

Compressive Massive Random Access for Massive Machine-Type Communications (mMTC)

Malong Ke^{†*}, Zhen Gao^{†*}, Yongpeng Wu[‡], and Xiangming Meng[§]

[†]Advanced Research Institute of Multidisciplinary Science, Beijing Institute of Technology, Beijing 100081, China

^{*}School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

[‡]Shanghai Jiao Tong University, Shanghai 200240, China

[§]Huawei Technologies Co.Ltd., Shanghai 201206, China

Email: gaozhen16@bit.edu.cn

Abstract—In future wireless networks, one fundamental challenge for massive machine-type communications (mMTC) lies in the reliable support of massive connectivity with low latency. Against this background, this paper proposes a compressive sensing (CS)-based massive random access scheme for mMTC by leveraging the inherent sporadic traffic, where both the active devices and their channels can be jointly estimated with low overhead. Specifically, we consider devices in the uplink massive random access adopt pseudo random pilots, which are designed under the framework of CS theory. Meanwhile, the massive random access at the base stations (BS) can be formulated as the sparse signal recovery problem by leveraging the sparse nature of active devices. Moreover, by exploiting the structured sparsity among different receiver antennas and subcarriers, we develop a distributed multiple measurement vector approximate message passing (DMMV-AMP) algorithm for further improved performance. Additionally, the state evolution (SE) of the proposed DMMV-AMP algorithm is derived to predict the performance. Simulation results demonstrate the superiority of the proposed scheme, which exhibits a good tightness with the theoretical SE.

Index Terms—Massive random access, massive machine-type communications (mMTC), compressive sensing (CS).

I. INTRODUCTION

Driven by the Internet-of-Things (IoT), how to support massive machine-type communications (mMTC) enabling massive connectivity with tens of billions of machine-type devices has been challenging the current wireless networks [1]. For mMTC, how to reliably support the access and estimate the associated channels of active devices triggered by external events dynamically performs an important role in uplink systems.

Conventional grant-based access control requires the additional control signaling and the prediction of uplink access requests for the granting of resources [2]. One representative is the strongest-user collision resolution (SUCRe) scheme [3], which designs a sophisticated protocol efficiently supporting devices in overloaded networks. However, the grant-based solutions may suffer from the inefficient schedule and difficult design for the access in mMTC [2]. As an alternative, grant-free access protocol without prior scheduling permission has recently attracted significant attention [4]. By exploiting the inherent sporadic traffic, a compressive sensing (CS)-based uplink grant-free non-orthogonal multiple access (NOMA)

scheme has been proposed to further reduce the overhead [5], [6], while only single-antenna is considered at the base station (BS). In multi-antenna systems, a modified Bayesian compressive sensing (BCS) algorithm is proposed for NOMA [7], which exploits the structured sparsity among different receiver antennas to enhance the performance. However, the BCS algorithm may not work efficiently in mMTC. Recently, a low-complexity iterative algorithm termed approximate message passing (AMP) has been proposed for massive random access [8]–[10], which can support the massive connectivity efficiently. However, most prior work [2]–[5], [7], [8] are limited to the single-carrier systems. Besides, the work in [8] assumes the full knowledge of the prior distribution and noise variance, which is an impractical assumption.

In this paper, we make further research in the direction of massive random access for mMTC, where we consider the more practical frequency-selective fading channels and the work is extended to the multi-carrier multi-antenna systems. Specifically, by exploiting the structured sparsity among different receiver antennas and subcarriers, we develop a distributed multiple measurement vector approximate message passing (DMMV-AMP) algorithm for further improved performance. Moreover, resorting to the expectation maximization (EM) algorithm, the proposed DMMV-AMP algorithm can learn unknown hyperparameters of the prior distribution and noise variance. We further derive the state evolution (SE) of DMMV-AMP algorithm to characterize the performance of the proposed scheme.

