Received 12 July 2023; revised 7 August 2023; accepted 20 August 2023. Date of publication 1 September 2023; date of current version 18 September 2023.
Digital Object Identifier 10.1109/OJCOMS.2023.3309383

Massive Access Made Simple: Bloom Filter-Based Coding Under Hard-Decision Envelope Detection

RUI DENG [1,2] AND WENYI ZHANG [1,2] (Senior Member, IEEE)

[1] Key Laboratory of Wireless-Optical Communications, Chinese Academy of Sciences, Hefei 230027, China
[2] School of Information Science and Technology, University of Science and Technology of China, Hefei 230027, China
CORRESPONDING AUTHOR: W. ZHANG (e-mail: wenyizha@ustc.edu.cn)

This work was supported by the National Key Research and Development Program of China under Grant 2018YFA0701603.
This work is an expanded version of the 23rd International Workshop Signal Processing Advances Wireless Communications (SPAWC), Oulu, Finland, [1] [DOI: 10.1109/SPAWC51304.2022.983399].

ABSTRACT Motivated by the demand of cheap hardware and low computational complexity in massive connectivity, we propose a scheme in which transmitters send data encoded by Bloom filter with On-Off Keying (OOK) modulation and receiver performs hard-decision envelope detection on received signals. For frequency-flat fading scenarios with inter-user synchronization (IUS), we develop a Noisy-Combinatorial Orthogonal Matching Pursuit (NCOMP) decoding strategy, and for frequency-selective fading scenarios with IUS, we propose an improved decoding strategy, called multipath-NCOMP (MNCOMP). In addition, we propose a sliding window method to adapt the two decoding strategies to scenarios without IUS. Based on a many-access channel (MnAC) model, we study the asymptotic performance of the proposed scheme for activity recognition and message transmission tasks. Theoretical analysis guarantees that the error probability of our scheme vanishes asymptotically with the number of users, and also shows that the MNCOMP strategy is superior to the NCOMP strategy in frequency-selective fading MnACs. In addition, the proposed sliding window method can ensure that our scheme is immune to the lack of IUS. Numerical experiments corroborate our asymptotic analytical results, for finite number of users.

INDEX TERMS Activity recognition cost, Bloom filter, envelope detection, frequency-flat fading, frequency-selective fading, inter-user synchronization, many-access channel, message transmission cost, noisy-combinatorial orthogonal matching pursuit, sliding window.

I. INTRODUCTION

MASSIVE machine-type communication (mMTC) is one of the three key application scenarios of 5G system [2]. In mMTC, there exist a massive number of potential users connecting to a base station (BS). However, there are often sporadic activities of these users, rendering the access scheme of traditional cellular networks inefficient or even infeasible [3]. In this regard, grant-free random access (GFRA) is recommended by 3GPP [4], which allows active users to transmit signatures and messages directly to BS without permission, thus ensuring low communication latency and high spectral efficiency [5]. However, as it is impossible to provide orthogonal pilots for all users in mMTC, it is a big challenge for BS to accomplish activity recognition and channel estimation with high accuracy in GFRA.

The standard multiuser information theoretic analysis for multi-access channels (MAC) is only applicable for a fixed and finite number of users and not suitable for massive connectivity systems. For this, to describe massive connectivity systems with random access from information theory, a novel channel model, named many-access channel (MnAC) has been proposed [6], where the number of users grows without bound, as the coding blocklength grows. MnAC has been extended to multiple-input multiple-output (MIMO) systems [7].

The episodic nature of active users allows the problem of recognizing active users to be formulated as a compressed sensing (CS) problem. In a series of works [8], [9], [10], CS algorithms have been applied to activity recognition, or joint activity recognition and message decoding, under the assumption that the receiver knows the exact channel state information (CSI) of all users. But this assumption is difficult to meet in practical scenarios with a huge number of potential users. A remedy is that BS performs joint activity recognition and channel estimation (JARCE); see, e.g., [11], [12], [13], [14], [15], [16], [17], [18]. In [11], [12], [13], Gaussian signature arrays were constructed and various approximate message passing (AMP) algorithms were developed to recover the CSI; in [14], Reed-Muller (RM) sequences were applied as signature arrays to reduce the required storage space; an orthogonal AMP algorithm was proposed in [15] for channels with spatial and temporal correlations; in [16], [17], generalized multiple measurement vector AMP and multi-rank aware method were respectively proposed for broadband systems; in [18], considering the estimation of large-scale fading coefficients only, covariance-based algorithms were utilized to alleviate the limitation on the number of active users due to the finite coherence time.

The joint approach of JARCE inevitably increases the overhead and complexity before message transmission. Furthermore, in some scenarios with high mobility, such as smart transportation, it is extremely difficult to acquire accurate CSI because of very short coherence time [19]. An alternative solution is to use noncoherent receivers. Early application of noncoherent detection was mainly motivated by simplified receiver circuits [20], and there have been recent applications for massive MIMO systems [21], [22]. Noncoherent receivers are attractive in mMTC scenarios since they have no requirement of CSI and possess low hardware complexity. Some studies [23], [24], [25] considered modulating messages to different signature arrays or power levels, and using AMP algorithms for joint activity recognition and message decoding in MIMO MnACs. However, those schemes need high cost for storing signal sequences [23], [24] and a large number of receiver antennas [25] increasing with the number of potential users.

Recently, another paradigm of grant-free access scheme named unsourced random access (URA) has been proposed in [26]. Unlike previous works, the URA problem formulation allows the BS to eliminate the requirement of activity recognition. Specifically, all users are assigned the same codebook, without signature identity. The URA has been extended to quasi-static fading channels [27] and block fading massive MIMO channels [28]. Some works studied coding schemes based on traditional channel coding [29], [30] or CS [31], [32], [33] for the URA.

Most existing GFRA schemes achieve their promised performance in mMTC scenarios at a cost of high complexity and idealistic modeling assumptions. Specifically, to mitigate the impact of channel fading, most works consider MIMO, e.g., [9], [11] or even massive MIMO, e.g., [12], [16], [25] systems to improve decoding performance via channel hardening effects. However, the increase in the number of antennas leads to a higher computational complexity of the CS algorithm in recovering the user state matrix. Noncoherent detection schemes do not require CSI, but still need massive antennas and sophisticated decoding algorithms to guarantee performance. Meanwhile, a sizable storage space may be required. The URA schemes remove the need of activity recognition, but still require sophisticated coding schemes to ensure performance. Furthermore, the above GFRA schemes all assume perfect inter-user synchronization (IUS). However asynchronism is a fundamental issue of multi-access [34], and it becomes extremely severe in mMTC due to the huge number of potential users. Phase uncertainty is an important aspect of asynchronism. In massive connectivity systems, it is difficult to know phase shifts at the users due to the delay and resource limits in feedback transmission. In particular, even without resource limits, in highly mobile scenarios, feedback delay may cause the phase information to be outdated when it reaches the users [35]. Meanwhile, time asynchronism is also an inherent challenge because users are geographically distributed. In multi-access systems, there have been some studies assuming time asynchronism [36], [37], but the existing GFRA schemes are still difficult to work without synchronous users. On the other hand, most existing GFRA schemes were proposed based on the frequency-flat fading channel model or even the block fading model. How to extend to the more practical frequency-selective fading channel has not been thoroughly treated.

A novel low-complexity coding scheme on OR MnACs has been proposed in [38]. The scheme considers a Bloom filter based coding and the standard Bloom filter verification procedure. Based on the proposed scheme, achievable bounds of the activity recognition task and the message transmission task have been derived. Note that an OR MnAC describes the input-output relationship in conditions where each user utilizes on-off signaling, and BS exploits envelope detection with negligible additive noise [39]. This observation provides us with inspirations for simplifying massive access, aiming at applying hard-decision envelope detection and Bloom filter based coding to reduce the complexity of transceiver hardware and of encoding and decoding algorithms. In [40], a similar scheme with Bernoulli distribution based coding and Noisy-COMP (NCOMP) based group testing (GT) has been proposed for activity recognition in noncoherent frequency-flat Rayleigh fading MnACs.

