Massive Access in Media Modulation Based Massive Machine-Type Communications

Li Qiao, Jun Zhang, Zhen Gao, Derrick Wing Kwan Ng, Marco Di Renzo, and Mohamed-Slim Alouini

June 24, 2021

Outline

Outline

Introduction of mMTC

- The mMTC are characterized by uplink transmissions with short packets from massively deployed machine type devices (MTDs) whose data traffic is sporadic
- **Grant-based approaches**
- Allocating orthogonal radio resources to different active MTDs via some sophisticated scheduling algorithms is necessary
- Prohibitive signaling overhead and latency, even the access congestion
- **Grant-free approaches**
- MTDs can transmit in the uplink without waiting for permission
- Simplifies the uplink access procedure and hence reduces the access latency
- The sporadic traffic characteristics of MTDs in mMTC motivates the application of compressive sensing (CS) techniques to tackle the challenging **device activity and data detection** (DADD) problem.

Outline

Introduction of Spatial Modulation

- Spatial modulation utilizes a single or fewer radio frequency (RF) chains than the number of antenna elements
- Part of the information bits are encoded onto the activated antenna elements to enhance the data rate
- Low-complexity and energy-efficient multiple-antenna scheme
- The application and suitability of spatial modulation for IoT is also discussed and proved experimentally

Outline

- **Extensions**
- Many progresses have been made in order to enhance the spectral efficiency of spatial modulation without compromising the low-complexity and energy-efficiency that originates from using a single-RF chain
- Generalized spatial modulation, media modulation, and, more recently, metasurface-based modulation
- **Media Modulation**
- We focus on **media modulation**, which employs a single RF chain, a single radiating element, and several low-cost RF mirrors
- Information bits are encoded into the active/inactive (or ON/OFF) status of the RF mirrors, which determines the resulting radiation pattern of the entire structure
- In contrast to spatial modulation, the number of spatial bits encoded in media modulation is larger and depends on the number of distinguishable radiation patterns that can be realized

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		✓	✓	✓	✓	✓

Outline

Our Contributions

- **Uncoded media modulation based mMTC**
- A doubly structured AMP (DS-AMP) algorithm was proposed for uncoded media modulation based mMTC
- We derive the theoretical state evolution (SE), which closely matches the simulated results of DS-AMP algorithm
- **Coded media modulation based mMTC**
- Bit-interleaved coded media modulation (BICMM) was designed for coded media modulation based mMTC
- Successive interference cancellation (SIC)-based iterative DS-AMP (IDS-AMP) scheme was developed for coded media modulation based mMTC, achieving improved the data decoding performance

- All K MTDs employ media modulation for enhancing throughput and the BS employs mMIMO with $N_r \gg 1$ antenna elements for reliable massive access
- Only K_a out of K ($K \gg K_a$) MTDs are simultaneously active
- Each MTD is equipped with an RF chain, a transmit antenna, and N_{RF} low-cost RF mirrors
- Each device has $N_t = 2^{N_{\mathrm{RF}}}$ different kinds of MAPs (i.e., N_t different channel realizations), where $\log_2 N_t = N_{\mathrm{RF}}$ extra information bits can be achieved

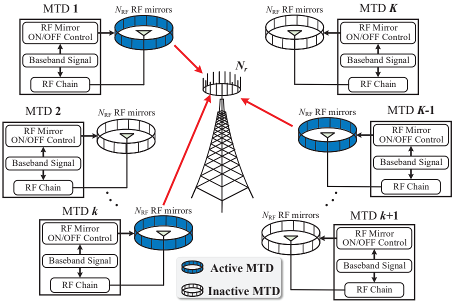

Fig. 1. For media modulation based mMTC scheme, MTDs adopt media modulation to access the mMIMO BS.

- The received signal at the BS in the j-th, $\forall j \in [J]$, time slot, can be written as

$$\mathbf{y}_j = \sum_{k=1}^{K} a_k s_{k,j} \mathbf{H}_k \mathbf{d}_{k,j} + \mathbf{w}_j = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{x}_{k,j} + \mathbf{w}_j = \mathbf{H}\tilde{\mathbf{x}}_j + \mathbf{w}_j, \tag{1}$$

- The binary activity indicator $a_k \in \{0, 1\}$ equals one (zero) as long as the k-th MTD is active (inactive)
- $s_{k,j}$ associated with the k-th MTD in the j-th time slot is selected from the M-QAM set $\mathbb{S}$
- $\mathbf{d}_{k,j} \in \mathbb{C}^{N_t \times 1}$ is the media modulated symbol
- $\mathbf{x}_{k,j} = a_k s_{k,j} \mathbf{d}_{k,j} \in \mathbb{C}^{N_t \times 1}$ is the effective uplink transmitted symbols
- $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is the multiple input multiple output (MIMO) channel matrix associated with the k-th MTD
- $\mathbf{w}^j \in \mathbb{C}^{N_r \times 1}$ is the Gaussian noise with its elements following the independent and identically distributed (i.i.d.) complex Gaussian distribution $\mathcal{CN}([\mathbf{w}^j]_n; 0, \sigma_w^2)$
- $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_K] \in \mathbb{C}^{N_r \times (KN_t)}$ and $\tilde{\mathbf{x}}_j = [(\mathbf{x}_{1,j})^T, (\mathbf{x}_{2,j})^T, ..., (\mathbf{x}_{K,j})^T]^T \in \mathbb{C}^{(KN_t) \times 1}$ are the aggregated channel matrix and uplink access signal, respectively

- Due to the media modulation property, only one entry of the media modulated symbol $\mathsf{d}_{k,j}$, $\forall j \in [J]$ and $\forall k \in [K]$, equals one and the others are zeros, i.e.,

$$\mathrm{supp}\left\{\mathsf{d}_{k,j}\right\} \in [N_t], \quad \|\mathsf{d}_{k,j}\|_0 = 1, \quad \|\mathsf{d}_{k,j}\|_2 = 1, \tag{2}$$

where $\mathrm{supp}\{\mathsf{d}_{k,j}\}$ denotes the support set of $\mathsf{d}_{k,j}$. We refer to this property as the *structured sparsity in the modulation domain*.

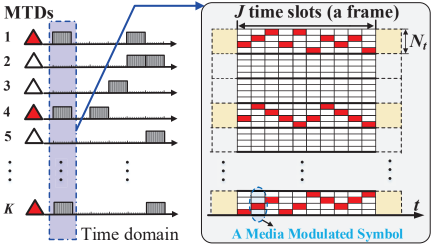

Fig. 2. We consider media modulation based mMTC with the slotted access frame structure, where the invariant active/inactive status of MTDs within a frame forms the *structured sparsity in the time domain*, and the media modulated symbols possess the *structured sparsity in the modulation domain*.

