quad\quad$ **while** 1 **do**
8 $\quad\quad\quad [\mathbf{p}]_{\widetilde{\Lambda}} = (\mathbf{H}_{[:,\widetilde{\Lambda}]})^H \mathbf{r}^{(i-1)}$, $[\mathbf{p}]_{[KN_t]\setminus\widetilde{\Lambda}} = 0$; {Correlation}
9 $\quad\quad\quad \tau_n^\star = \arg\max_{\widehat{\tau}_n \in \widetilde{\Lambda}_n} |[\mathbf{p}]_{\widehat{\tau}_n}|^2$, for $n \in [|\Lambda|_c]$;
10 $\quad\quad\quad \Omega = \{\tau_n^\star + (\Lambda[n]-1)N_t\}_{n=1}^{|\Lambda|_c}$;{$|\Lambda|_c$ most likely MAPs}
11 $\quad\quad\quad \Omega' = \Omega \cup \Psi^{(i-1)}$;{Preliminary support estimate}
12 $\quad\quad\quad [\mathbf{e}]_{\Omega'} = (\mathbf{H}_{[:,\Omega']})^\dagger \mathbf{r}^{(0)}$, $[\mathbf{e}]_{[KN_t]\setminus\Omega'} = 0$;{Coarse LS}
13 $\quad\quad\quad \eta_n^\star = \arg\max_{\widehat{\eta}_n \in \widetilde{\Lambda}_n} |[\mathbf{e}]_{\widehat{\eta}_n}|^2$, for $n \in [|\Lambda|_c]$;
14 $\quad\quad\quad \Psi^{(i)} = \{\eta_n^\star + (\Lambda[n]-1)N_t\}_{n=1}^{|\Lambda|_c}$; {Pruning support set}
15 $\quad\quad\quad [\mathbf{e}]_{\Psi^{(i)}} = (\mathbf{H}_{[:,\Psi^{(i)}]})^\dagger \mathbf{r}^{(0)}$,$[\mathbf{e}]_{[KN_t]\setminus\Psi^{(i)}} = 0$;{Fine LS}
16 $\quad\quad\quad \mathbf{r}^{(i)} = \mathbf{r}^{(0)} - \mathbf{H}\mathbf{e}$;{Residue Update}
17 $\quad\quad\quad$ **if** $i \geq \widehat{K_a}$ or $\Psi^{(i)} = \Psi^{(i-1)}$ **then**
18 $\quad\quad\quad\quad \Psi = \Psi^{(i)}$, $n^\star = \arg\max_{\widehat{n}\in[|\Lambda|_c]} |[\mathbf{e}]_{\Psi[\widehat{n}]}|^2$;
19 $\quad\quad\quad\quad \mathbf{v} = \mathbf{v} - \mathbf{H}_{[:,\Psi[n^\star]]}[\mathbf{e}]_{\Psi[n^\star]}$;{Measurement vector update}
20 $\quad\quad\quad\quad [\widetilde{\mathbf{x}}^j]_{\Psi[n^\star]} = [\mathbf{e}]_{\Psi[n^\star]}$, $\Lambda = \Lambda \setminus \{\Lambda[n^\star]\}$;
21 $\quad\quad\quad\quad$ **break**; {Terminates the while-loop}
22 $\quad\quad\quad$ **end**
23 $\quad\quad\quad i = i + 1$;
24 $\quad\quad$ **end**
25 $\quad$ **end**
26 **end**
Result: $\mathbf{X} = [\widetilde{\mathbf{x}}^1, \widetilde{\mathbf{x}}^2, ..., \widetilde{\mathbf{x}}^J]$.