In this paper, our objective is to design a GFRA scheme with low hardware cost and computational complexity, applicable to various channel scenarios. We propose a scheme in Section II where each user utilizes Bloom filter based coding and OOK modulation, and BS utilizes hard-decision envelope detection, to treat both activity recognition and message transmission tasks in massive access. This scheme is built upon a noisy OR channel model which allows us to utilize the NCOMP decoding strategy. In Section III, we provide necessary technical tools for facilitating subsequent analysis. In Section IV, we derive the efficiency of NCOMP for frequency-flat fading MnACs with IUS and it shows that the scheme can guarantee performance without

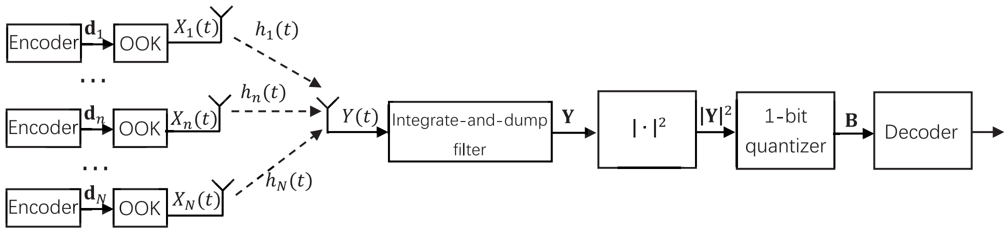


FIGURE 1. The transceiver structure.

relying on MIMO and any CSI.[1] In Section V, we propose a multipath-NCOMP (MNCOMP) decoding strategy to mitigate the effects of frequency-selective fading on performance of the proposed scheme, and we prove that it is better than NCOMP by asymptotic analysis. In Section VI, we propose a sliding window method to modify the above decoding strategies to adapt to asynchronous communication as well. Asymptotic analysis shows that the method incurs no loss in performance due to user asynchrony. In Section VII, we analyze the complexity of our scheme, and conduct numerical experiments to verify the feasibility of our scheme with a finite number of users.

Notation: In this paper, boldface letters denote vectors and calligraphic letters denote sets. We use $\text{Exp}(x)$ to denote a random variable that satisfies an exponential distribution with rate parameter x, $\Gamma(n, \beta)$ to denote a random variable that satisfies a gamma distribution with parameters n and β, $H_2(x)$ to denote the information entropy of Bernoulli random variable with parameter x, $\bar{\mathsf{E}}$ to denote the complementary event of E, $\mathbb{E}[\cdot]$ to denote the expectation operator, and $|\cdot|$ to denote the modulus of a complex variable, respectively.

II. SYSTEM MODEL

A. TRANSCEIVER STRUCTURE

Consider an uplink system consisting of N potential single-antenna users, denoted by the set $\mathcal{N} = \{1, \ldots, N\}$, and a BS equipped with one antenna. We assume that, in each communication block, each user accesses the channel with probability N_a/N, independently with others. Thus, the number of active users at the same time is a binomial random variable with mean N_a. Define $\mathbf{S} = [S_1, \ldots, S_N]$ as the activity indicator of N users. Specifically, for user $n \in \mathcal{N}$, if it is active, $S_n = 1$; otherwise, $S_n = 0$. We assume that users may not be synchronized, and denote the propagation time of user n as T_n. The channel input-output relationship is modeled as

$$Y(t) = \sum_{n=1}^{N} \sum_{\iota} S_n \alpha_{n,\iota}(t) X_n\big(t - T_n - T_{n,\iota}\big) + Z(t), \tag{1}$$

where $X_n(t)$ is the transmitted signal of user n, $\alpha_{n,\iota}(t)$ and $T_{n,\iota}$ are respectively the complex gain coefficient and the delay of the ι-th multipath component of user n, and $Z(t)$ is the complex additive white Gaussian noise (AWGN) with power spectral density (PSD) N_0. The delay spread of user n is $\tau_n = \max_{\iota} T_{n,\iota}$ and the maximum multipath delay is $T_{\max} = \max_{n}(T_n + \tau_n)$. The average power of each user satisfies the power constraint

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X_n^2(t)dt \leq \rho \qquad n \in \mathcal{N}. \tag{2}$$

The SNR is defined as $\mathsf{SNR} := \frac{\rho}{N_0 W_{\mathrm{N}}}$, where W_{N} is the receiver bandwidth.

As shown in Fig. 1, the proposed scheme utilizes OOK modulation at each user and hard-decision envelop detection at the BS to help reduce the detection and hardware complexity. In the following, we will introduce the transceiver structure in detail.

1) TRANSMITTER

Each of the N users is assigned a unique codebook containing J codewords, denoted as $\mathbf{C}_n = [\mathbf{c}_{n,1}, \ldots, \mathbf{c}_{n,J}]$, $n \in \mathcal{N}$, and each of the J codewords is a length-L binary array, denoted as $\mathbf{c}_{n,j} = [c_{n,j,1}, \ldots, c_{n,j,L}] \in \{0, 1\}^{1\times L}$, $j \in \{1, \ldots, J\}$. We assume that the codebooks of all users are known to the BS. During each transmission block, each active user transmits one codeword selected uniformly randomly from its codebook with OOK modulation, and the transmitted array of user n is denoted as $\mathbf{d}_n = [d_n[1], \ldots, d_n[L]]$. Thus, $\mathbf{d}_n = \sum_{j=1}^{J} \mathbb{1}_{\mathbb{J}_n}(j)\mathbf{c}_{n,j}$, where $\mathbb{1}_{\mathbb{J}_n}(j)$ is an indicator defined as

$$\mathbb{1}_{\mathbb{J}_n}(j) := \begin{cases} 1, & j = \mathbb{J}_n \\ 0, & j \neq \mathbb{J}_n, \end{cases} \tag{3}$$

and $\mathbb{J}_n$ is the message transmitted by user n, which is an independently and identically distributed (IID) random variable distributed over $\mathcal{J}_n = \{1, \ldots, J\}$. The baseband signal of user n can be represented as

$$X_n(t) = \sum_{l=1}^{L} d_n[l] p_{\mathrm{tr}}(t - (l-1)T_s), \tag{4}$$

where $p_{\mathrm{tr}}(t)$ is the rectangular pulse with amplitude A and pulse duration T_s. Note that, if user n is inactive, i.e., $S_n = 0$ in (1), it can be thought of as transmitting a length-L all-"0" codeword.

1. Multi-antenna BS can be introduced into our scheme to further improve performance, which further benefits from the channel hardening effect as the number of antennas gets large (see Section IV-D for details).

2) RECEIVER

The baseband received signal $Y(t)$ is fed into an integrate-and-dump filter with duration T_s per integration for filtering and sampling, yielding a length-$(L + L_\mathrm{d})$ array $\mathbf{Y} = [Y[1], \ldots, Y[L + L_\mathrm{d}]]$, where $L_\mathrm{d} = \lceil T_{\max}/T_s \rceil$ is the discretized value of the maximum multipath delay $T_{\max}$. An integrate-and-dump filter is the combination of an integrator and a sampler, which resets the integrator to 0 every T_s and its equivalent noise bandwidth $W_\mathrm{N} = 1/T_s$. Then, a one-bit quantizer with threshold $\eta \geq 0$ performs hard decision on the energy of samples $|\mathbf{Y}|^2$, yielding an output denoted as $\mathbf{B} = [B[1], \ldots, B[L+L_\mathrm{d}]]$, which is a binary array satisfying $B[\ell] = \mathrm{sgn}(|Y[\ell]|^2)$, $1 \leq \ell \leq L + L_\mathrm{d}$, where

$$\mathrm{sgn}(y) := \begin{cases} 1, & y > \eta \\ 0, & y \leq \eta. \end{cases} \quad (5)$$

The hard-decision threshold η needs to be optimally chosen based on appropriate performance metrics.[2] The decoder decides which users are active and which messages they transmit with high probability according to $\mathbf{B}$.