Outline

- We introduce an activity indicator vector $\mathrm{a} = [a_1, a_2, ..., a_K]^T \in \mathbb{C}^{K \times 1}$, which is sparse as the number of active MTDs $K_a = \|\mathrm{a}\|_0 \ll K$
- We collectively refer to the *s*tructured sparsity in the time domain due to the slotted access frame structure and the *s*tructured sparsity in the modulation domain shown in (2) as the *d*oubly structured sparsity
- To exploit the *s*tructured sparsity in the time domain, we rewrite the received signals of J successive time slots as

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

where we have $\mathrm{Y} = [\mathrm{y}_1, \mathrm{y}_2, ..., \mathrm{y}_J] \in \mathbb{C}^{N_r \times J}$, $\mathrm{H} \in \mathbb{C}^{N_r \times (KN_t)}$, $\mathrm{X} = [\tilde{\mathrm{x}}_1, \tilde{\mathrm{x}}_2, ..., \tilde{\mathrm{x}}_J] \in \mathbb{C}^{(KN_t) \times J}$, and $\mathrm{W} = [\mathrm{w}_1, \mathrm{w}_2, ..., \mathrm{w}_J] \in \mathbb{C}^{N_r \times J}$.
- The massive access problem can be formulated as the following optimization problem

$$\begin{aligned} \min_{\mathrm{X}} \|\mathrm{Y} - \mathrm{H}\mathrm{X}\|_F^2 &= \min_{\{\tilde{\mathrm{x}}_j\}_{j=1}^J} \sum_{j=1}^{J} \|\mathrm{y}_j - \mathrm{H}\tilde{\mathrm{x}}_j\|_2^2 \\ &= \min_{\{a_k, \mathrm{d}_{k,j}, s_{k,j}\}_{j=1,k=1}^{J,K}} \sum_{j=1}^{J} \left\| \mathrm{y}_j - \sum_{k=1}^{K} a_k s_{k,j} \mathrm{H}_k \mathrm{d}_{k,j} \right\|_2^2 \end{aligned} \tag{4}$$
$$\text{s.t. } (2),\ \ \|\mathrm{a}\|_0 \ll K,\ \text{ and } s_{k,j} \in \mathbb{S}, k \in [K], j \in [J].$$

Problem Formulation Based on Factor Graph

- Minimizing the mean square error between Y and HX is equivalent to estimating the *a* posteriori mean of the uplink access signal X
- The *a* posteriori mean of $[\mathrm{x}_{k,j}]_i$, $\forall k \in [K]$, $\forall j \in [J]$, $\forall i \in [N_t]$, can be expressed as

$$[\widehat{\mathrm{x}}_{k,j}]_i = \sum_{[\mathrm{x}_{k,j}]_i \in \overline{\mathbb{S}}} [\mathrm{x}_{k,j}]_i \, p\left([\mathrm{x}_{k,j}]_i \,|\mathrm{y}_j\right), \tag{5}$$

where $\overline{\mathbb{S}} = \{\mathbb{S}, 0\}$, $p\left([\mathrm{x}_{k,j}]_i \,|\mathrm{y}_j\right)$ is the marginal distribution of $p\left(\tilde{\mathrm{x}}_j|\mathrm{y}_j\right)$ and it can be expressed as

$$p\left([\mathrm{x}_{k,j}]_i \,|\mathrm{y}_j\right) = \sum_{\sim\left\{[\mathrm{x}_{k,j}]_i\right\}} p\left(\tilde{\mathrm{x}}_j|\mathrm{y}_j\right). \tag{6}$$

- Based on the Bayes' theorem, the joint posterior distribution $p\left(\tilde{\mathrm{x}}_j|\mathrm{y}_j\right)$ can be expressed as

$$p\left(\tilde{\mathrm{x}}_j|\mathrm{y}_j;\sigma_w^2,\mathrm{a}\right) = \frac{p\left(\mathrm{y}_j|\tilde{\mathrm{x}}_j;\sigma_w^2\right)p\left(\tilde{\mathrm{x}}_j;\mathrm{a}\right)}{p\left(\mathrm{y}_j\right)} = \frac{1}{p\left(\mathrm{y}_j\right)}\prod_{n=1}^{N_r} p\left([\mathrm{y}_j]_n\,|\tilde{\mathrm{x}}_j;\sigma_w^2\right)\prod_{k=1}^{K} p\left(\mathrm{x}_{k,j};a_k\right), \tag{7}$$

where the likelihood function can be expressed as

$$p\left([\mathrm{y}_j]_n\,|\tilde{\mathrm{x}}_j;\sigma_w^2\right) = \frac{1}{\pi\sigma_w^2}\exp\left(-\frac{1}{\sigma_w^2}\left|[\mathrm{y}_j]_n - \sum_{k=1}^{K}[\mathrm{H}_k\mathrm{x}_{k,j}]_n\right|^2\right). \tag{8}$$

Problem Formulation Based on Factor Graph

- According to the structured sparsity in the modulation domain and the discrete distribution of QAM alphabet, the *a* prior distribution $p\left(\mathrm{x}_{k,j}; a_k\right)$ in (7) is formulated as

$$
\begin{aligned}
p\left(\mathrm{x}_{k,j}; a_k\right) = {} & (1-a_k)\prod_{i=1}^{N_t}\delta\left(\left[\mathrm{x}_{k,j}\right]_i\right) + \\
& a_k\left\{\frac{1}{N_t}\sum_{i=1}^{N_t}\left[\frac{1}{M}\sum_{s\in\mathbb{S}}\delta\left(\left[\mathrm{x}_{k,j}\right]_i - s\right)\prod_{g\in[N_t],g\neq i}\delta\left(\left[\mathrm{x}_{k,j}\right]_g\right)\right]\right\},
\end{aligned}
\tag{9}
$$

where $M = |\mathbb{S}|_c$ and $\delta(\cdot)$ is the Dirac delta function.

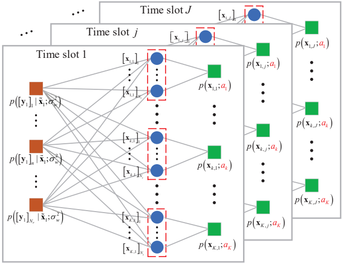

Fig. 3. Factor graph of the joint posterior distribution $p\left(\tilde{\mathrm{x}}_j|\mathrm{y}_j;\sigma_w^2,\mathrm{a}\right)$. The circles represent variable nodes and the squares represent factor nodes.

- However, calculating the marginal distribution $p\left([\mathrm{x}_{k,j}]_i\,|\mathrm{y}_j\right)$, $\forall k \in [K]$, $\forall j \in [J]$, and $\forall i \in [N_t]$, from the joint posterior distribution $p\left(\tilde{\mathrm{x}}_j|\mathrm{y}_j;\sigma_w^2,\mathrm{a}\right)$ is extremely complicated in massive access, due to the exceedingly large value of K.
- Fortunately, AMP can provide an effective approximation of the marginal distributions with a lower complexity, yet achieving the near MMSE performance.
- The DS-AMP algorithm is proposed for handling such a massive access problem
- We apply the DS-AMP algorithm to calculate the *a* posteriori mean of media modulated signals $\mathrm{x}_{k,j}$ $(k \in [K],\, j \in [J])$, and resort to the EM algorithm for estimating the activity indicators a_k as well as the variance σ_w^2 of the complex Gaussian noise.

- AMP decouples the matrix estimation problem of (3) into KJN_t uncoupled scalar problems in the asymptotic regime, i.e.,

$$\mathrm{Y} = \mathrm{HX} + \mathrm{W} \ \rightarrow \ r_{l,j} = [\mathrm{x}_{k,j}]_i + \widehat{w}_{l,j}, \ \forall i, j, k, \tag{10}$$

where $l = (k-1)N_t + i$, $r_{l,j}$ is the mean of $[\mathrm{x}_{k,j}]_i$ estimated by AMP algorithm, $\widehat{w}_{l,j} \sim \mathcal{CN}\left(\widehat{w}_{l,j}; 0, \phi_{l,j}\right)$ is the equivalent noise, and $\phi_{l,j}$ is its variance.