II. SYSTEM MODEL

We consider a typical uplink mMTC system with one BS equipped with M antennas and K single-antenna devices, where OFDM with N subcarriers is adopted to combat the time dispersive channels, and P pilots are uniformly allocated across N subcarriers. For the subchannel at the p -th subcarrier ($1 \leq p \leq P$), the received signal $\mathbf{y}_{p,k}^t \in \mathbb{C}^{M \times 1}$ at the BS from the k -th device during the t -th OFDM symbol can be expressed as

$$\mathbf{y}_{p,k}^t = \mathbf{h}_{p,k} s_{p,k}^t + \mathbf{w}_p^t, \quad (1)$$

in which $\mathbf{h}_{p,k} \in \mathbb{C}^{M \times 1}$ is the subchannel associated with the k -th device at the p -th subcarrier, $s_{p,k}^t$ generated from i.i.d.

standard complex Gaussian distribution, is the random access pilot from the k -th device, and \mathbf{w}_p^t denotes the additive white Gaussian noise (AWGN) at the BS. The channel vectors can be modeled as $\mathbf{h}_{p,k} = \tau_k \tilde{\mathbf{h}}_{p,k}$, where τ_k is the large-scale fading caused by the path loss and shadowing fading, and $\tilde{\mathbf{h}}_{p,k}$ is the small-scale fading channels [11]. For the typical mMTC system, only a small subset of devices are activated to access the BS within a given coherent time interval, and the device activity indicator can be denoted as

$$\alpha_k = \begin{cases} 1, & \text{if the } k\text{th device is active,} \\ 0, & \text{otherwise,} \end{cases} \quad \forall k. \quad (2)$$

Meanwhile, we define the set of active devices as $\mathcal{K} = \{k : \alpha_k = 1, k = 1, \dots, K\}$ and the number of active devices is $K_a = |\mathcal{K}|$. Hence, the received signals from all devices can be written as

$$\mathbf{y}_p^t = \sum_{k=1}^K \alpha_k \mathbf{h}_{p,k} s_{p,k}^t + \mathbf{w}_p^t = \mathbf{H}_p \mathbf{s}_p^t + \mathbf{w}_p^t, \quad (3)$$

in which $\mathbf{H}_p = [\alpha_1 \mathbf{h}_{p,1}, \dots, \alpha_K \mathbf{h}_{p,K}] \in \mathbb{C}^{M \times K}$, $\mathbf{s}_p^t = [s_{p,1}^t, \dots, s_{p,K}^t]^T \in \mathbb{C}^{K \times 1}$. Furthermore, the received signals over G ($G \ll K$) successive time slots are jointly exploited to detect the uncoordinated random access devices as

$$\mathbf{Y}_p = \mathbf{S}_p \mathbf{X}_p + \mathbf{W}_p, \quad (4)$$

where $\mathbf{Y}_p = [\mathbf{y}_p^1, \dots, \mathbf{y}_p^G]^T \in \mathbb{C}^{G \times M}$, $\mathbf{S}_p = [\mathbf{s}_p^1, \dots, \mathbf{s}_p^G]^T \in \mathbb{C}^{G \times K}$, $\mathbf{X}_p = [\mathbf{x}_{p,1}, \dots, \mathbf{x}_{p,K}]^T \in \mathbb{C}^{K \times M}$, and $\mathbf{x}_{p,k} = \alpha_k \mathbf{h}_{p,k}$ denotes the channel vector of the k -th device at the p -th subcarrier.

In this paper, we consider the grant-free random access, where the indexes set \mathcal{K} and channel vector $\mathbf{h}_{p,k}$ associated with the active devices can be jointly estimated from the noisy measurements \mathbf{Y}_p and the random access pilot matrix \mathbf{S}_p known at the BS. On the other hand, due to the sporadic traffic of devices in typical mMTC, only a smaller number of devices are active, i.e., $K_a \ll K$. This observation motivates us to formulate the joint active devices detection and channel estimation for massive random access as a CS problem. Furthermore, we observe that different columns of \mathbf{X}_p share the common support, and $\{\mathbf{X}_p\}_{p=1}^P$ have the common sparsity pattern, namely,

$$\text{supp}\{\mathbf{X}_1\} = \text{supp}\{\mathbf{X}_2\} = \dots = \text{supp}\{\mathbf{X}_P\}, \quad (5)$$

which inspires us to solve (4) with the distributed multiple measurement vector (DMMV) CS theory for further improved performance [12].

III. CS-BASED MASSIVE RANDOM ACCESS SCHEME

In this section, we propose the DMMV-AMP algorithm for massive random access, where the sparse traffic observed from multiple antennas and multiple subcarriers is considered. Especially, we integrate the hyperparameters learning manner of AMP with nearest neighbor sparsity pattern learning (AMP-NNSPL) algorithm [13], [14] into our scheme for further improved support estimation performance.