B. NOISY OR CHANNEL

In this subsection, we derive a noisy OR channel from the proposed transceiver. The noisy OR channel is an OR channel followed by a binary asymmetric channel (BAC). Specifically, the OR channel is a memoryless noiseless MAC as

$$\mathsf{Y} = \mathsf{X}_1 \vee \mathsf{X}_2 \vee \cdots \vee \mathsf{X}_N, \quad \mathsf{X}_n, \mathsf{Y} \in \{0, 1\}. \quad (6)$$

That is to say, the output is "0" only if all the inputs are "0"'s; otherwise the output is "1". BAC is a memoryless channel with a binary input and a binary output. The probability of the channel outputting "1" under input "0" is generally different from that of the channel outputting "0" under input "1".

At the receiver, the output of the integrate-and-dump filter is

$$\begin{aligned} Y[\ell] &= \int_{(\ell-1)T_s}^{\ell T_s} Y(t)dt \\ &= \int_{(\ell-1)T_s}^{\ell T_s} \sum_{n=1}^{N} \sum_{\iota} S_n \alpha_{n,\iota}(t) \sum_{l=1}^{L} d_n[l] \\ &\quad p_{\mathrm{tr}}\big(t - (l-1)T_\mathrm{s} - T_n - T_{n,\iota}\big)dt \\ &\quad + \int_{(\ell-1)T_s}^{\ell T_s} Z(t)dt \\ &= X[\ell] + Z[\ell], \quad \ell = 1, \ldots, L + L_\mathrm{d}, \end{aligned} \quad (7)$$

where $Z[\ell] \sim \mathcal{CN}(0, N_0 T_s)$ is IID [42].[3] The tap number of the integrate-and-dump filter of user n is $\lceil \tau_n/T_s \rceil$. We assume user n has M_n non-zero taps. Thus, $X[\ell]$ can be written as the standard uniform tapped delay line model[4]

$$X[\ell] = \sum_{n=1}^{N} \sum_{m=1}^{M_n} S_n h_{n,m}[\ell] d_n\big[\ell - \zeta_n - \ell_{n,m}\big], \quad (8)$$

where $h_{n,m}[\ell]$ is the gain coefficient of the m-th non-zero tap of user n on the ℓ-th sample

$$\begin{aligned} h_{n,m}[\ell] = &\int_{(\ell-1)T_s}^{\ell T_s} \sum_{\iota} \alpha_{n,\iota}(t) \\ &p_{\mathrm{tr}}\big(t - \big(\ell - \zeta_n - \ell_{n,m} - 1\big)T_\mathrm{s} - T_n - T_{n,\iota}\big)dt, \end{aligned} \quad (9)$$

where $\zeta_n = \lceil T_n/T_s \rceil$ is the discretized propagation time of user n and $\ell_{n,m}$ is the discrete delay of the m-th non-zero taps of user n. Define the tap-delay vector of user n as $\mathbf{l}_n = [\ell_{n,1}, \ldots, \ell_{n,M_n}]$. Note that $\ell_{n,1} = 0$. For the sake of simplicity, we assume that average received powers of all non-zero taps of all users are the same, and $h_{n,m}[\ell]$ is IID satisfying $h_{n,m}[\ell] \sim \mathcal{CN}(0, A^2 T_s^2)$. Therefore, the output of the one-bit quantizer is

$$\begin{aligned} B[\ell] &= \mathrm{sgn}\Big(|Y[\ell]|^2\Big) \\ &= \mathrm{sgn}\left(\left|\sum_{n=1}^{N} \sum_{m=1}^{M_n} S_n h_{n,m}[\ell] d_n\big[\ell - \zeta_n - \ell_{n,m}\big] + Z[\ell]\right|^2\right). \end{aligned} \quad (10)$$

The existence of AWGN may cause errors in hard decision of the one-bit quantizer. Specifically, conditioned on the event that $S_n d_n[\ell - \zeta_n - \ell_{n,m}] = 0$ for all $m \in \{1, \ldots, M_n\}$, $n \in \mathcal{N}$, the hard decision of the energy of ℓ-th sample $|Y[\ell]|^2$ should be $B[\ell] = 0$. But due to the existence of AWGN, $|Y[\ell]|^2$ may exceed the threshold η, and thus the hard decision $B[\ell] = 1$. This error event is called zero-to-one flip and its probability is[5]

$$p_o = \Pr\Big(|Z[\ell]|^2 > \eta\Big) = \Pr(\mathrm{Exp}(1) > \eta) = e^{-\eta}. \quad (11)$$

Similarly, conditioned on the event that at least one $S_n d_n[\ell - \zeta_n - \ell_{n,m}] = 1$ for $m \in \{1, \ldots, M_n\}$, $n \in \mathcal{N}$, the hard decision of $|Y[\ell]|^2$ should be $B[\ell] = 1$. But, since $|Y[\ell]|^2$ may not exceed the threshold η, the hard decision $B[\ell] = 0$. This error event is called one-to-zero flip and its probability is

$$\begin{aligned} p_v &\leq \Pr\Big(|AT_s h_n[\ell] + Z[\ell]|^2 \leq \eta\Big) \\ &= \Pr\Big(\mathrm{Exp}\Big(A^2 T_s/N_0 + 1\Big) \leq \eta\Big) \\ &= 1 - e^{-\frac{\eta N_0}{A^2 T_s + N_0}}. \end{aligned} \quad (12)$$

Thus, (10) can be equivalent to a noisy OR channel, which is an OR channel with inputs $S_n d_n[\ell - \zeta_n - \ell_{n,m}] \in \{0, 1\}$,

2. In our paper, we consider the hard-decision threshold η optimization problem in the many-access regime (see remarks of each proposition for details), and the optimal and sub-optimal η for finite N can be obtained by an exhaustive search and a heuristic method respectively (for details, see Section VII-C).

3. It should be noted that if we utilize an envelope detector to obtain the envelope of $Y(t)$, e.g., [21], we cannot obtain such an IID Gaussian noise term.

4. We assume symbol synchronism between users, and for the scenario without symbol synchronization, it can be classified as the multi-tap case (see Section V).

5. We normalize the energy of all samples by $N_0 T_s$.

$m = 1, \ldots, M_n$, $n = 1, \ldots, N$, and an output

$$B^{[n]}[\ell] = \operatorname{sgn}\left(\sum_{n=1}^{N}\sum_{m=1}^{M_n} S_n d_n\big[\ell - \zeta_n - \ell_{n,m}\big]\right) \in \{0,1\}, \quad (13)$$

followed by a BAC with an input $B^{[n]}[\ell]$, an output $B(\ell) \in \{0, 1\}$ and transition probabilities p_v and p_o.

C. TASKS OF MASSIVE ACCESS SYSTEM

We consider activity recognition task and message transmission task, which are defined as follows.