- The joint posterior distribution (7) can be approximated as

$$p\left(\tilde{\mathrm{x}}_j | \mathrm{y}_j; \sigma_w^2, \mathrm{a}\right) \approx q\left(\tilde{\mathrm{x}}_j | \mathrm{y}_j; \sigma_w^2, \mathrm{a}\right) = \prod_{k=1}^{K} \prod_{i=1}^{N_t} q\left([\mathrm{x}_{k,j}]_i \,|\, r_{l,j}, \phi_{l,j}; \sigma_w^2, a_k\right). \tag{11}$$

- Based on the Bayes' theorem, the approximated marginal *a* posteriori distribution, denoted as $q\left([\mathrm{x}_{k,j}]_i \,|\, r_{l,j}, \phi_{l,j}; \sigma_w^2, a_k\right)$, $\forall i, j, k$, is given as follows

$$q\left([\mathrm{x}_{k,j}]_i \,|\, r_{l,j}, \phi_{l,j}; \sigma_w^2, a_k\right) = \frac{q\left(r_{l,j} | \, [\mathrm{x}_{k,j}]_i \,; \sigma_w^2\right) \ p\left([\mathrm{x}_{k,j}]_i \,; a_k\right)}{q\left(r_{l,j}; \sigma_w^2, a_k\right)}, \tag{12}$$

where

$$q\left(r_{l,j} \middle| \left[x_{k,j}\right]_i\right) = \frac{1}{\pi\phi_{l,j}} \exp\left(-\frac{1}{\phi_{l,j}} \left|r_{l,j} - \left[x_{k,j}\right]_i\right|^2\right), \tag{13}$$

$$\begin{aligned} p\left(\left[x_{k,j}\right]_i; a_k\right) &= \sum_{\sim\left\{\left[x_{k,j}\right]_i\right\}} p\left(x_{k,j}; a_k\right) \\ &= \left(1 - \frac{a_k}{N_t}\right)\delta\left(\left[x_{k,j}\right]_i\right) + \frac{a_k}{N_t M}\sum_{s\in\mathbb{S}}\delta\left(\left[x_{k,j}\right]_i - s\right), \end{aligned} \tag{14}$$

$$q\left(r_{l,j}; \sigma_w^2, a_k\right) = \sum_{\left[x_{k,j}\right]_i \in \overline{\mathbb{S}}} q\left(r_{l,j} \middle| \left[x_{k,j}\right]_i; \sigma_w^2\right) \; p\left(\left[x_{k,j}\right]_i; a_k\right). \tag{15}$$

- Then, the *a* posteriori mean and variance of $\left[x_{k,j}\right]_i$, $\forall i, j, k$, are respectively given as

$$\left[\widehat{x}_{k,j}\right]_i = f_m\left(r_{l,j}, \phi_{l,j}\right) = \sum_{\left[x_{k,j}\right]_i \in \overline{\mathbb{S}}} \left[x_{k,j}\right]_i \; q\left(\left[x_{k,j}\right]_i \middle| r_{l,j}, \phi_{l,j}; \sigma_w^2, a_k\right), \tag{16}$$

$$\left[\widehat{v}_{k,j}\right]_i = f_v\left(r_{l,j}, \phi_{l,j}\right) = -\left|\left[\widehat{x}_{k,j}\right]_i\right|^2 + \sum_{\left[x_{k,j}\right]_i \in \overline{\mathbb{S}}} \left|\left[x_{k,j}\right]_i\right|^2 q\left(\left[x_{k,j}\right]_i \middle| r_{l,j}, \phi_{l,j}; \sigma_w^2, a_k\right), \tag{17}$$

where $l = (k-1)N_t + i$.

Update Rules of DS-AMP Algorithm

- Note that $r_{l,j}$, $\phi_{l,j}$, $\widehat{\mathrm{x}}_{k,j}$, and $\widehat{\mathrm{v}}_{k,j}$ are updated iteratively in the AMP algorithm.
- In the factor graph in Fig. 5, $r_{l,j}$ and $\phi_{l,j}$, $\forall l, j$, are updated iteratively at the variable nodes. The update rules in the t-th iteration are expressed as

$$\phi_{l,j}^t = \left(\sum_{n=1}^{N_r} \frac{\left| \mathrm{H}_{[n,l]} \right|^2}{\sigma_w^2 + V_{n,j}^t} \right)^{-1}, \tag{18}$$

$$r_{l,j}^t = \left[\widehat{\mathrm{x}}_{k,j}^t \right]_i + \phi_{l,j}^t \sum_{n=1}^{N_r} \frac{\mathrm{H}_{[n,l]}^* \left([\mathrm{y}_j]_n - Z_{n,j}^t \right)}{\sigma_w^2 + V_{n,j}^t}, \tag{19}$$

- $V_{n,j}^t$ and $Z_{n,j}^t$, $\forall n, j$, are updated at the factor nodes of the factor graph as

$$V_{n,j}^t = \sum_{k=1}^{K} \left| \mathrm{H}_{k\,[n,:]} \right|^2 \widehat{\mathrm{v}}_{k,j}^t, \tag{20}$$

$$Z_{n,j}^t = \sum_{k=1}^{K} \mathrm{H}_{k\,[n,:]} \widehat{\mathrm{x}}_{k,j}^t - V_{n,j}^t \frac{[\mathrm{y}_j]_n - Z_{n,j}^{t-1}}{\sigma_w^2 + V_{n,j}^{t-1}}. \tag{21}$$

Parameters Estimation

- The EM algorithm is an iterative approach to find the maximum likelihood solutions for probabilistic models with the unknown parameters
- Define $\boldsymbol{\theta} = \left\{\sigma_w^2, a_k, k \in [K]\right\}$, EM algorithm updates the parameter set $\boldsymbol{\theta}$ as follows

$$Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^t\right) = \mathbb{E}\left\{\ln\, p\left(\mathrm{X}, \mathrm{Y}; \boldsymbol{\theta}\right) | \mathrm{Y}; \boldsymbol{\theta}^t\right\}, \tag{22}$$

$$\boldsymbol{\theta}^{t+1} = \arg\max_{\boldsymbol{\theta}} Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^t\right), \tag{23}$$

  where $\boldsymbol{\theta}^t$ is the parameter set estimated in the t-th iteration, $\mathbb{E}\left\{\cdot | \mathrm{Y}; \boldsymbol{\theta}^t\right\}$ denotes the expectation conditioned on the received signal Y under $\boldsymbol{\theta}^t$.
- The update rules of the noise variance σ_w^2 and the activity indicator a_k, $\forall k$, as follows

$$\left(\sigma_w^2\right)^{t+1} = \frac{1}{JN_r} \sum_{j=1}^{J} \sum_{n=1}^{N_r} \left[\frac{\left([\mathrm{y}_j]_n - Z_{n,j}^t\right)^2}{\left(1 + \frac{V_{n,j}^t}{\left(\sigma_w^2\right)^t}\right)^2} + \frac{\left(\sigma_w^2\right)^t V_{n,j}^t}{V_{n,j}^t + \left(\sigma_w^2\right)^t} \right], \tag{24}$$

$$a_k^{t+1} = f_a\left(r_{l,j}^t, \phi_{l,j}^t; a_k^t\right) = \frac{1}{J} \sum_{j=1}^{J} \sum_{\mathrm{x}_{k,j} \in \Gamma_0} \prod_{i=1}^{N_t} q\left([\mathrm{x}_{k,j}]_i | r_{l,j}^t, \phi_{l,j}^t; a_k^t\right), \tag{25}$$

  where $l = (k-1)N_t + i$ and Γ_0 is the set of all possible $\mathrm{x}_{k,j}$ when the k-th MTD is active

Arithmetic Flow of the Proposed DS-AMP Algorithm

Algorithm 1: Proposed DS-AMP Algorithm

Input: The received signals $\mathbf{Y}=[\mathbf{y}_1,...,\mathbf{y}_J]\in\mathbb{C}^{N_r\times J}$, the channel matrix $\mathbf{H}=[\mathbf{H}_1,...,\mathbf{H}_K]\in\mathbb{C}^{N_r\times(KN_t)}$, and the maximum iteration number T_0.