B. The SIC-SSP Algorithm Proposed for Data Detection

Based on the estimated active MTDs Γ obtained from Algorithm 1, the data detection problem in formula (4) reduces to the same CS problem as in [8] (i.e., Eq. (10) for $J = 1$ in [8]), which can be solved by the group subspace pursuit (GSP) algorithm of [8]. To further improve the performance, the proposed SIC-SSP algorithm, as listed in Algorithm 2, intrinsically integrates the idea of successive interference cancellation (SIC) with the GSP algorithm. Specifically, the outer for-loop recovers $\{\widetilde{\mathbf{x}}^j\}_{j=1}^J$ separately. For each $\widetilde{\mathbf{x}}^j$ with $j \in [J]$, the inner for-loop recovers a structured sparse signal with $\widehat{K_a}$ sparsity by performing $(\widehat{K_a} - 1)$ SIC operations. In contrast to the existing GSP algorithm, the inner for-loop of the proposed algorithm incorporates the SIC operation (lines 17∼22). Specifically, line 18 selects the index of the maximum element of the finely estimated signal $\mathbf{e}$ and subsequently line 19 eliminates it from the measurement vector $\mathbf{v}$; line 20 records the maximum element in $\widetilde{\mathbf{x}}^j$ ($j \in [J]$) and reduces the size of the remaining set of active MTDs Λ by 1, which corresponds to reducing the column dimension of the channel matrix in the next iteration for improving the data detection performance. Moreover, lines 9 and 13 improve the performance by exploiting the signal's structured sparsity. Finally, the algorithm is terminated when $\mathbf{X}$ is fully reconstructed.

C. Computational Complexity

1) The computational complexity of the proposed StrOMP algorithm (**Algorithm 1**) in the i-th iteration mainly depends on the following operations.
Signal correlation (line 3): The matrix multiplication used has a complexity order of $\mathcal{O}(JKN_tN_r)$.
Coarse signal estimate via LS (line 6): The coarse LS solution has a complexity order of $\mathcal{O}[J(2N_r(iN_t)^2 + (iN_t)^3)]$.
Fine signal estimate via LS (line 9): The fine-grained LS solution has a complexity order of $\mathcal{O}[J(2N_r i^2 + i^3)]$.

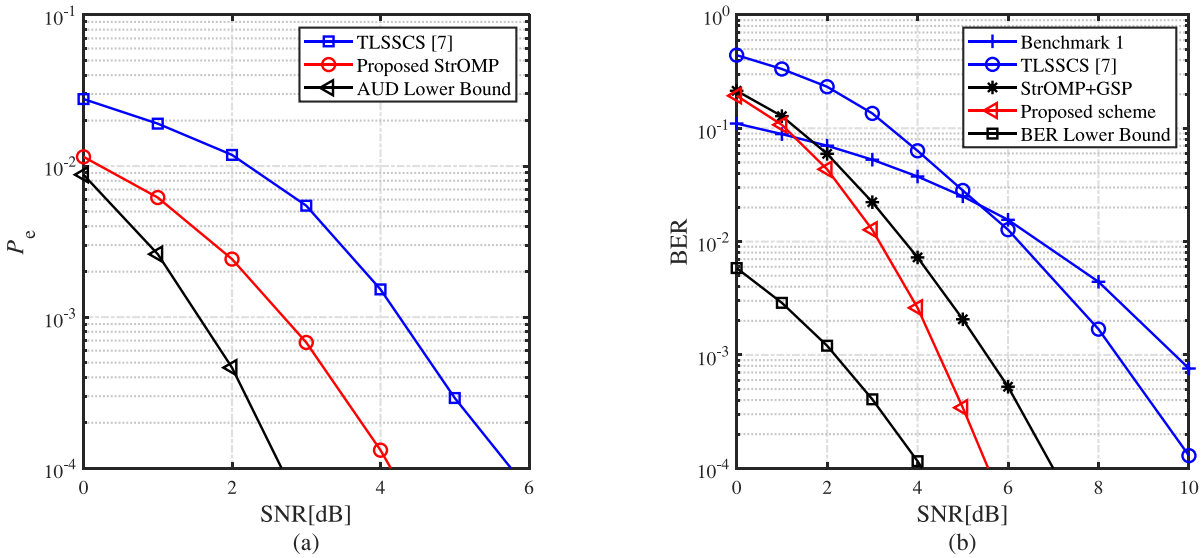

Fig. 2. Performance comparison of different solutions versus the SNR ($N_r = 50$, $J = 12$): (a) AUD performance; (b) BER performance.