Definition 1 (Activity Recognition Task): Activity recognition task is to utilize a coding scheme for the transceiver to recognize active users with a sufficiently low error probability. In the coding scheme, the codebook of each user has only one codeword, i.e., its signature array, whose length is denoted as L_0 instead of L in order to distinguish it from the coding blocklength in the message transmission task. The coding scheme consists of the following mappings:

- Encoding functions $\mathcal{E}_n(S_n)$: $\{0, 1\} \to \{\mathbf{0}, \mathbf{a}_n^{\mathrm{s}}\}$ for each user $n \in \mathcal{N}$, where $\mathbf{0}$ is a length-L_0 all-"0" array and $\mathbf{a}_n^{\mathrm{s}} = [a_{n,1}^{\mathrm{s}}, \ldots, a_{n,L_0}^{\mathrm{s}}] \in \{0, 1\}^{1\times L_0}$ is the signature array of user n.
- Decoding function $\mathcal{D}(\mathbf{B})$: $\{0, 1\}^{1\times(L_0+L_{\mathrm{d}})} \to \{0, 1\}^{1\times N}$, which is a deterministic rule, mapping each possible received vector $\mathbf{B}$ to a decision of activity states $\hat{\mathbf{S}} = [\hat{S}_1, \ldots, \hat{S}_N]$.

The probability of erroneous recognition of the coding scheme is $\Pr[\mathsf{E}^{[\mathrm{r}]}] = \Pr\{\hat{\mathbf{S}} \neq \mathbf{S}\}$.

Definition 2 (Message Transmission Task): Message transmission task is to utilize a coding scheme for the transceiver to recognize active users and transmit their messages with a sufficiently low error probability. The coding scheme consists of the following mappings:

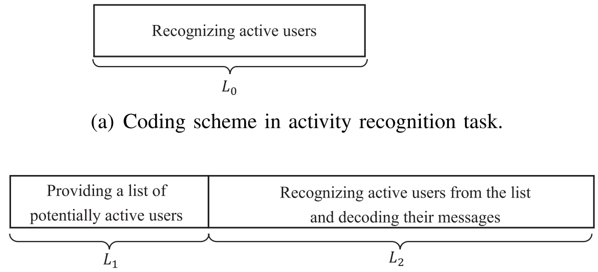
- Encoding functions $\mathcal{E}_n(\tilde{\mathfrak{J}}_n)$: $\{0, 1, \ldots, J\} \to \{\mathbf{0}, \mathbf{c}_{n,1}, \ldots, \mathbf{c}_{n,J}\}$ for every user $n \in \mathcal{N}$, where $\tilde{\mathfrak{J}}_n = S_n\mathfrak{J}_n$ and $\mathbf{0}$ is a length-L all-"0" array. Specifically, $\mathcal{E}_n(0) = \mathbf{0}$ and $\mathcal{E}_n(j) = \mathbf{c}_{n,j}$, $\forall j = 1, \ldots, J$.
- Decoding function $\mathcal{D}(\mathbf{B})$: $\{0, 1\}^{1\times(L+L_{\mathrm{d}})} \to \{0, 1, \ldots, J\}^{1\times N}$, which is a deterministic rule, mapping each possible received vector $\mathbf{B}$ to a decision of message states $\hat{\mathbf{J}} = [\hat{\mathfrak{J}}_1, \ldots, \hat{\mathfrak{J}}_N]$.

Let $\tilde{\mathbf{J}}_n = [\tilde{\mathfrak{J}}_1, \ldots, \tilde{\mathfrak{J}}_N]$. The probability of erroneous decoding of the coding scheme is $\Pr[\mathsf{E}^{[\mathrm{m}]}] = \Pr\{\hat{\mathbf{J}} \neq \tilde{\mathbf{J}}\}$.

To facilitate subsequent analysis, this paper considers the many-access regime where the number of users N grows without bound. We dedicate to a scenario satisfying the following conditions. The average number of active users, N_a, satisfies $N_a = \Theta(N^{\beta})$ for some $0 < \beta < 1$. Note that N_a grows without bound, while the activity ratio N_a/N asymptotically vanishes, with N. The number of messages per user, J, satisfies $J = \Theta(N^{\gamma})$ for some $\gamma > 0$. In this scenario, the total information that needs to be transmitted in activity recognition task and message transmission task are $NH_2(N_a/N) = (1 - \beta)N_a[\log_2 N + O(1)]$ and $NH_2(N_a/N) + N_a \log_2 J = (1 - \beta + \gamma)N_a[\log_2 N + O(1)]$, respectively. We follow the cost definitions in [38] to characterize the efficiency of activity recognition and message transmission. The two costs are defined as follows.

Definition 3 (Feasible Activity Recognition Cost): An activity recognition cost Ω_a is called feasible, if there exists a sequence of length-$(\Omega_a N_a \log_2 N)$ signature arrays such that, the probability of erroneous recognition $\Pr[\mathsf{E}^{[\mathrm{r}]}]$ vanishes, as N grows without bound.

Definition 4 (Feasible Message Transmission Cost): A message transmission cost Ω_m is called feasible, if there exists a sequence of codebooks with length-$(\Omega_m N_a \log_2 N)$ codewords such that, the probability of erroneous decoding $\Pr[\mathsf{E}^{[\mathrm{m}]}]$ vanishes, as N grows without bound.



(a) Coding scheme in activity recognition task.

(b) Coding scheme in message transmission task.

FIGURE 2. Coding scheme.

D. CODING SCHEME

Our scheme utilizes Bloom filter as signature arrays and codewords. The Bloom filter can be defined as follows [41].

Definition 5 (Bloom Filter): A Bloom filter of parameters (L, K), denoted as $\mathsf{BF}(L, K)$, is a length-L array, generated independently by the following algorithm:

- Create an all-"0" length-L array;
- Take K independent hash functions, and each hash function uniformly selects one among L positions to set it to "1".

Note that when a position has already been set as "1", it will remain to be "1" if it is selected later by another hash function.

For activity recognition task, the Bloom filter based coding scheme is as shown in Fig. 2(a). Each of the N users is assigned a length-L_0 Bloom filter of parameters (L_0, K_0) as its unique signature array $\mathbf{a}_n^{\mathrm{s}}$. Active users transmit their non-orthogonal length-L_0 signature arrays to the BS, and if the decoder decides user n sends its signature array, it sets $\hat{S}_n = 1$; otherwise, it sets $\hat{S}_n = 0$.

For message transmission task, the Bloom filter based coding scheme adopts the partial activity recognition coding proposed in [38] consisting of two phases, as shown in Fig. 2(b).[6] Specifically, each of the N users is assigned a

6. The partial activity recognition coding is proven to be more beneficial in our scheme compared with the two-phase scheme with complete separation [6] (see Section V).

length-L_1 Bloom filter of parameters ($L_1 = \Omega_1 N_a \log_2 N$, K_1) as its unique signature array, denoted as $\mathbf{a}_n^{\mathrm{m}} = [a_{n,1}^{\mathrm{m}}, \ldots, a_{n,L_1}^{\mathrm{m}}] \in \{0,1\}^{1\times L_1}$, and is assigned J length-L_2 Bloom filters of parameters ($L_2 = \Omega_2 N_a \log_2 N$, K_2) as its J message arrays, denoted as $\mathbf{C}_n^{\mathrm{m}} = [\mathbf{c}_{n,1}^{\mathrm{m}}, \ldots, \mathbf{c}_{n,J}^{\mathrm{m}}]$, where $\mathbf{c}_{n,j}^{\mathrm{m}} = [c_{n,j,1}^{\mathrm{m}}, \ldots, c_{n,j,L_2}^{\mathrm{m}}] \in \{0,1\}^{1\times L_2}$, $j = 1, \ldots, J$. The codebook of user n, $\mathbf{C}_n$, consists of $\mathbf{a}_n^{\mathrm{m}}$ and $\mathbf{C}_n^{\mathrm{m}}$, where each of the length-L ($L = L_1 + L_2$) codewords $\mathbf{c}_{n,j} = [\mathbf{a}_n^{\mathrm{m}}, \mathbf{c}_{n,j}^{\mathrm{m}}] \in \{0,1\}^{1\times L}$, $j = 1, \ldots, J$. In Phase 1, active users transmit their non-orthogonal length-L_1 signature arrays to the BS, and the decoder provides a list of potentially active users by deciding whether signature arrays of users are sent, and sets $\hat{\mathsf{J}}_n = 0$ for each user n not contained in the list. In Phase 2, active users transmit their non-orthogonal length-L_2 messages to the BS, and the decoder resolves the ambiguity of activity recognition by decoding messages. Specifically, for each user n in the list, if the decoder decides user n only sends its j-th message, it sets $\hat{\mathsf{J}}_n = j$; otherwise, it sets $\hat{\mathsf{J}}_n = 0$.