A. DMMV-AMP Algorithm

We first consider the massive random access problem for the p -th subcarrier, where the solution is listed in **Algorithm 1** named MMV-AMP algorithm. According to the theory of statistical signal processing, the minimum mean square error (MMSE) estimation of (4) is the posterior means, which can be expressed as

$$\hat{x}_{km} = \int x_{km} p(x_{km} | \mathbf{Y}) dx_{km}, \quad \forall k, m, \quad (6)$$

where the subcarrier index p in $x_{p,km}$ and \mathbf{Y}_p is dropped to simplify the notations, and $x_{p,km}$ is the (k, m) -th element of the matrix \mathbf{X}_p . The marginal posterior probability is calculated by $p(x_{km} | \mathbf{Y}) = \int p(\mathbf{X} | \mathbf{Y}) d\mathbf{X}_{\setminus km}$, which involves the multi-dimensional integrals, and the $\mathbf{X}_{\setminus km}$ denotes the collection of $\{x_{ij}\}_{1 \leq i \leq K, i \neq k, 1 \leq j \leq M, j \neq m}$. The joint posterior probability can be computed according to Bayesian rule in a factored form as

$$p(\mathbf{X} | \mathbf{Y}) = \frac{1}{Z} \prod_{g=1}^G \prod_{m=1}^M p(y_{gm} | \mathbf{X}) \prod_{k=1}^K \prod_{m=1}^M p_0(x_{km}), \quad (7)$$

in which Z is the normalization factor. In this paper, we consider a flexible spike and slab prior distribution which can match the real distribution of channels well

$$p_0(\mathbf{X}) = \prod_{k=1}^K \prod_{m=1}^M [(1 - \lambda_{km}) \delta(x_{km}) + \lambda_{km} f(x_{km})], \quad (8)$$

where $\lambda_{km} \in (0, 1)$ is the sparse ratio, i.e., the probability of x_{km} being nonzero, $\delta(x_{km})$ is the Dirac delta function, $f(x_{km})$ is the distribution of the nonzero entries. Under the assumption of AWGN, the likelihood function of y_{gm} can be expressed as

$$p(y_{gm} | \mathbf{X}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \left| y_{gm} - \sum_k S_{gk} x_{km} \right|^2\right), \quad (9)$$

where σ^2 is the variance of noise. As the marginal probability is hard to compute, we resort to the belief propagation (BP), which provides low-complexity heuristics for approximating $p(x_{km} | \mathbf{Y})$. A key remark is that AMP decouples the matrix estimation problem (4) into KM scalar problems in the asymptotic regime, i.e., $K \rightarrow \infty$ while λ is fixed, and the posterior distribution of x_{km} is approximated as [?], [15]

$$p(x_{km} | R_{km}^t, \Sigma_{km}^t) = \frac{1}{Z} p_0(x_{km}) \mathcal{CN}(x_{km}; R_{km}^t, \Sigma_{km}^t), \quad (10)$$

where t denotes the t -th iteration, R_{km}^t and Σ_{km}^t are calculated as step 4 in **Algorithm 1**. We have the assumption that $f(x_{km}) = \mathcal{CN}(x_{km}; \mu, \tau)$, which is a flexible prior model for the channels. By exploiting this prior model, the posterior distributions are obtained by (10) as

$$p(x_{km} | R_{km}^t, \Sigma_{km}^t) = (1 - \pi_{km}^t) \delta(x_{km}) + \pi_{km}^t \mathcal{N}(x_{km}; A_{km}^t, \Delta_{km}^t), \quad (11)$$

where

$$A_{km}^t = \frac{\tau R_{km}^t + \Sigma_{km}^t \mu}{\Sigma_{km}^t + \tau}, \quad \Delta_{km}^t = \frac{\tau \Sigma_{km}^t}{\tau + \Sigma_{km}^t},$$

$$\mathcal{L} = \frac{1}{2} \ln \frac{\Sigma_{km}^t}{\Sigma_{km}^t + \tau} + \frac{|R_{km}^t|^2}{2\Sigma_{km}^t} - \frac{|R_{km}^t - \mu|^2}{2(\Sigma_{km}^t + \tau)},$$

$$\pi_{km}^t = \frac{\lambda_{km}^t}{\lambda_{km}^t + (1 - \lambda_{km}^t) \exp(-\mathcal{L})}, \quad (12)$$

in which π_{km}^t is the equivalent sparse ratio. The posterior mean and posterior variance can now be explicitly calculated as

$$g_a(R_{km}^t, \Sigma_{km}^t) = \pi_{km}^t A_{km}^t, \quad (13)$$

$$g_c(R_{km}^t, \Sigma_{km}^t) = \pi_{km}^t (|A_{km}^t|^2 + \Delta_{km}^t) - |g_a|^2. \quad (14)$$