Residue update (line 10): Since signal $\mathbf{A}$ acquired in line 9 is represented by a sparse matrix, the complexity of computing the residual is $\mathcal{O}(JN_r i)$.

2) The computational complexity of the proposed SIC-SSP algorithm (**Algorithm 2**) in the s-th ($1 \leq s \leq \widehat{K_a}$) inner for-loop mainly depends on the following operations.
Correlation (line 8): The matrix multiplication involved has a complexity order of $\mathcal{O}[(\widehat{K_a} - s + 1)N_t N_r]$.
Coarse LS (line 12): The coarse LS solution has a complexity order of $\mathcal{O}[2N_r(2(\widehat{K_a} - s + 1))^2 + (2(\widehat{K_a} - s + 1))^3]$.
Fine LS (line 15): The fine LS solution has a complexity order of $\mathcal{O}[2N_r(\widehat{K_a} - s + 1)^2 + (\widehat{K_a} - s + 1)^3]$.
Residue update (line 16): The complexity of computing the residual is $\mathcal{O}[(\widehat{K_a} - s + 1)N_r]$.

IV. Simulation Results

Let us now evaluate the probability of AUD error rate (P_e) and the bit error rate (BER) for the proposed CS-based massive access solution. Here we have $P_e = \frac{E_u + E_f}{K}$, and $\text{BER} = \frac{E_u J\eta + B_m + B_c}{K_a J\eta}$, where E_u is the number of active MTDs missed by activity detection, E_f is the number of falsely detected inactive MTDs, B_m and B_c are the total number of error bits in the media-modulated symbols and conventional symbols for detected active MTDs within a frame, respectively, and $K_a J\eta$ is the total number of bits transmitted by K_a active MTDs within a frame. In our simulations, the total number of MTDs is $K = 100$ with $K_a = 8$ active MTDs. Furthermore, each media modulation based MTD adopts $M_r = 2$ RF mirrors and 4-QAM ($M = 4$), hence the overall throughput becomes $\eta = M_r + \log_2 M = 4$ bpcu. Finally, P_{th} in the proposed StrOMP algorithm is set to 2, which is selected experimentally.

For comparison, we consider the following benchmarks. **Benchmark 1** : Zero forcing multi-user detector for the traditional mMIMO UL [8] supporting K_a single-antenna users adopting 16-QAM to achieve the same 4 bpcu. **TLSSCS** : TLSSCS detector from [7], and using the scaling factor of $\alpha = 4$ (i.e., α in Eq. (6) of [7]). **StrOMP+GSP** : The proposed StrOMP algorithm and the existing GSP algorithm of [8] are successively used to detect the active MTDs and the data. **AUD lower bound** : A modified StrOMP algorithm relying on the perfect knowledge of K_a, which performs the iterations including lines 3∼10 and lines 14∼15 for K_a times, and the estimated output support set is $\Gamma^{(K_a)}$ containing K_a elements. **BER lower bound** : The Oracle LS based detector relying on the perfect known index set of active MTDs and the support set of media-modulated symbols, is considered as the BER lower bound of the proposed mMTC scheme.

From Fig. 2(b), Fig. 3(b), and Fig. 4(b), it is obvious that the BER performance of the proposed mMTC scheme outperforms the traditional mMIMO UL (Benchmark 1) for the same throughput when P_e is small enough, thanks to the extra bits introduced by media modulation. Note that it is actually unfair for the proposed scheme to be compared with the benchmark 1 in BER performance since the latter does not consider the AUD error.

Fig. 2(a) and Fig. 2(b) compare the AUD performance and BER performance versus the signal-to-noise ratio (SNR), respectively. It is clear that the AUD performance of the proposed StrOMP algorithm is better than that of the TLSSCS algorithm, and it is hence closer to the AUD lower bound. We find that the BER performance of our "StrOMP+SIC-SSP" solution outperforms the TLSSCS detector, and the "StrOMP+GSP" solution, which demonstrates the efficiency of the proposed solution. Moreover, compared to the "StrOMP+GSP" solution, the BER performance of our "StrOMP+SIC-SSP" solution is getting better and better with the increase of SNR, which proves the the efficiency of the SIC operation.