The noisy OR channel model derived from the proposed transceiver structure allows the decoder to utilize the NCOMP decoding strategy to decide which arrays are sent in frequency-flat fading channels (see Section IV-B for detail). In frequency-selective fading channels, the output of the OR channel $\mathbf{B}^{[n]}$ is a superposition of all delayed replicas of Bloom filters transmitted by all active users.[7] We call the superposition of multiple delayed replicas of a Bloom filter a generalized Bloom filter. The improved decoding strategy for frequency-selective fading channels, MNCOMP, will be proposed in Section V-B.

The parameter K_i, $i = 0, 1, 2$, of the Bloom filters is set as $K_i = L_i\kappa_i/N_a$, and $\kappa_i > 0$ should be carefully chosen under different decoding strategies to optimize the performance of the scheme. Meanwhile, to satisfy (2), the amplitude of the rectangular pulse, A, is constrained as

$$A^2\frac{K_i}{L_i} = \rho. \tag{14}$$

In the many-access regime, note that

$$\frac{L_i}{K_i} = \frac{N_a}{\kappa_i} = \frac{N^\beta}{\kappa_i} \tag{15}$$

vanishes asymptotically with N, and thus, A grows without bound with N.

Table 1 summarizes the key parameters used in this paper.

III. PRELIMINARY LEMMAS FOR ANALYSIS

Definition 6: A generalized Bloom filter, denoted as $\mathsf{GBF}(L + L_{\mathrm{d}}, K_g, \zeta_g, \mathbf{l}_g : g = 1, \ldots, G)$, can be obtained by superposing all delayed replicas of G length-L Bloom filters, $\mathsf{BF}(L, K_g)$, $g = 1, \ldots, G$, with the discrete propagation time ζ_g and the tap-delay vector $\mathbf{l}_g = [\ell_{g,1}, \ldots, \ell_{g,M_g}]$. It can also be constructed by the following rule:

- First create a length-($L + L_{\mathrm{d}}$) all-“0” array, and set index $g = 1$.

7. Here “superposition” is OR operations.

TABLE 1. A summary of key parameters.

Notation	Parameter Description
N	Total number of users
S_n	Activity indicator of user n
N_a	Mean of the number of active users
β	Parameter related to N_a: $N_a = \Theta(N^\beta)$
J	Number of messages
γ	Parameter related to J: $J = \Theta(N^\gamma)$
M_n	Number of non-zero taps of user n
$\mathbf{l}_n$	Tap-delay vector of user n
ζ_n	Discrete propagation time of user n
L	Coding blocklength
K	Number of hash functions for Bloom filters
κ	The adjustment parameter of K
η	Quantization threshold

- Take K_g independent hash functions. Each hash function uniformly selects a position ℓ among first L positions and sets $(\ell+\zeta_g+\ell_{g,1})$-th,..., $(\ell+\zeta_g+\ell_{g,M_g})$-th positions to “1”.
- Set $g = g + 1$, and repeat the 2-nd step until $g > G$.

Note that, the Bloom filter is a special case of the generalized Bloom filter.

We next present properties of generalized Bloom filters to facilitate subsequent analysis. These properties can be generalized from properties of Bloom filters in [38].

Lemma 1 (Superposition Property): Consider superposing two generalized Bloom filters, $\mathsf{GBF}(L, K_1, \zeta_1, \mathbf{l}_1)$ and $\mathsf{GBF}(L, K_2, \zeta_2, \mathbf{l}_2)$. If $\zeta_1 = \zeta_2$ and $\mathbf{l}_1 = \mathbf{l}_2$, the resulting array will be a generalized Bloom filter of parameters $(L, K_1 + K_2, \zeta_1, \mathbf{l}_1)$; otherwise, it will be a generalized Bloom filter of parameters $(L, K_g, \zeta_g, \mathbf{l}_g : g = 1, 2)$.

Lemma 2 (Occupancy Concentration Property): Let $\mathbf{B}^{[n]} = \mathsf{GBF}(L + L_{\mathrm{d}}, K_g, \zeta_g, \mathbf{l}_g : g = 1, \ldots, G)$. Define weight W of $\mathbf{B}^{[n]}$ as the Hamming weight of the first L elements of $\mathbf{B}^{[n]}$ and $\mathbf{l}_g$-weight W_g of $\mathbf{B}^{[n]}$ as the number of ℓ's within $\{1, \ldots, L\}$ which satisfy $B^{[n]}[\ell + \ell_{g,m}] = 1$ for all $m = 1, \ldots, M_g$. For any $\epsilon > 0$, parameter $\mathsf{Z}_g = L - W_g$ of $\mathbf{B}^{[n]}$ satisfies

$$\Pr\left[\left|\mathsf{Z}_g - \mathbb{E}[\mathsf{Z}_g]\right| \le \epsilon L\right] < 2\exp\left(-\frac{\epsilon^2 L^2}{2M_{\max}^2 K}\right), \tag{16}$$

where $K = \sum_{g=1}^{G} K_g$ is the number of hash functions of $\mathbf{B}^{[n]}$, and $M_{\max} = \max_g M_g$.

Lemma 3 (Incorrect Flipping Concentration Property): Define false-number O of $\mathbf{B}$ as the number of zero-to-one flips in $\mathbf{B}$ and $\mathbf{l}_g$-false-number O_g of $\mathbf{B}$ as the number of ℓ's within $\{1, \ldots, L\}$ which satisfy that at least one $B^{[n]}[\ell + \ell_{g,m}]$, $m = 1, \ldots, M_g$, experiences zero-to-one flip. Define miss-number V of $\mathbf{B}$ as the number of one-to-zero flips in $\mathbf{B}$ and $\mathbf{l}_g$-miss-number V_g of $\mathbf{B}$ as the number of ℓ's within $\{1, \ldots, L\}$ which satisfy that at least one $B^{[n]}[\ell + \ell_{g,m}]$, $m = 1, \ldots, M_g$, experiences one-to-zero flip.

The $\mathbf{l}_g$-false-number O_g of $\mathbf{B}$ satisfies for any $\epsilon_1 > 0$,

$$\Pr\left[\left|O_g - \mathbb{E}[O_g]\right| \le \epsilon_1 L \middle| W_g = w_g\right] > 1 - 2\exp\left(-\frac{2\epsilon_1^2 L^2}{{M_g}^2 (L - w_g)}\right), \tag{17}$$

and the $\mathbf{l}_g$-miss-number V_g of $\mathbf{B}$ satisfies for any $\epsilon_2 > 0$,

$$\Pr\left[\left|V_g - \mathbb{E}[V_g]\right| \le \epsilon_2 L \middle| W_g = w_g\right] > 1 - 2\exp\left(-\frac{2\epsilon_2^2 L^2}{M_g^3 w_g}\right). \tag{18}$$

Proof: Note that $W_g \le W \le M_g W_g$, $O/M_g \le O_g \le M_g O$, and $V/M_g \le V_g \le M_g V$. We can obtain (17) and (18) following from Hoeffding's inequality. ■

IV. FREQUENCY-FLAT FADING CHANNELS WITH IUS

In this section, we analyze the performance of the proposed scheme with NCOMP in activity recognition task and message transmission task on frequency-flat fading MnACs with IUS.