Thus, the AMP iterations can be calculated in the matrix-vector forms instead of the integrals. However, conventional AMP algorithm assumes full knowledge of the prior distribution and noise variance, which is an impractical assumption. Resorting to the EM algorithm, the DMMV-AMP algorithm can learn the unknown hyperparameters, i.e., $\theta = \{\mu, \tau, \sigma^2, \lambda_{km}, \forall k, m\}$, which consists of two steps

$$Q(\theta, \theta^t) = \mathbb{E}(\ln p(\mathbf{X}|\mathbf{Y})|\mathbf{Y}; \theta^t), \quad (15)$$

$$\theta^{t+1} = \arg \max_{\theta} Q(\theta, \theta^t), \quad (16)$$

where $\mathbb{E}(\cdot|\mathbf{Y}; \theta^t)$ denotes expectation conditioned on measurements \mathbf{Y} with parameters θ^t . Two problems arise in the EM algorithm, the computation of $p(\mathbf{X}|\mathbf{Y}; \theta^t)$ is high complexity and the joint optimization of θ is difficult. However, the approximation of $p(\mathbf{X}|\mathbf{Y}; \theta^t)$ is given by the AMP as (10). Moreover, we resort to incremental EM algorithm [16], i.e., θ is updated one element at a time while other parameters are fixed. By taking the derivative of (15) where the fixed parameters are considered as constant and zeroing the derivative, the update rules of the hyperparameters are obtained as

$$(\sigma^2)^{t+1} = \frac{1}{GM} \sum_g \sum_m \left[\frac{|y_{gm} - Z_{gm}^t|^2}{|1 + V_{gm}^t/(\sigma^2)^t|^2} + \frac{(\sigma^2)^t V_{gm}^t}{(\sigma^2)^t + V_{gm}^t} \right], \quad (17)$$

$$\tau^{t+1} = \frac{\sum_k \sum_m \pi_{km}^t [|\mu^t - A_{km}^t|^2 + \Delta_{km}^t]}{\sum_k \sum_m \pi_{km}^t}, \quad (18)$$

$$\mu^{t+1} = \frac{\sum_k \sum_m \pi_{km}^t A_{km}^t}{\sum_k \sum_m \pi_{km}^t}, \quad \lambda_{km}^{t+1} = \pi_{km}^t. \quad (19)$$

As the EM algorithm may converge to the local extremum of the likelihood function, the initialization of the hyperparameters is very important, which are given as

$$\lambda_{km}^0 = \lambda^0 = \frac{G}{K} \left\{ \max_{c>0} \frac{1-2K[(1+c)^2\Phi(-c)-c\phi(c)]/G}{1+c^2-2[(1+c)^2\Phi(-c)-c\phi(c)]} \right\}, \quad (20)$$

$$(\sigma^2)^0 = \frac{1}{M} \sum_m \frac{\|\mathbf{y}_m\|^2}{(\text{SNR}^0+1)G}, \quad (21)$$

$$\tau^0 = \frac{1}{M} \sum_m \frac{\|\mathbf{y}_m\|_2^2 - M(\sigma^2)^0}{\|\mathbf{S}\|_F^2 \lambda^0}, \quad \mu^0 = 0, \quad (22)$$

where $\Phi(-c)$ and $\phi(c)$ are, respectively, the cumulative distribution function and the probability density function of the

Algorithm 1 MMV-AMP Algorithm

Input: Noisy observation \mathbf{Y}_p , pilot matrix \mathbf{S}_p , the maximum number of iteration T_{\max} and termination threshold ε .

Output: The estimated channel vectors $\hat{\mathbf{x}}_k, \forall k$.