Output: The set of active MTDs Ω and the reconstructed media modulation signal $\mathbf{X}\in\mathbb{C}^{KN_t\times J}$.

$\forall i,j,k,n$: We initialize the iterative index t=1, the activity indicator $a_k^1=0.5$, $Z_{n,j}^0=[\mathbf{y}_j]_n$, $V_{n,j}^0=1$, the noise variance $\left(\sigma_w^2\right)^1=100$, the reconstructed signal $\mathbf{X}=\mathbf{0}_{KN_t\times J}$, $\left[\widehat{\mathbf{x}}_{k,j}^1\right]_i=a_k^1\sum\limits_{s\in\mathbb{S}}s/MN_t$, and $\left[\widehat{\mathbf{v}}_{k,j}^1\right]_i=a_k^1\sum\limits_{s\in\mathbb{S}}|s|^2/MN_t-\left|\left[\widehat{\mathbf{x}}_{k,j}^1\right]_i\right|^2$;

for $t=1$ to T_0 **do**

  %AMP operation:

  $\forall i,j,k,n$: Compute $V_{n,j}^t$, $Z_{n,j}^t$, $\phi_{l,j}^t$, and $r_{l,j}^t$ by using (20), (21), (18), and (19), respectively, where $l=(k-1)N_t+i$; {Decoupling step}

  $\forall i,j,k,n$: Compute $\left[\widehat{\mathbf{x}}_{k,j}^{t+1}\right]_i$ and $\left[\widehat{\mathbf{v}}_{k,j}^{t+1}\right]_i$ by using (16) and (17), respectively; {Denoising step}

  %EM operation:

  $\forall k$: Compute $(\sigma_w^2)^{t+1}$ and a_k^{t+1} by using (24) and (25);

end for

%Min-max normalization:

Let $\tilde{\mathbf{a}}=\frac{\widehat{\mathbf{a}}-\min(\widehat{\mathbf{a}})}{\max(\widehat{\mathbf{a}})-\min(\widehat{\mathbf{a}})}$, where $\widehat{\mathbf{a}}=[\widehat{a}_1,...,\widehat{a}_K]^T=\left[a_1^{T_0},...,a_K^{T_0}\right]^T$, $\min(\cdot)$ and $\max(\cdot)$ are the minimum value and maximum value of the arguments, respectively;

%Extract the active MTDs and their MAPs:

$\forall k$: The set of active MTDs $\Omega=\{k|\,[\tilde{\mathbf{a}}]_k>0.5\}$;

$\forall k,j$: $\eta^*=\arg\max_{\widehat{\eta}\in[N_t]}\left[\widehat{\mathbf{x}}_{k,j}^{T_0}\right]_{\widehat{\eta}}$;

$\forall k\in\Omega,\forall j$: The reconstructed signal is $\mathbf{X}_{[(k-1)N_t+\eta^*,j]}=\left[\widehat{\mathbf{x}}_{k,j}^{T_0}\right]_{\eta^*}$.

Outline

- SE is a tool for analyzing the performance of AMP algorithms in the large system limit, i.e., $KN_t \to \infty$, by tracking the mean-square errors (MSE) of each iteration
- To start with, the MSE and average variance are respectively defined as

$$e^t = \frac{1}{KJN_t} \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{i=1}^{N_t} \left| [\widehat{\mathrm{x}}_{k,j}^t]_i - [\mathrm{x}_{k,j}]_i \right|^2 , \tag{26}$$

$$v^t = \frac{1}{KJN_t} \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{i=1}^{N_t} [\widehat{\mathrm{v}}_{k,j}^t]_i , \tag{27}$$

- In large system limit and the elements of the measurement matrix obey the i.i.d. distribution with zero mean and variance γ, the value of r_0^t in SE can be expressed as

$$r_0^t = x_0 + \sqrt{\frac{\sigma_w^2 + \gamma K N_t e^t}{N_r \gamma}} z, \tag{28}$$

where x_0 is a realization of X_0, $z \sim \mathcal{CN}(z; 0, 1)$, and the quantity of ϕ_0^t in SE can be shown as

$$\phi_0^t \approx \frac{\sigma_w^2 + \gamma K N_t v^t}{N_r \gamma}. \tag{29}$$

- The MSE and the average variance in the $(t+1)$-th iteration can be expressed as

$$e^{t+1} = \int dx_0 p_0(x_0) \int \mathcal{D}z \left| f_m(r_0^t, \phi_0^t) - x_0 \right|^2, \tag{30}$$

$$v^{t+1} = \int dx_0 p_0(x_0) \int \mathcal{D}z f_v(r_0^t, \phi_0^t), \tag{31}$$

where $p_0(x_0)$ is the *a* prior distribution as indicated in (14), $\mathcal{D}z = e^{-|z|^2}/\pi dz$, $f_m(r_0^t, \phi_0^t)$ and $f_v(r_0^t, \phi_0^t)$ are defined in (16) and (17), respectively

- Monte Carlo simulations are adopted to generate a large number of realizations of the transmit signals, where the sporadic traffic and the doubly structured sparsity are fully embodied

State Evolution of DS-AMP Algorithm

Algorithm 2: State Evolution of DS-AMP Algorithm

Input: The noise variance σ_w^2, the sparsity level $\lambda = \frac{K_a}{K}$, the number of MAPs N_t, the frame length J, the order of QAM, the variance of the elements in the measurement matrix γ, the number of Monte Carlo simulations N_{MC}, the maximum SE iterations T_{SE}, and ε.

Output: The theoretically predicted MSE $\widehat{e}$.

$\forall m \in [N_{\text{MC}}]$: Generate N_{MC} realizations of transmit signals $\mathbf{X}^m \in \mathbb{C}^{KN_t \times J}$, based on the *a prior* distribution (9).

$\forall m, k$: Define $\mathbf{e}^1 = \mathbf{0}_{N_{\text{MC}} \times 1}$ and $\mathbf{v}^1 = \mathbf{0}_{N_{\text{MC}} \times 1}$ to record the predicted MSE and average variance of the m-th Monte Carlo realization. We initialize the iteration number $t = 1$, the predicted MSE $e^1 = 1$, the average variance $v^1 = 1$, and the activity indicators for the m-th signal realization $a_{k,m}^1 = 0.5$;

for $t = 1$ to T_{SE} **do**

    for $m = 1$ to N_{MC} **do**

        $\forall i, j, k$: $r_{l,j}^{m,t} = \left[\mathbf{x}_{k,j}^m\right]_i + \sqrt{\frac{\sigma_w^2 + \gamma K N_t e^t}{N_r \gamma}} z$, $\phi_{l,j}^{m,t} = \frac{\sigma_w^2 + \gamma K N_t v^t}{N_r \gamma}$;

        $\forall i, j, k$: $\left[\widehat{\mathbf{x}}_{k,j}^m\right]_i = f_m(r_{l,j}^{m,t}, \phi_{l,j}^{m,t})$, $\left[\widehat{\mathbf{v}}_{k,j}^m\right]_i = f_v(r_{l,j}^{m,t}, \phi_{l,j}^{m,t})$;

        $\forall k$: $a_{k,m}^{t+1} = f_a(r_{l,j}^{m,t}, \phi_{l,j}^{m,t}; a_{k,m}^t)$;