A. REVIEW OF OR MANY-ACCESS CHANNELS

In [38], bounds of minimum feasible activity recognition cost and minimum feasible message transmission cost on OR MnACs are respectively derived as

$$1 - \beta \le \Omega_a \le 1/\ln 2, \quad 1 - \beta + \gamma \le \Omega_m \le (1+\gamma)/\ln 2. \tag{19}$$

The above upper bounds are achieved by utilizing Bloom filter based coding and classic Bloom filter verification procedure.[8] The above lower bounds are respectively achieved by allowing all users to fully cooperate to send a codeword informing the receiver about their activity states in activity recognition task and messages of active users in message transmission task.[9] We remark that the lower bounds are loose, because of the assumption of full cooperation. In the rest of this paper, the lower bounds are called full cooperation lower bounds.

We remark that the above bounds are the performance of our scheme in the absence of noise. In the following, we turn to the noisy case.

B. NCOMP DECODING STRATEGY

Note that, in frequency-flat fading channels with IUS, i.e., $M_n = 1$, $\zeta_n = 0$, $\forall n \in \mathcal{N}$, the OR channel output $\mathbf{B}^{[n]}$ of the activity recognition or each phase of the message transmission is the superposition of Bloom filters transmitted by all active users. Due to the BAC model, the hard-decision output $\mathbf{B}$ can be obtained by performing random flips on $\mathbf{B}^{[n]}$. Recalling the classic Bloom filter verification procedure [41]: Verify whether an item has been superposed in $\mathbf{B}^{[n]}$, by checking whether the Bloom filter of this item is contained in $\mathbf{B}^{[n]}$ (i.e., $\mathbf{B}^{[n]}$ containing all "1"s of the Bloom filter), we can provide a decoding strategy that can be utilized in noisy OR channels as follows: For the activity recognition or each phase of the message transmission, the decoder decides that an array $\mathbf{a}$ is sent when there are at most λ "1"s in the array not contained in $\mathbf{B}$, where λ is the decoding threshold. Specifically, if $\mathbf{a}$ satisfies

$$\sum_{w=1}^{W} B[l_w] \ge W - \lambda, \tag{20}$$

the decoder decides the array is sent, where W is the Hamming weight of the array and l_w is the position of the w-th "1" of the array. We note that the above decoding strategy is called NCOMP decoding strategy in GT problems [43]. We remark that, our scheme also provides a noisy non-adaptive GT protocol.

C. FEASIBLE COSTS OF PROPOSED SCHEME

In the frequency-flat fading MnACs with IUS scenario, the NCOMP decoding strategy is used in each phase of tasks to recognize active users or decode messages. Let us start with the activity recognition task. In our scheme, activity recognition has two types of error events: 1) At least one active user is recognized as an inactive user, E^{MR}; 2) At least one inactive user is falsely recognized as an active user, E^{FR}. We have the following result on the activity recognition task.

Proposition 1: With the NCOMP decoding strategy, the proposed scheme achieves a feasible activity recognition cost $\Omega_a = 1/\ln 2$.

Proof: See Appendix A. ■

From the proof of Prop. 1, we remark that the decoding threshold λ should satisfy $\lambda > 0$ and $\lambda = o(\log_2 N)$, the hard-decision threshold η should satisfy $\eta = o(N^{\frac{\beta\lambda}{\lambda+1}}/\log_2 N)$ and $\lim_{N\to\infty} e^{-\eta} \to 0$, and κ_0 should be chosen as $\kappa_0 = \ln 2$ to minimize the feasible activity recognition cost. This implies that the decoding threshold λ can be set to a modest constant and the hard-decision threshold η should increase with N. The minimum cost $1/\ln 2$ is consistent with the result in [38], which is equivalent to the minimum cost achieved by our scheme when noise is absent. According to Def. 3, the coding blocklength corresponding to $\Omega_a = 1/\ln 2$ is $L_0 = N_a \log_2 N/\ln 2$. It has the same asymptotic growth rate $N_a \log_2 N$ as the full cooperation lower bound $(1-\beta)N_a \log_2 N$ in [38]. Especially, when the average number of active users N_a grows slowly with N, i.e., β is close to zero, the increase is only 45% compared with the full cooperation lower bound.

In [6], the minimum coding blocklength required for activity recognition task in Gaussian MnAC is given as

$$L_0 = \frac{(1-\beta)N_a \log N}{\log(1 + N_a \rho)}. \tag{21}$$

8. We note that the procedure is called COMP decoding strategy in GT problems.

9. The required coding blocklengths of the scheme in activity recognition task and message transmission task are $NH_2(N_a/N)$ and $NH_2(N_a/N) + N_a \log_2 J$, respectively.

Compared with the result of Prop. 1, we see that both our result and (21) have the same asymptotic growth rate $N_a \log N$. The main difference between our result and (21) is that the denominator of (21) is the sum capacity of an ideal Gaussian multiple access channel. In order to achieve (21), we need soft-decision detection at the receiver, and further need all users be perfectly synchronized in time, carrier frequency, and phase. In addition, a lower bound of the minimum coding blocklength in non-coherent Rayleigh-fading many-access channel is given as

$$L_0 \geq \frac{(1-\beta)N_a \log N}{N_a C_{\text{non}}(\rho)}, \tag{22}$$

where $C_{\text{non}}(\rho)$ is the capacity of a single-user non-coherent Rayleigh fading channel. We see that, similar to (21), the lower bound in (22) is still the same asymptotic growth rate $N_a \log N$ divided by the sum capacity bound.

Then we consider the message transmission task. There are two types of error events: 1) At least one active user is missed, i.e., recognized as inactive or its messages are incorrectly decoded, E^{ME}; 2) At least one inactive user is falsely recognized as active, denoted as E^{FR}. Note that the error event E^{ME} consists of the following three events: 1) At least one active user is missed, i.e., recognized as inactive in Phase 1, E^{MR}; 2) At least one active user's transmitted message is missed, i.e., not decoded in Phase 2, E^{MD}; 3) At least one active user's non-transmitted messages are falsely decoded in Phase 2, E^{FD}.

We have the following result on message transmission task.

Proposition 2: With the NCOMP decoding strategy, the proposed scheme achieves a feasible message transmission cost $\Omega_m = (1+\gamma)/\ln 2$, where the cost in Phase 1 is $\Omega_1 = (1-\beta)/\ln 2$ and the cost in Phase 2 is $\Omega_2 = (\beta+\gamma)/\ln 2$.

Proof: See Appendix B. ■

From the proof of Prop. 2, we remark that the decoding threshold of Phase i, λ_i $(i = 1, 2)$, and the hard-decision threshold η should meet certain conditions similar to those in Prop. 1 (see Appendix B for details), and κ_i $(i = 1, 2)$ should be chosen as $\kappa_i = \ln 2$ to minimize the feasible message transmission cost. The minimum cost is also consistent with the result in [38], which is equivalent to the minimum cost achieved by our scheme when noise is absent. In contrast, it can be proved that the cost of a separated two-phase coding scheme is $\Omega_m > (1+\beta+\gamma)/\ln 2$, which is strictly larger than that of Prop. 2. We also remark that Prop. 1 and Prop. 2 indicate that our scheme can guarantee performance without relying on multiple antennas.