Fig. 3(a) and Fig. 3(b) compare the AUD performance and BER performance versus the frame length J, respectively. Owing to the exploitation of the block sparsity, it can be seen that the AUD performance of the proposed StrOMP improves upon increasing J. Furthermore, as for the AUD performance, the advantage of the proposed StrOMP algorithm over the TLSSCS algorithm becomes more obvious upon increasing J. We also find that except for the Oracle LS (BER lower bound), the proposed "StrOMP+SIC-SSP" solution has the lowest BER floor, for sufficiently large J.

Fig. 4(a) and Fig. 4(b) compare the AUD performance and BER performance versus the number of receive antennas N_r, respectively. Observe from Fig. 4 that when N_r becomes large, the AUD performance or BER of the proposed "StrOMP+SIC-SSP" solution is better than that of the TLSSCS detector and the "StrOMP+GSP" solution. This indicates the superiority of the proposed solution for employment in mMIMO.

The computational complexity of different solutions in our simulations are compared in Table II, where the different algorithms are divided into two parts based on their functions (i.e., AUD or data detection). It is obvious that the number of complex-valued multiplications

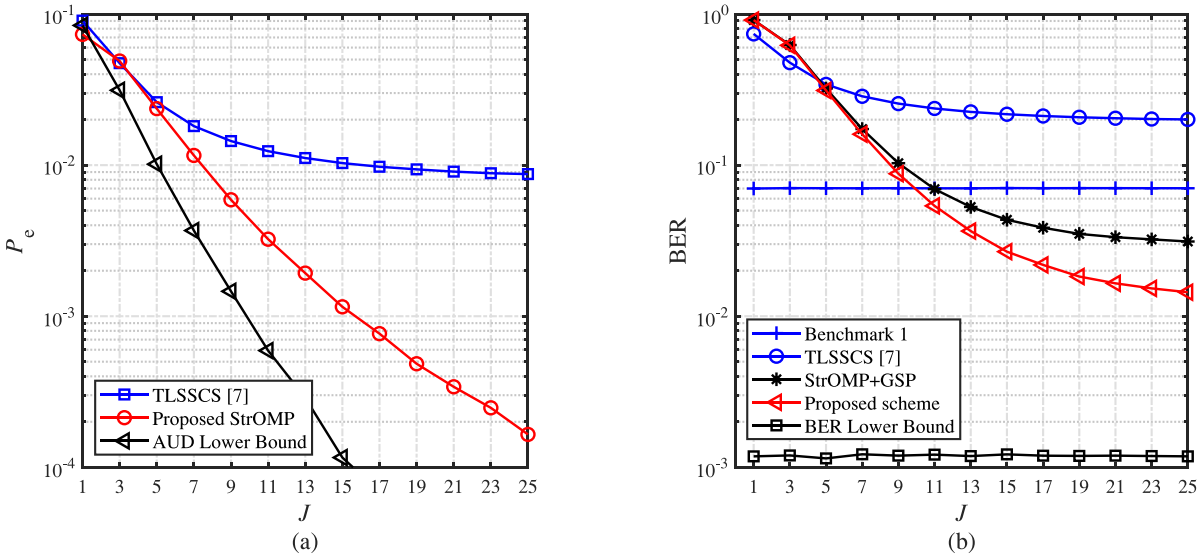

Fig. 3. Performance comparison of different solutions versus the frame length J (SNR $= 2$ dB, $N_r = 50$): (a) AUD performance; (b) BER performance.