        Calculating $\left[\mathbf{e}^{t+1}\right]_m$ and $\left[\mathbf{v}^{t+1}\right]_m$ referring to (27) and (28), respectively;

    end for

    $e^{t+1} = \frac{1}{N_{\text{MC}}} \sum_{m=1}^{N_{\text{MC}}} \left[\mathbf{e}^{t+1}\right]_m$, $v^{t+1} = \frac{1}{N_{\text{MC}}} \sum_{m=1}^{N_{\text{MC}}} \left[\mathbf{v}^{t+1}\right]_m$;

    $\widehat{e} = e^{t+1}$;

    if $\left|e^{t+1} - e^t\right| < \varepsilon$ **then**

        break; {End the SE iterations}

    end if

end for

Computational Complexity of DS-AMP Algorithm

- The computational complexity of the proposed DS-AMP algorithm mainly arises from the complex-valued matrix multiplications of the following operations in each iteration
- **AMP decoupling step**: The complexity of performing the AMP decoupling step, i.e., (18)−(21), is $\mathcal{O}(\frac{5}{2}JKN_tN_r)$
- **AMP denoising step**: The complexity of performing the AMP denoising step, i.e., (15)−(17), is $\mathcal{O}\left[JKN_t(|\mathbb{S}|_c+\frac{1}{4})\right]$
- Simulation results demonstrate that the predefined maximum iteration number T_0 can be small to guarantee the convergence of the proposed DS-AMP algorithm. Hence the overall complexity is on the order of $\mathcal{O}\left[T_0JKN_t(\frac{5}{2}N_r+|\mathbb{S}|_c+\frac{1}{4})\right]$, which scales linearly with the number of MTDs, the number of MAPs in media modulation, the order of QAM modulation, and the number of receive antennas at the BS
- This linear complexity is appealing for efficiently processing the massive access of future IoT

Outline

- **Data Packet Structure**
- The proposed data packet structure is dedicated for the implementation of an SIC-based IDS-AMP detector at the receiver
- Specifically, the data packet is composed of two parts, including the former L_s-bit signature sequence part and the following L_d-bit payload data
- We consider that all the MTDs share the same binary signature sequence $\mathrm{b}_s \in \mathbb{N}^{L_s \times 1}$, which is a pre-defined pseudo-random 0/1 sequence known at the transceiver
- At the receiver, we consider the Hamming distance $D(\mathrm{b}_s, \widehat{\mathrm{b}}_s)$ between b_s and $\widehat{\mathrm{b}}_s$ as a metric to evaluate the decoding quality of the associated MTD's payload data part, where $\widehat{\mathrm{b}}_s \in \mathbb{N}^{L_s \times 1}$ is the estimated binary signature sequence of any detected MTD
- As a result, the error propagation in the SIC processing can be mitigated with the aid of the proposed signature sequence

- **BICMM**
- The BICMM scheme includes an encoder, a bit-wise interleaver, and a media modulation module
- Specifically, after the channel coding, the length of one data packet is expanded from L bits to L' bits, and then this L'-bit data packet is delivered to the following bit-wise interleaver module
- We consider a **block interleaver** with η columns by $J = L'/\eta$ rows, and the L'-bit data packet is read into the interleaver by rows and read out by columns
- Every η bits of the interleaved L'-bit data packet are sequentially modulated into J media modulation symbols and are transmitted in J successive time slots (i.e., a frame)
- The bit-wise interleaver module can provide the diversity to overcome the dramatic spatial-selective fading channels of media modulation for improving data decoding performance
- This interleaver module is different from the interleaver used in the channel coding (e.g., Turbo coding), where the latter is designed for AWGN channels rather than spatial-selective fading channels

Dedicated Data Packet Structure and BICMM at MTDs

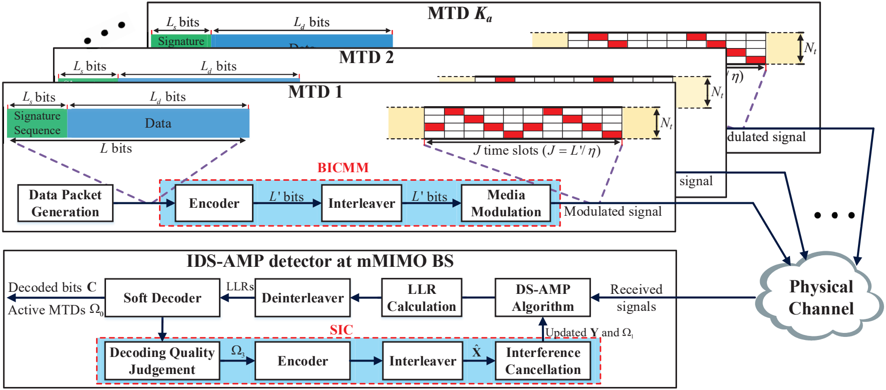

Fig. 4. Communication process of the proposed massive access solution for coded media modulation based mMTC.

The proposed IDS-AMP detector has 8 modules, including

- **DS-AMP algorithm module**
- **LLR calculation module**
- **A deinterleaver**
- **A soft decoder**
- **Decoding quality judgement module**
- **An encoder** {For signals reconstruction}
- **An interleaver** {For signals reconstruction}
- **Interference cancellation module**

Proposed IDS-AMP Detector at the BS

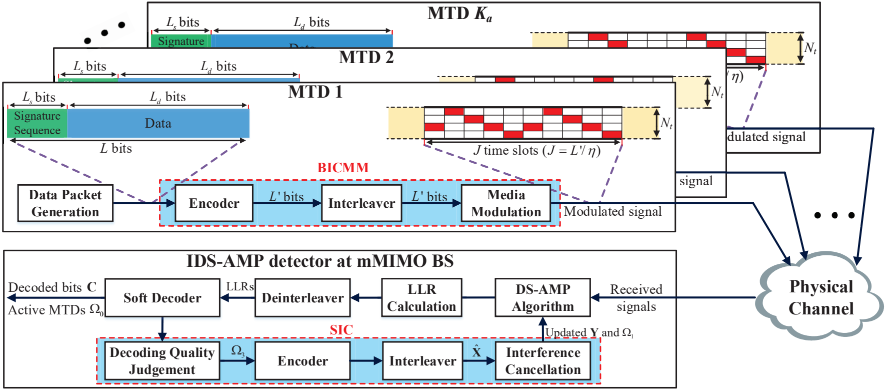

Fig. 4. Communication process of the proposed massive access solution for coded media modulation based mMTC.

- **DS-AMP algorithm module**
- We obtain the approximated *a* posteriori distribution $q\left(\mathrm{x}_{k,j}|\mathrm{y}_j;\sigma_w^2,a_k\right)$, $\forall k,j$, and the activity indicator vector $\widehat{\mathrm{a}}=[\widehat{a}_1,...,\widehat{a}_K]^T$, which can be acquired by calling the DS-AMP algorithm
- If the iteration index $i=0$, we acquire the indices of active MTDs detected, denoted as Ω_0
- Meanwhile, the index set of MTDs remaining to be iteratively decoded, denoted as Ω_1, is assigned to be equivalent to Ω_0 in the first iteration (i.e., the iteration index $i=0$) and will be updated in the following SIC.
- We update the iteration index $i=i+1$, and then select the $\overline{N}$ MTDs most likely to be active based on the quantities of activity indicators for the following SIC, where these $\overline{N}$ MTDs' index set is denoted as $\Omega_2{=}\Theta([\widehat{\mathrm{a}}]_{\Omega_1},\overline{N})$
- If $|\Omega_1|_c$ is smaller than the predefined constant $\overline{N}$, let $\Omega_2=\Omega_1$