We can also evaluate the efficiency of message transmission in terms of the asymptotically achievable message length defined in [6], which can be derived from message transmission costs. From the full cooperation lower bound in [38], upper bound of the asymptotically achievable message length of the proposed scheme is derived as

$$\log_2 J = L/N_a - (1-\beta)\log_2 N. \tag{23}$$

A heuristic explanation of (23) is as follows: When a genie accurately informs the BS of which users are active, the total number of bits that can be transmitted would be approximately L, and hence the asymptotically achievable message length is L/N_a. The total uncertainty in the activity of all N users is $H_2(N_a/N) = (1-\beta)N_a \log_2 N$, and thus the message length penalty on each of the N_a active users is $(1-\beta)\log_2 N$. From Prop. 2, the asymptotically achievable message length of the proposed scheme with NCOMP decoding strategy is derived as

$$\log_2 J = \frac{L \ln 2}{N_a} - \log_2 N. \tag{24}$$

We also have a similar heuristic explanation of (24) as follows: The asymptotically achievable message length with a genie is LL_2/N_a, which is 31% smaller than that of (23). The penalty on each active user is $\log_2 N$ which corresponds to the result in Prop. 1. Note that, the asymptotically achievable sum rate of the proposed scheme with NCOMP decoding strategy is

$$R_{\text{sum}} = \frac{N_a \log_2 J}{L} = \ln 2 - \frac{\ln 2}{1+\gamma}. \tag{25}$$

In the limit of large γ, $R_{\text{sum}} = \ln 2$, which is consistent with the sum capacity of an OR MAC without joint decoding.

D. PROPOSED SCHEME WITH MULTI-ANTENNA BS

The proposed scheme can be extended to the case of multi-antenna BS, by modifying the input of one-bit quantizer to the average energy of samples from all antennas, to further improve performance. In the multi-antenna case, assuming that the BS is equipped with N_r antennas, the output of the one-bit quantization is

$$B[\ell] = \text{sgn}\left(\frac{1}{N_r}\sum_{n_r=1}^{N_r}\left|\sum_{n=1}^{N} S_n h_{n,n_r}[\ell] d_n[\ell] + Z_{n_r}[\ell]\right|^2\right), \tag{26}$$

where h_{n,n_r} and $Z_{n_r}[\ell]$ are the gain coefficient of user n and AWGN on the n_r-th antenna of the BS, respectively. The probability of zero-to-one flip is

$$\begin{aligned} p_o^{[\text{mul}]} &= \Pr\left(\frac{1}{N_r}\sum_{n_r=1}^{N_r}\left|Z_{n_r}[\ell]\right|^2 > \eta\right) \\ &= \Pr(\Gamma(N_r, 1/N_r) > \eta), \end{aligned} \tag{27}$$

and the probability of one-to-zero flip is

$$\begin{aligned} p_v^{[\text{mul}]} &\leq \Pr\left(\frac{1}{N_r}\sum_{n_r=1}^{N_r}\left|AT_s h_{n,n_r}[\ell] + Z_{n_r}[\ell]\right|^2 \leq \eta\right) \\ &= \Pr\Big(\Gamma\Big(N_r, \Big(A^2 T_s/N_0 + 1\Big)/N_r\Big) \leq \eta\Big). \end{aligned} \tag{28}$$

We note that as N_r grows without bound, the noisy OR channel model (24) asymptotically becomes an OR channel. So in this asymptotic regime, the feasible activity recognition cost Ω_a and the feasible message transmission cost

Ω_m in Prop. 1 and Prop. 2 still hold. Therefore, the asymptotic performance with multi-antenna BS is identical to that of the single-antenna BS case. For finite number of users, especially in the low SNR regime, however, multiple antennas still effectively improve the performance (see Fig. 8 in Section VII-C). It should be noted that the extension of the scheme with multi-antenna BS does not incur extra computational complexity, since the input to the one-bit quantizer is simply the aggregated energy from receive antennas and the subsequent processing is exactly identical to that for the single-antenna BS case.

The extension can also be applied directly in schemes on frequency-selective fading channels and scenarios without IUS presented later, which will not be repeated in the subsequent sections.

V. FREQUENCY-SELECTIVE FADING CHANNELS WITH IUS

In this section we proceed to the frequency-selective fading channels with IUS. We assume that the channel model has G different tap-delay vectors, denoted as $\mathcal{G} = \{\mathbf{l}_1, \ldots, \mathbf{l}_G\}$, $\mathbf{l}_g = [\ell_{g,1}, \ldots, \ell_{g,M_g}]$, where the maximum and minimum of M_g are denoted respectively as $M_{\max}$ and $M_{\min}$. Each user's tap-delay vector is $\mathbf{l}_g$ with probability α_g. Thus, the average number of non-zero taps of users is written as

$$\mathbb{E}[M] = \sum_{g=1}^{G} \alpha_g M_g. \tag{29}$$

For the frequency-selective fading MnACs, they can be divided into two categories: (1) We call dense multipath, if $\mathbf{l}_g = [1, \ldots, M_g]$, $g = 1, \ldots, G$; (2) We call sparse multipath, otherwise.

A. PERFORMANCE OF NCOMP

Using the proposed scheme with NCOMP decoding strategy, we have the following result.

Proposition 3: With the NCOMP decoding strategy, the proposed scheme achieves a feasible activity recognition cost

$$\Omega_a = \frac{\mathbb{E}[M]}{\ln 2}, \tag{30}$$

and a feasible message transmission cost

$$\Omega_m = \frac{(1+\gamma)\mathbb{E}[M]}{\ln 2}. \tag{31}$$

Proof: See Appendix C. ∎

From the proof of Prop. 3, we remark that the hard-decision threshold and decoding threshold should satisfy the same conditions as those in the remarks of Prop. 1 and Prop. 2 and $\kappa_i = \ln 2/\mathbb{E}[M]$ $(i = 0, 1, 2)$, to minimize the feasible activity recognition cost and the feasible message transmission cost (see Appendix C for details). The above minimum costs Ω_a and Ω_m increase by a factor of $\mathbb{E}[M]$ times, compared with the results in frequency-flat fading MnACs. The performance loss is due to the inter-user interference induced by multi-taps and is independent of categories of frequency-selective fading. From Ω_m in Prop. 3, the asymptotically achievable message length is derived as

$$\log_2 J = \frac{L\ln 2}{N_a\mathbb{E}[M]} - \log_2 N, \tag{32}$$

and the asymptotically achievable sum rate is derived as

$$R_{\text{sum}} = \frac{\ln 2}{\mathbb{E}[M]} - \frac{\ln 2}{(1+\gamma)\mathbb{E}[M]} \tag{33}$$

B. IMPROVED DECODING STRATEGY

Then we consider proposing an improved decoding strategy to reduce the performance loss. Note that, the feasible costs achieved by our proposed scheme with NCOMP decoding strategy are limited by the false recognition and false decoding probabilities, according to the results of Prop. 3. Thus, if the receiver knows the tap-delay vector of each user, we can utilize all non-zero taps in decoding to reduce the false recognition and false decoding probabilities. Based on the heuristic idea, we propose the following MNCOMP decoding strategy: The decoder decides that an array $\mathbf{a}$ is sent when there are at most λ "1"s in the array whose delayed replicas are not completely contained in $\mathbf{B}$, where λ is the decoding threshold. Specifically, if $\mathbf{a}$ satisfies

$$\sum_{w=1}^{W} \prod_{m=1}^{M_g} B\left[l_w + \ell_{g,m}\right] \geq W - \lambda, \tag{34}$$

the decoder decides the array is sent, where W, $\mathbf{l}_g$ and l_w are the Hamming weight, the tap-delay vector and the position of the w-th "1" of the array respectively. To evaluate the performance of MNCOMP, we first give the following Lemma.