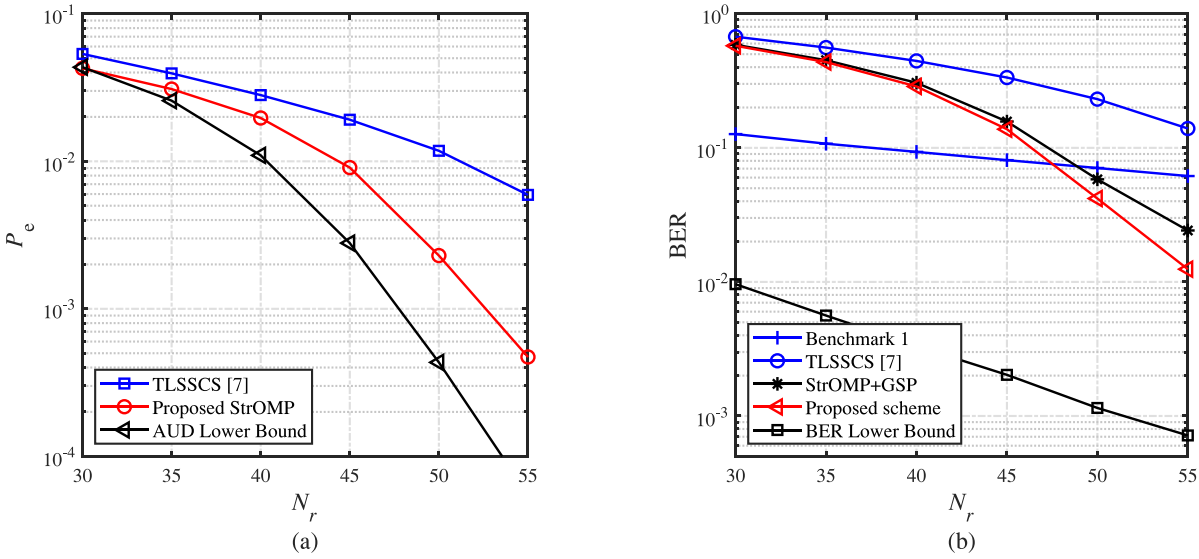

Fig. 4. Performance comparison of different solutions versus the the number of receive antennas N_r (SNR $= 2$ dB, $J = 12$): (a) AUD performance; (b) BER performance.

TABLE II
COMPUTATIONAL COMPLEXITY COMPARISON OF DIFFERENT ALGORITHMS

Algorithms		**Computational complexity**	**Complex-valued multiplications**[1](10^6)	
			$N_r = 50$	$N_r = 100$
AUD	Proposed StrOMP	$\mathcal{O}\{(K_a+1)JKN_tN_r + \sum_{s=1}^{K_a+1}[JN_r(s+2s^2+2(sN_t)^2)+J(s^3+(sN_t)^3)]\}$	9.6	17.6
	AUD part of TLSSCS [7]	$\mathcal{O}\{(K_a+1)[{N_r}^2(KN_t+J)+N_rJKN_t] + \sum_{s=1}^{K_a+1}[{N_r}^2+2N_r(sN_t)^2+(sN_t)^3]\}$	12.5	44.2
	AUD lower bound	$\mathcal{O}\{K_aJKN_tN_r + \sum_{s=1}^{K_a}[JN_r(s+2s^2+2(sN_t)^2)+J(s^3+(sN_t)^3)]\}$	7.1	13.2
Data detection	Proposed SIC-SSP	$\mathcal{O}\{J\sum_{s=1}^{K_a}[2sN_r(N_t+1)+14N_rs^2+11s^3]\}$	2.1	4.0
	Data detection part of TLSSCS [7]	$\mathcal{O}[JN_rK_aN_t+2N_r(K_aN_t)^2+(K_aN_t)^3]$	0.15	0.28
	GSP [8]	$\mathcal{O}\{J[2sN_r(N_t+1)+14N_r{K_a}^2+11{K_a}^3]\}$	0.65	1.2
	BER lower bound	$\mathcal{O}(JN_rK_a+2N_r{K_a}^2+{K_a}^3)$	0.01	0.02
	Benchmark 1	$\mathcal{O}(JN_rK_a+2N_r{K_a}^2+{K_a}^3)$	0.01	0.02

[1] The number of the complex-valued multiplications is calculated under the parameters $J = 12$, $N_t = 4$, $K = 100$, $K_a = 8$.