- **LLR calculation module**
- For any media modulation symbol $\mathrm{x}_{k,j}$, $\forall k, j$, the LLR of the associated media modulated bit $B_{k,j,b}^{\mathrm{MED}}$, $\forall b \in [\log_2 N_t]$, and the LLR of the associated M-QAM bit $B_{k,j,d}^{\mathrm{QAM}}$, $\forall d \in [\log_2 M]$, can be respectively expressed as

$$\mathrm{LLR}\left(B_{k,j,b}^{\mathrm{MED}}\right) = \log \frac{\sum_{\mathrm{x}_{k,j} \in \Phi_0^b} q\left(\mathrm{x}_{k,j} | \mathrm{y}_j; \sigma_w^2, a_k\right)}{\sum_{\mathrm{x}_{k,j} \in \Phi_1^b} q\left(\mathrm{x}_{k,j} | \mathrm{y}_j; \sigma_w^2, a_k\right)}, \tag{32}$$

$$\mathrm{LLR}\left(B_{k,j,d}^{\mathrm{QAM}}\right) = \log \frac{\sum_{\mathrm{x}_{k,j} \in \Psi_0^d} q\left(\mathrm{x}_{k,j} | \mathrm{y}_j; \sigma_w^2, a_k\right)}{\sum_{\mathrm{x}_{k,j} \in \Psi_1^d} q\left(\mathrm{x}_{k,j} | \mathrm{y}_j; \sigma_w^2, a_k\right)}, \tag{33}$$

where Φ_0^b (Φ_1^b) is the set of $\mathrm{x}_{k,j}$ for which the media modulated bit $B_{k,j,b}^{\mathrm{MED}}$, $\forall b$, equals zero (one), and Ψ_0^d (Ψ_1^d) is the set of $\mathrm{x}_{k,j}$ for which the M-QAM bit $B_{k,j,d}^{\mathrm{QAM}}$, $\forall d$, equals zero (one)
- For example, supposing that $N_t = 2$, $M = 2$, $\mathbb{S} = \{+1, -1\}$, $b \in [1]$, and $d \in [1]$, then we can get

$$\Phi_0^1 = \left\{ \begin{bmatrix} +1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\}, \Phi_1^1 = \left\{ \begin{bmatrix} 0 \\ +1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}, \tag{34}$$

$$\Psi_0^1 = \left\{ \begin{bmatrix} +1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ +1 \end{bmatrix} \right\}, \Psi_1^1 = \left\{ \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}. \tag{35}$$

- **Deinterleaver and soft decoder modules**
- Firstly, the LLR information of MTDs with indices in Ω_2 is respectively deinterleaved
- Secondly, for each MTD with indices in Ω_2, the soft decoder decodes the deinterleaved LLR information to get the data packet of $L_s + L_d$ bits
- **Decoding quality judgement module**
- Firstly, for each MTD with its index $\Omega_2|_{\overline{n}}$, $\forall \overline{n} \in [|\Omega_2|_c]$, we calculate the Hamming distances between the decoded signature sequence and the true signature sequence
- Based on the Hamming distances recorded, whether and how to perform the SIC processing will be judged
- If none of the Hamming distances equals zero, we skip the interference cancellation module, then decode the remaining MTDs indexed by $\{\Omega_1 \setminus \Omega_2\}$
- If there exist zero Hamming distance, it indicates there is one or several MTDs with almost perfect decoding quality. we extract those well decoded MTDs, whose index set is denoted as Ω_3, and continue to perform the following signals reconstruction module and interference cancellation module

- **Signals reconstruction module**
- For each MTD with indices in Ω_3, to reconstruct the signal components, we sequentially perform encoding, interleaving, and media modulation (i.e., repeat the BICMM scheme at MTD) according to the decoded bits
- **Interference cancellation module**
- Firstly, we subtract the signal components $\hat{\mathrm{X}}_{[\widetilde{\Omega_3},:]}$ from the received signals Y
- Secondly, we subtract the index set Ω_3 of the MTDs cancellated in the current iteration from the index set Ω_1 of MTDs to be decoded in the following iterations, updated as $\Omega_1 = \{\Omega_1 \setminus \Omega_3\}$
- Meanwhile, we obtain the index set of MTDs that are already subtracted in the previous iterations, denoted as $\Lambda = \{\Omega_0 \setminus \Omega_1\}$
- Thirdly, we update the measurement matrix as $\mathrm{H} = \mathrm{H}_{0[:,\{[KN_t]\setminus\widetilde{\Lambda}\}]}$, where $\mathrm{H}_0 \in \mathbb{C}^{N_r \times (KN_t)}$ is the input channel matrix, $\widetilde{\Lambda}$ denotes the MAPs' index of MTDs indexed by Λ

- **Why does the "Interference cancellation module" help to improve the data decoding performance?**
- Since the sparsity level in the next iteration is reduced with constant dimension N_r of the observations (i.e., the number of receive antennas at the BS), the proposed IDS-AMP detector is capable of achieving the improved decoding performance by using the aforementioned interference cancellation module.
- **Is the "Decoding quality judgement module" necessary?**
- Yes, it prevents the error propagation in the SIC processing

Outline

- An extensive simulation investigation is carried out to evaluate the MTDs' activity detection error rate (ADER), the symbol error rate (SER), and the bit error rate (BER) of the proposed massive access solution
- The ADER is defined as

$$\mathrm{ADER} = \frac{E_m + E_f}{K}, \tag{36}$$

  where E_m is the number of active MTDs missed to be detected, E_f is the number of inactive MTDs falsely detected to be active
- The SER is defined as

$$\mathrm{SER} = \frac{JE_m + E_{\mathrm{symbol}}}{JK_a}, \tag{37}$$

  where E_{symbol} is the number of error symbols of the detected active MTDs, and JK_a is the total number of symbols transmitted by K_a active MTDs within one frame
- The BER is defined as

$$\mathrm{BER} = \frac{\eta JE_m + E_{\mathrm{MED}} + E_{\mathrm{QAM}}}{\eta JK_a}, \tag{38}$$

  where E_{MED} and E_{QAM} are the overall error numbers of media modulated bits and of quadrature amplitude modulated bits for detected active MTDs within one frame, respectively, and ηJK_a is the total bits transmitted by K_a active MTDs

- The number of MTDs is $K = 500$ with $K_a = 50$ active MTDs, where each MTD adopts $N_{\mathrm{RF}} = 2$ RF mirrors for media modulation and 4-QAM ($M = 4$), the throughput is $\eta = N_{\mathrm{RF}} + \log_2 M = 4$ bpcu
- The number of receive antennas is $N_r = 256$, the maximum iteration number is set to $T_0 = 15$, and the Rayleigh MIMO channel model is considered, and the frame length J is set to 12
- **SE of DS-AMP algorithm**
- The number of Monte Carlo simulations is $N_{\mathrm{MC}} = 500$, the maximum number of iterations is $T_{\mathrm{SE}} = 50$, and the terminal threshold is $\varepsilon = 10^{-5}$. Note that since we can obtain the *a* posteriori estimation of the media modulation signals $\mathrm{x}_{k,j}$, $\forall k, j$, in each Monte Carlo simulation, the ADER, BER, and SER of the theoretical SE can be calculated in the same way as those in the DS-AMP algorithm, and then averaged over all the Monte Carlo simulations
- **Coded scenario**
- We consider the Turbo coding with $1/3$ rate and 12 tail bits. The length of the data packet is $L = 120$ with the length of the signature sequence being $L_s = 20$. Hence, after channel encoding, the length of the data packet is $L' = 3L + 12 = 372$ and the frame length is $J = L'/\eta = 93$ for coded media modulation based mMTC. Finally, $\overline{N}$ is set to 5