- 1: Set $t = 1$, $\hat{x}_{km}^1 = \int x_{km} p_0(x_{km}) dx_{km}$, $v_{km}^1 = \int |x_{km} - \hat{x}_{km}^1|^2 p_0(x_{km}) dx_{km}$, $V_{gm}^0 = 1$, $Z_{gm}^0 = y_{gm}$, $\forall k, m, g$. Initialize the hyperparameters as (20)-(22).
 - 2: **repeat**
 - 3: **Factor nodes update:** $g = 1, \dots, G$; $m = 1, \dots, M$
 $V_{gm}^t = \sum_k |S_{gk}|^2 v_{km}^t$
 $Z_{gm}^t = \sum_k S_{gk} \hat{x}_{km}^t - \frac{V_{gm}^t}{(\sigma^2)^t + V_{gm}^t} (y_{gm} - Z_{gm}^{t-1})$
 - 4: **Variable nodes update:** $k = 1, \dots, K$; $m = 1, \dots, M$
 $\Sigma_{km}^t = \left[\sum_g \frac{|S_{gk}|^2}{(\sigma^2)^t + V_{gm}^t} \right]^{-1}$
 $R_{km}^t = \hat{x}_{km}^t + \Sigma_{km}^t \sum_g \frac{S_{gk}^* (y_{gm} - Z_{gm}^t)}{(\sigma^2)^t + V_{gm}^t}$
 $\hat{x}_{km}^{t+1} = g_a(R_{km}^t, \Sigma_{km}^t)$
 $v_{km}^{t+1} = g_c(R_{km}^t, \Sigma_{km}^t)$
 - 5: Update the hyperparameters as (17)-(19), $t = t + 1$.
 - 6: **until** $t > T_{\max}$ or $\left\| \hat{\mathbf{X}}_p^{t+1} - \hat{\mathbf{X}}_p^t \right\|_F < \varepsilon \left\| \hat{\mathbf{X}}_p^t \right\|_F$.
 - 7: **return** $\hat{x}_{km}, \forall k, m$
-

standard normal distribution, c denotes the maximum sparse ratio with fixed G/K , and $\text{SNR}^0 = 100$.

Moreover, the MMV-AMP algorithm is further extended to the distributed model with all subchannel matrix \mathbf{X}_p , $\forall p \in [P]$ being jointly estimated, which is termed as the DMMV-AMP algorithm. For all subcarriers, the messages are parallelly updated as step 3, step 4 in **Algorithm 1**. Notice that the columns of channel matrix \mathbf{X}_p share the common support, which can be utilized to improve the accuracy of the support estimation. We further exploit the common sparsity described in (5), and the estimated sparse ratio is updated by $\hat{\lambda}_{km}^{t+1} = \hat{\lambda}_k^{t+1} = \frac{1}{MP} \sum_p \sum_m \pi_{p,km}^t$. The modified update rules of the mean and the variance of nonzero entries are written as

$$\mu^{t+1} = \frac{1}{P} \sum_p \mu_p^t, \quad \tau^{t+1} = \frac{1}{P} \sum_p \tau_p^t. \quad (23)$$

Additionally, compared with conventional AMP-based solutions, we consider the more practical frequency-selective fading channels, thus the equivalent signal-to-noise ratio (SNR) varies with the subcarrier index p . We assume there is an adaptive power control at the BS to eliminate the effect of the large scale fading, and all the subcarriers have the same signal power, thus the frequency-selective fading effect is included in the different noise variances $(\sigma^2)_p^t$ which are respectively calculated as (17). With the estimated channel matrix $\hat{\mathbf{X}}_n$ and the characteristic of the sparsity pattern, we develop a threshold-based activity detector defined as follows

$$\hat{\alpha}_k = \begin{cases} 1, & \sum_p \sum_m r(\hat{x}_{p,km}) \geq p_{\text{th}} MP, \\ 0, & \sum_p \sum_m r(\hat{x}_{p,km}) < p_{\text{th}} MP. \end{cases} \quad (24)$$

In (24), $r(\cdot)$ denotes a mapper where $r(\hat{x}_{km}) = 1$ if $|\hat{x}_{km}| > 10^{-10}$, otherwise $r(\hat{x}_{km}) = 0$, and p_{th} is a tunable parameter, which will influence the probability of missed detection and false alarm. In this paper, we consider $p_{th} = 0.99$. Finally, given that the k -th device is declared active, its channel is estimated as $\hat{\mathbf{h}}_{p,k} = \hat{\mathbf{x}}_{p,k}$, where $\hat{\mathbf{x}}_{p,k}$ is the k -th row of the estimated channel matrix $\hat{\mathbf{X}}_p$.