of the proposed StrOMP algorithm is a little lower than that in the AUD part of the TLSSCS algorithm (i.e., lines 1-14 of Algorithm 1 in [7]) when $N_r = 50$. If N_r is doubled, the number of complex-valued multiplications of the proposed StrOMP algorithm increases linearly with N_r, whereas the complexity of the AUD part of the TLSSCS algorithm is nearly proportional to the square of N_r. Hence, it is clear that our StrOMP algorithm is more suitable for mMIMO in conjunction with large antenna arrays. Furthermore, after obtaining the active MTDs, the data detection part of the TLSSCS algorithm becomes an LS operation (i.e., line 15 of Algorithm 1 in [7]) associated with a limited BER performance for the media-modulated signal. Hence, our proposed SIC-SSP algorithm attains a better data detection performance at the cost of a higher computational complexity.

V. Conclusion

A media modulation based mMTC UL scheme relying on mMIMO detection at the BS was proposed for achieving reliable massive access with an enhanced throughput. The sparse nature of the mMTC traffic motivated us to propose a CS-based solution. First, an StrOMP algorithm was proposed to detect the active MTDs exhibiting block-sparsity and structured sparsity of the UL signals, which improved the performance. Then, an SIC-SSP algorithm was proposed for detecting the data of the detected MTDs by exploiting the structured sparsity of media-modulated symbols for enhancing the performance. Furthermore, we analysed the computational complexity of the proposed algorithms. Finally, our simulation qualified the benefits of the proposed solution.

References

[1] C. Bockelmann *et al.*, "Massive machine-type communications in 5G: Physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, Sep. 2016.

[2] B. K. Jeong, B. Shim, and K. B. Lee, "MAP-based active user and data detection for massive machine-type communications," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8481–8494, Sep. 2018.

[3] 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.

[4] Y. Du *et al.*, "Block-sparsity-based multiuser detection for uplink grant-free NOMA," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 7894–7909, Dec. 2018.

[5] B. Wang, L. Dai, Y. Zhang, T. Mir, and J. Li, "Dynamic compressive sensing-based multi-user detection for uplink grant-free NOMA," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2320–2323, Nov. 2016.

[6] Y. Du, B. Dong, Z. Chen, X. Wang, Z. Liu, P. Gao, and S. Li, "Efficient multi-user detection for uplink grant-free NOMA: Prior-information aided adaptive compressive sensing perspective," *IEEE J. Select. Areas Commun.*, vol. 35, no. 12, pp. 2812–2828, Jul. 2017.

[7] X. Ma, J. Kim, D. Yuan, and H. Liu, "Two-level sparse structure based compressive sensing detector for uplink spatial modulation with massive connectivity," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1594–1597, Sep. 2019.

[8] Z. Gao, L. Dai, Z. Wang, S. Chen, and L. Hanzo, "Compressive-sensing based multiuser detector for the large-scale SM-MIMO uplink," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 1860–1865, Feb. 2017.

[9] L. Xiao *et al.*, "Efficient compressive sensing detectors for generalized spatial modulation systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1284–1298, Feb. 2017.

[10] L. Xiao, Y. Xiao, P. Yang, J. Liu, S. Li, and W. Xiang, "Space-time block coded differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8821–8834, Oct. 2017.

[11] L. Zhang, M. Zhao and L. Li, "Low-complexity multi-user detection for MBM in uplink large-scale MIMO systems," *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1568–1571, Aug. 2018.

[12] B. Shamasundar, S. Jacob, L. N. Theagarajan, and A. Chockalingam, "Media-based modulation for the uplink in massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8169–8183, Sep. 2018.

[13] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1527–1550, 2017.

[14] A. K. Khandani, "Media-based modulation: A new approach to wireless transmission," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 3050–3054.

[15] Y. Naresh and A. Chockalingam, "On media-based modulation using RF mirrors," *IEEE Trans. Veh. Tech.*, vol. 66, no. 6, pp. 4967–4983, Jun. 2017.

[16] E. Basar, "Media-based modulation for future wireless systems: A tutorial," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 160–166, Oct. 2019.

[17] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.