Benchmarks

- **Benchmark 1**: LMMSE multi-user detector for a traditional uplink mMIMO system, where K_a single-antenna users (after the grant-based scheduling) adopting 16-QAM (for achieving the same throughput 4 bpcu) is supported by mMIMO BS with $N_r = 256$ receive antennas
- **Benchmark 2**: The StrOMP algorithm is used for activity detection and the SIC-SSP algorithm is used for data detection, where the terminal threshold P_{th} for the StrOMP algorithm is set to 1.5
- **Benchmark 3**: A modified DS-AMP algorithm without executing the min-max normalization, where the activity detection method is $\{k|[\widehat{\mathrm{a}}]_k > 0.5\}, \forall k \in [K]$
- **AMP**: Conventional AMP algorithm, where the sparsity level is $\lambda = \frac{K_a}{K}$ and the noise variance σ_w^2 are perfectly known in advance, and the *a* prior probability is replaced by $p\left([\mathrm{x}_{k,j}]_i\right) = (1-\lambda)\delta\left([\mathrm{x}_{k,j}]_i\right) + \frac{\lambda}{M}\sum\limits_{s\in\mathbb{S}}\delta\left([\mathrm{x}_{k,j}]_i - s\right)$
- **TLSSCS**: The cutting-edge TLSSCS algorithm, where the scaling factor $\alpha = 4$
- **PIA-MSMP**: The state-of-the-art PIA-MSMP algorithm with the perfectly known sparsity level

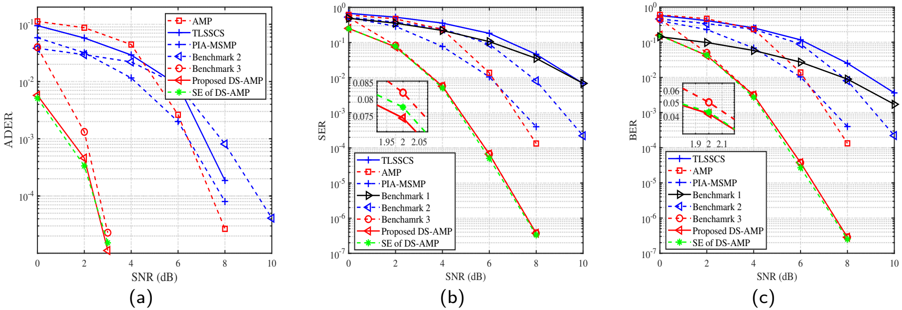

Fig. 5. Performance comparison of different solutions versus SNR: (a) ADER performance comparison; (b) SER performance comparison; (c) BER performance comparison.

- The proposed DS-AMP algorithm outperforms the TLSSCS algorithm, the PIA-MSMP algorithm, and benchmark 2 in terms of ADER, SER, and BER performance
- Our proposed DS-AMP algorithm outperforms the conventional AMP algorithm in ADER, SER, and BER, thanks to the exploitation of the doubly structured sparsity
- The proposed DS-AMP algorithm outperform benchmark 3 at lower SNR (i.e., 0 dB∼2 dB), which verifies the effectiveness of the proposed min-max normalization
- We observe that the SE offers a good tightness compared with the proposed DS-AMP algorithm in ADER, SER, and BER performance

Performance of the Proposed DS-AMP Algorithm

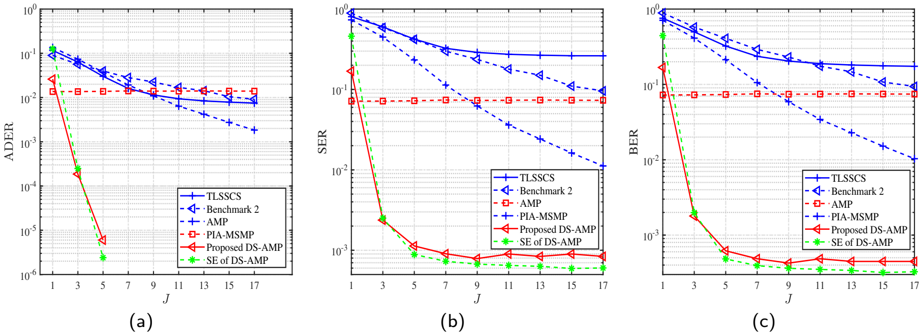

Fig. 6. Performance comparison of different solutions versus the numbers of time slots J within a frame at SNR $=$ 5 dB: (a) ADER performance comparison; (b) SER performance comparison; (c) BER performance comparison.

- Owing to the exploitation of the structured sparsity in the time domain, it can be seen that the advantage of the proposed DS-AMP algorithm over other algorithms in ADER performance becomes more obvious upon increasing J
- If J is small (i.e., $J < 5$), we can obtain improved BER and SER performance with the improvement of ADER performance
- If J is large (i.e., $J > 9$), the BER and SER performance almost stay unaltered against different J due to the unchanged perfect ADER performance

Performance of the Proposed DS-AMP Algorithm

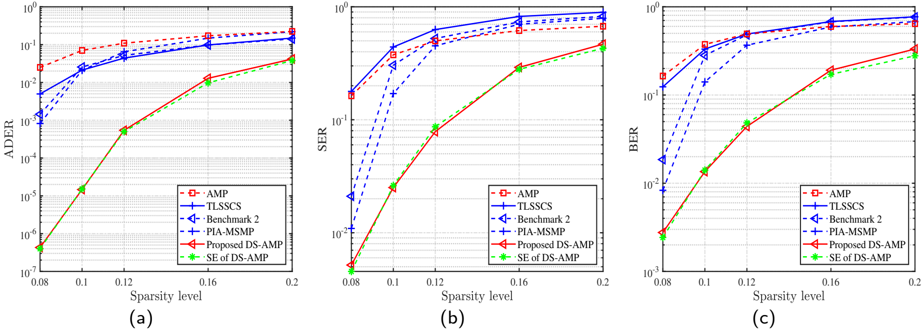

Fig. 7. Performance comparison of different solutions versus the sparsity level $\lambda = \frac{K_a}{K}$, given $K = 500$ and SNR $= 3$ dB: (a) ADER performance comparison; (b) SER performance comparison; (c) BER performance comparison.

- The proposed DS-AMP algorithm outperforms the conventional AMP algorithm, the TLSSCS algorithm, the PIA-MSMP algorithm, and benchmark 2 in ADER, SER, and BER performance

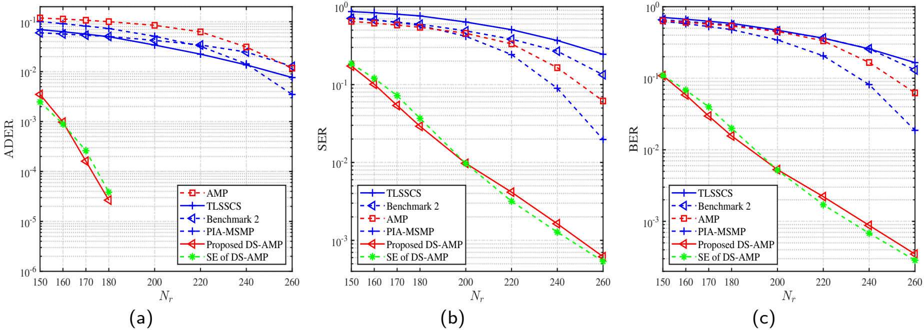

Fig. 8. Performance comparison of different solutions versus the numbers of receive antennas N_r at SNR = 5 dB: (a) ADER performance comparison; (b) SER performance comparison; (c) BER performance comparison.