B. State Evolution

Statistical properties of AMP algorithms allow us to accurately analyze their performance in the asymptotic regime [17]. In this subsection, we use SE to characterize the mean square error (MSE) performance of the proposed scheme. The MSE of the estimation and the variance of the estimated signal are defined as

$$E^t = \frac{1}{KM} \sum_k \sum_m |\hat{x}_{km}^t - x_{km}^t|^2, \quad V^t = \frac{1}{KM} \sum_k \sum_m v_{km}^t.$$

Define the random variable $X_0 \sim p_0(\mathbf{X})$, $Z \sim \mathcal{CN}(z; 0, 1)$, then the mean and variance of the posterior distribution are respectively expressed as [17]

$$R^t = x_0 + \sqrt{\frac{\sigma_0^2 + E^t}{G/K}} z, \quad \Sigma^t = \frac{(\sigma^2)^t + V^t}{G/K}, \quad (25)$$

which show that the AMP algorithm decouples the vector or matrix estimation problem into independent scalar problems, $\sigma_0^2 = 1/\text{SNR}$, and the E^t and V^t can be updated as

$$E^t = \mathbb{E} \left[|g_a(R^t, \Sigma^t) - x_0|^2 \right], \quad V^t = \mathbb{E} [g_c(R^t, \Sigma^t)], \quad (26)$$

with $\mathbb{E}[\cdot]$ denoting the expectation with respect to the random variable X_0 and Z . In contrast to the conventional AMP algorithms assuming full knowledge of the prior distribution, the SE of DMMV-AMP algorithm also need to track the update rules of the hyperparameters θ , which are given as follows

$$\pi^{t+1} = \mathbb{E}(\pi^t), \quad \tau^{t+1} = \mathbb{E}(\tau^t), \quad \mu^{t+1} = \mathbb{E}(\mu^t), \quad (27)$$

$$(\sigma^2)^{t+1} = \frac{\sigma_0^2 + E^t}{[1 + V^t / (\sigma^2)^t]} + \frac{(\sigma^2)^t V^t}{(\sigma^2)^t + V^t}, \quad (28)$$

where π^t , τ^t , μ^t are calculated as (12), (18) and (19).

IV. SIMULATION RESULTS

In this section, we provide the simulated and analytical results of the proposed scheme. Consider an uplink mMTC system with one BS equipped with $M = 32$ antennas and P pilots are uniformly allocated in $N = 2048$ subcarriers. We assume $K = 1000$ potential devices are randomly distributed in the cell with radius 1km and $K_a = 100$ devices are active at a time. The carrier frequency is 2GHz, the bandwidth is 10MHz, and the received signal-to-noise ratio is $\text{SNR} = 20\text{dB}$. We set $T_{\max} = 200$ and $\varepsilon = 10^{-5}$. The performance is evaluated by the probability of error detection P_e and the normalized MSE

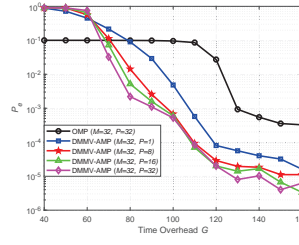


Fig. 1. The probability of error detection comparison as the function of G .

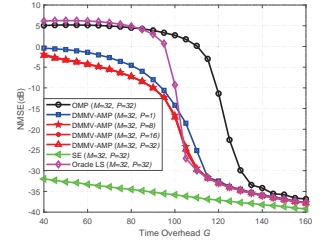


Fig. 2. The NMSE performance comparison and the SE of DMMV-AMP.

(NMSE) under various time overhead G and pilot sequence P , which are defined as

$$P_e = \frac{\sum_k |\hat{\alpha}_k - \alpha_k|}{K}, \quad \text{NMSE} = 10 \log \frac{\sum_p \|\hat{\mathbf{X}}_p - \mathbf{X}_p\|_F^2}{\sum_p \|\mathbf{X}_p\|_F^2}.$$

Fig. 1 examines the performance of device activity detector achieved by the DMMV-AMP algorithm and OMP algorithm [18]. It can be observed that the P_e of DMMV-AMP based scheme decreases over G rapidly, but P_e of the OMP based scheme remains unchanged when $G < 110$. There exists a significant performance gap between the DMMV-AMP based scheme and the OMP based scheme when $G > 80$, which shows that DMMV-AMP based scheme can significantly reduce the access latency when the same performance is considered. Further, with the fixed M , when the common support among multiple carriers is leveraged, the performance can be further improved.