- Similar conclusion as observed in Fig. 7 can be obtained
- In particular, both Fig. 7 and Fig. 8 verify the superiority and robustness of the proposed DS-AMP algorithm under different system parameters, i.e., the sparsity level or the number of receive antennas, in typical IoT scenarios

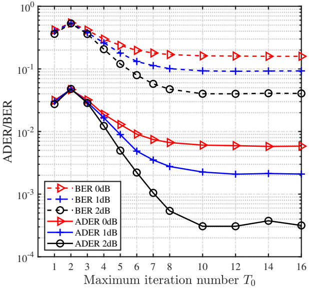

Fig. 9. Performance of the proposed DS-AMP algorithm versus the maximum iteration number T_0.

- We find that the ADER and BER performance of the proposed DS-AMP algorithm converges fast at various SNRs (usually requires less than 15 iterations), which indicates we can adopt the maximum iteration number $T_0 = 15$ for Algorithm 1

Performance of the Proposed DS-AMP Algorithm

Table I: Computational complexity comparison of different algorithms for uncoded media modulation based mMTC

Algorithms	Computational complexity	Complex-valued multiplications[1]	
		$N_r = 128$	$N_r = 256$
Benchmark 1	$\mathcal{O}(JN_rK_a + 2N_rK_a{}^2 + K_a{}^3)$	0.84×10^6	1.56×10^6
DS-AMP	$\mathcal{O}[T_0JKN_t(\frac{5}{2}N_r + \lvert\mathbb{S}\rvert_c + \frac{1}{4})]$	1.17×10^8	2.32×10^8
AMP	$\mathcal{O}[T_0JKN_t(\frac{5}{2}N_r + \lvert\mathbb{S}\rvert_c + \frac{1}{4})]$	1.17×10^8	2.32×10^8
Benchmark 3	$\mathcal{O}[T_0JKN_t(\frac{5}{2}N_r + \lvert\mathbb{S}\rvert_c + \frac{1}{4})]$	1.17×10^8	2.32×10^8
TLSSCS	$\mathcal{O}\{(JN_rK_a + 2N_rK_a{}^2 + K_a{}^3) + (K_a + 1)[N_r{}^2(KN_t + J) + N_rJKN_t] + \sum_{s=1}^{K_a+1}[N_r{}^2 + 2N_r(sN_t)^2 + (sN_t)^3]\}$	2.14×10^9	7.53×10^9
PIA-MSMP	$\mathcal{O}\{3JK_aN_r(N_t + 1) + (K_a + 1)[N_r{}^2(KN_t + J) + N_rJKN_t] + \sum_{s=1}^{K_a}[N_r{}^2 + 2N_r(sN_t)^2 + (sN_t)^3]\}$	2.12×10^9	7.50×10^9
Benchmark 2	$\mathcal{O}\{K_aJKN_tN_r + \sum_{s=1}^{K_a}[JN_r(s+2s^2+2(sN_t)^2)+J(s^3+(sN_t)^3)] + \sum_{s=1}^{K_a}[JN_r(s + 2s^2 + 2(sN_t)^2) + J(s^3 + (sN_t)^3)]\}$	4.82×10^9	8.16×10^9

[1] The order of complex-valued multiplications is obtained under parameters $J = 12$, $N_t = 4$, $K = 500$, $K_a = 50$, $T_0 = 15$, $\lvert\mathbb{S}\rvert_c = 4$.

- The complexity of the DS-AMP algorithm scales linearly with the number of receive antennas N_r
- The computational complexities of both TLSSCS and PIA-MSMP algorithms can be approximately proportional to the square of N_r

- **Benchmark 4**: The proposed DS-AMP algorithm adopting uncoded media modulation and a hard decision (i.e., perform a hard decision according to the output signal $\mathrm{X} \in \mathbb{C}^{KN_t \times J}$ from **Algorithm 1** to get the demodulated binary bits)
- **Benchmark 5**: The proposed DS-AMP algorithm adopting coded media modulation and soft decision, while the processing of interleaving/deterleaving and SIC is not adopted
- **Benchmark 6**: The proposed DS-AMP algorithm adopting coded media modulation, interleaving/deinterleaving, and soft decision, while the SIC processing is not adopted
- **Benchmark 7**: The proposed IDS-AMP scheme except that the proposed decoding quality judgement is removed and let Ω_3 equal Ω_2

Performance of the Proposed IDS-AMP Scheme

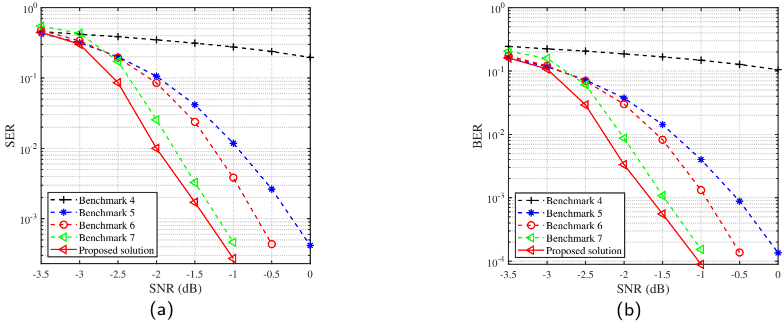

Fig. 10. Performance of the proposed SIC-based massive access solution in comparison with the benchmarks: (a) SER performance comparison; (b) BER performance comparison.

- The worst performance is achieved by benchmark 4, which indicates that the necessity in adopting the channel coding and soft decoding for the improving data decoding performance
- The superiority of benchmark 6 over benchmark 5 verifies the effectiveness of the proposed BICMM in overcoming the spatial-selective fading channels among different media modulation signals

- The superiority of benchmark 7 over benchmark 6 in high SNR regime (i.e., larger than -2 dB) verifies the effectiveness of SIC processing
- Benchmark 7 is observed to be inferior to benchmark 5 in low SNR regime (i.e., -3.5 dB$\sim$-2.5 dB), since the SIC at low SNR can degrade the performance
- The superiority of IDS-AMP scheme over benchmark 7 verifies the data decoding gain achieved by the proposed decoding quality judgement module

Conclusions

- **Uncoded media modulation based mMTC**
- The proposed DS-AMP algorithm reliably tackles the DADD problem, and outperforms the state-of-the-art algorithms
- The SE theoretically characterizes the DS-AMP algorithm
- **Coded media modulation based mMTC**
- The developed BICMM scheme is effective
- The SIC-based IDS-AMP scheme, including a dedicated data packet and an IDS-AMP detector, improves the data decoding performance

References

- L. Qiao and Z. Gao, "Joint active device and data detection for massive MTC relying on spatial modulation," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Seoul, Korea (South), 2020, pp. 1-6.
- 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.
- 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.
- 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, Sept. 2018.
- C. Wei, H. Liu, Z. Zhang, J. Dang, and L. Wu, "Approximate message passing-based joint user activity and data detection for NOMA," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 640-643, Mar. 2017.
- 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.

- Y. Du, B. Dong, Z. Chen *et al.*, "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.
- 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, Sept. 2019.
- L. Qiao, J. Zhang, Z. Gao, S. Chen, and L. Hanzo, "Compressive sensing based massive access for IoT relying on media modulation aided machine type communications," *IEEE Trans. Veh. Technol.*, vol. PP, no. PP, Jul. 2020.
- X. Ma, S. Guo, and D. Yuan, "Improved compressed sensing-based joint user and symbol detection for media-based modulation-enabled massive machine-type communications," *IEEE Access*, vol. 8, pp. 70058-70070, 2020.
- 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.
- X. Meng, S. Wu, L. Kuang, D. Huang, and J. Lu, "Multi-user detection for spatial modulation via structured approximate message passing," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1527-1530, Aug. 2016.

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.

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, Sept. 2018.

Thanks for your attention!

Q & A