Fig. 2 verifies the NMSE performance of the proposed scheme, OMP based scheme and the oracle LS based scheme with the known support set of the sparse channel matrix. It shows that when time overhead is large enough, both DMMV-AMP algorithm and OMP algorithm can approach the performance of the oracle LS based scheme, since the support is estimated exactly in this case. However, the proposed scheme outperforms the OMP based scheme when $G < 140$, and its performance becomes better when P increases. An important observation is that when the pilot length is less than the active devices ($G < K_a$), the proposed scheme can still work very well by exploiting the structured sparsity shared by different subcarriers and receiver antennas. The proposed DMMV-AMP based scheme can even outperform the oracle LS when $G < 100$. In addition, the performance of the proposed scheme is well predicted by the state evolution when the time overhead is large enough.

V. CONCLUSION

A CS-based massive random access scheme has been proposed for uplink mMTC systems, which can significantly reduce the access latency. By exploiting the structured sparsity among multiple BS antennas and multiple carriers, we propose a DMMV-AMP algorithm, and its SE is also derived to analyze the performance. Simulation results demonstrate that the pro-

posed scheme outperforms its counterparts with significantly reduced access latency.

REFERENCES

- [1] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59-65, Sep. 2016.
- [2] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86-93, Jun. 2013.
- [3] E. Björnson, E. de Carvalho, J. H. Sørensen, E. G. Larsson, and P. Popovski, "A random access protocol for pilot allocation in crowded massive MIMO systems," *IEEE Trans. Wireless. Commun.*, vol. 16, no. 4, pp. 2220-2234, Apr. 2017.
- [4] Z. Zhang, X. Wang, Y. Zhang, and Y. Chen, "Grant-free rateless multiple access: A novel massive access scheme for internet of things," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2019-2022, Oct. 2016.
- [5] 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, July. 2016.
- [6] Z. Gao, L. Dai, S. Han, C. I. Z. Wang, and L. Hanzo, "Compressive sensing techniques for next-generation wireless communications," *IEEE Wireless. Commun.*, vol. 25, no. 3, pp. 144-153, June. 2018.
- [7] X. Xu, X. Rao, and V. K. N. Lau, "Active user detection and channel estimation in uplink CRAN systems," in *Proc. Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 2727-2732.
- [8] L. Liu and W. Yu, "Massive connectivity with massive MIMO-Part I: Device activity detection and channel estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2933-2946, Jun. 2018.
- [9] X. Meng, S. Wu, L. Kuang, and J. Lu, "An expectation propagation perspective on approximate message passing," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1194-1197, Aug. 2015.
- [10] X. Meng, S. Wu, and J. Zhu, "A unified bayesian inference framework for generalized linear models," *IEEE Signal Process. Lett.*, vol. 25, no. 3, pp. 398-402, Mar. 2018.
- [11] X. Lin, S. Wu, L. Kuang, Z. Ni, X. Meng, and C. Jiang, "Estimation of sparse massive MIMO-OFDM channels with approximately common support," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1179-1182, May. 2017.
- [12] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169-6183, Dec. 2015.
- [13] X. Meng, S. Wu, L. Kuang, D. Huang, and J. Lu, "Approximate message passing with nearest neighbor sparsity pattern learning," [Online]. arXiv:1601.00543, 2016.
- [14] H. Cao, J. Zhu, and Z. Xu, "Adaptive one-bit quantisation via approximate message passing with nearest neighbour sparsity pattern learning," *IET Signal Process.*, vol. 12, no. 5, pp. 629-635, 2018.
- [15] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *Proc. Inf. Theory Workshop (ITW)*, Jan. 2010, pp. 1-5.
- [16] R. M. Neal and G. E. Hinton, "A view of the EM algorithm that justifies incremental, sparse, and other variants," in *Learning in graphical models*, Springer, 1998, pp. 355-368.
- [17] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: II. analysis and validation," in *Proc. Inf. Theory Workshop (ITW)*, Jan. 2010, pp. 1-5.
- [18] J. Lee, G. T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370-2386, Jun. 2016.