Lemma 4 (Probability of Tap-Delay Vector): Define p_g, $g = 1, \ldots, G$ as $\Pr\{B^{[n]}[\ell + \ell_{g,m}] = 1 : m = 1, \ldots, M_g\}\}$. We have

$$\begin{aligned} p_g = {} & 1 + (-1)^1 \sum_{m_1=1}^{M_g} \Pr\left\{B^{[n]}\left[\ell_{g,m_1}\right] = 0\right\} + \cdots \\ & + (-1)^{\mathsf{m}} \sum_{m_1=1}^{M_g-\mathsf{m}+1} \sum_{m_2=m_1+1}^{M_g-\mathsf{m}+2} \cdots \sum_{m_{\mathsf{m}}=m_{\mathsf{m}-1}+1}^{M_g} \\ & \Pr\left\{B^{[n]}\left[\ell_{g,m}\right] = 0 : m = m_1, m_2, \ldots, m_{\mathsf{m}}\right\} + \cdots \\ & + (-1)^{M_g} \Pr\left\{B^{[n]}\left[\ell_{g,m}\right] = 0 : m = 1, \ldots, M_g\right\}. \end{aligned} \tag{35}$$

Proof: We can obtain (35) following the rule of constructing generalized Bloom filters and the inclusion-exclusion principle. ∎

According to the MNCOMP decoding strategy and Lemma 4, we have the following result.

Proposition 4: With the MNCOMP decoding strategy, the proposed scheme achieves a feasible activity recognition cost and a feasible message transmission cost as

$$\Omega_a = -\frac{1}{\kappa_0 \ln p_{\max}}, \qquad \Omega_m = -\frac{(1+\gamma)}{\kappa_m \ln p_{\max}}, \tag{36}$$

TABLE 2. Performance comparison of two decoding strategies.

M_{min}	κ_0 of MNCOMP	Ω_a of MNCOMP	Ω_a of NCOMP
1	ln2	1/ln2	1/ln2
2	0.4142	1.9035	2.8854
3	0.2919	2.4047	4.3281
4	0.2245	2.9160	5.7708
5	0.1820	3.4311	7.2135
6	0.1529	3.9480	8.6562

where

$$p_{\max} = \max_{g=1,\ldots,G} p_g, \tag{37}$$

$$\kappa_0 = \kappa_m = \arg\max_{\kappa\in(0,1)} -\kappa\log(p_{\max}), \tag{38}$$

and $\kappa_1 = \kappa_2 = \kappa_m$.

In particular, on dense multipath fading MnACs, we have

$$p_{\max} = 1 - M_{\min}e^{-\kappa_i\mathbb{E}[M]} + (M_{\min}-1)e^{-\kappa_i(\mathbb{E}[M]+1)}, \quad i = 0,1,2. \tag{39}$$

Proof: See Appendix D. ■

From the proof of Prop. 4, we remark that users with different tap-delay vectors have different error probabilities with the MNCOMP decoding strategy, and the feasible costs are only limited by error probabilities of users whose tap-delay vectors have the largest p_g, i.e., $p_g = p_{\max}$. This is because different tap-delay vectors provide different extra information for decoding and the tap-delay vector with the largest p_g provides the least extra information. In dense multipath, the tap-delay vector with the number of taps $M_{\min}$ has the largest p_g. Therefore, under the same condition of $\mathbb{E}[M]$, the smaller the $M_{\min}$, the more feasible costs are required. Especially when $M_{\min} = 1$, the performance of MNCOMP degrades to that of NCOMP. Table 2 shows the performance comparison of two decoding strategies with different $M_{\min}$ on dense multipath, where $M_{\min} = \mathbb{E}[M]$. It can be seen from the table that the performance of the proposed scheme with MNCOMP is greatly improved compared to that with NCOMP.

Numerical experiments (see Section VII-C) demonstrate that, with the same number of taps, the performance of MNCOMP on sparse multipath may be better than that on dense multipath. From Ω_m in Prop. 4, the asymptotically achievable message length is derived as

$$\log_2 J = -\frac{L\kappa_m \ln p_{\max}}{N_a} - \log_2 N \tag{40}$$

and the asymptotically achievable sum rate is derived as

$$R_{\text{sum}} = -\kappa_m \ln p_{\max} + \frac{\kappa_m \ln p_{\max}}{1+\gamma}. \tag{41}$$

VI. FADING CHANNELS WITHOUT IUS

When the transmitted arrays of each user are not synchronized, all the above decoding strategies can not be utilized directly. Therefore, we propose a sliding window method. Assume the maximum propagation time for users is $\zeta_{\max}$, i.e., $\zeta_{\max} = \max_n \zeta_n$. The method performs $\zeta_{\max}+1$ NCOMP or MNCOMP decodings on $\zeta_{\max}+1$ windows of $\mathbf{B}$.

Take the activity recognition task as an example. First, the sliding window is placed on the 1-st to L_0-th positions of $\mathbf{B}$, and then the decoder performs 1-st decoding on it to obtain the set of active users $\mathcal{N}_1$; Then the window slides to the 2-nd to (L_0+1)-th positions of $\mathbf{B}$, and the decoder performs 2-nd decoding to obtain the set of active users $\mathcal{N}_2$; The rest of $\zeta_{\max}-1$ decodings are done in the same manner. After executing the decoding $\zeta_{\max}+1$ times, the decoder obtains $\zeta_{\max}+1$ sets and the final output is the union of all sets.

We have the following result for the sliding window strategy.

Proposition 5: The proposed scheme using the sliding window method on MnACs without IUS achieves the same feasible activity recognition cost and message transmission cost as that on MnACs with IUS.

Proof: See Appendix E. ■

We remark that the sliding window method is essentially an exhaustive search, and for sufficiently large N there is little loss in performance of the scheme. We also remark that the complexity of the method is $\zeta_{\max}+1$ times that of the original decoding strategy (see Section VII-A for details), which is usually tolerable since $\xi_{\max}$ is typically small or modest.

An alternative way to reduce complexity is to first estimate the propagation times of users, and then only one NCOMP/MNCOMP decoding is required for the window corresponding to the propagation time. Propagation times can be estimate in activity recognition and the method is as follows. Similar to the sliding window method, the decoder still performs $\zeta_{\max}$ decodings and obtains $\zeta_{\max}+1$ sets. For each user n, if it is only contained in one set $\mathcal{N}_\zeta$, we can declare the estimate of the propagation time of user n as $\bar{\zeta}_n = \zeta$.

To characterize the performance of the estimation method, we denote the event that propagation times of all active users are correctly estimated by $\bar{\mathsf{E}}^{\mathbf{EE}} = \{\bar{\zeta}_n = \zeta_n : \forall n \in \mathcal{N}, S_n = 1\}$ and extend definitions of feasible costs which need to satisfy that $\Pr\{\bar{\mathsf{E}}^{\mathbf{EE}}\} \to 1$ as $N \to \infty$, besides the conditions in Def. 3 and Def. 4. We have the following result using the estimation method.

Proposition 6: Define $p_{g,\zeta}$, $g = 1, \ldots, G$, $\zeta = 1, \ldots, \zeta_{\max}$ as $\Pr\{B[\ell+\ell_g[m]] = 1, B[\ell+\ell_g[m]+\zeta] = 1 : m = 1, \ldots, M_g\}$. On frequency-flat fading MnACs without IUS, the proposed scheme using the estimation method achieves a feasible activity recognition cost Ω_a and a feasible cost in Phase 1 of message transmission Ω_1, as

$$\Omega_a = 1/\ln 2, \qquad \Omega_1 = \max\{\beta/\ln 2, (1-\beta)\ln 2\}. \tag{42}$$

On frequency-selective fading MnACs without IUS, the proposed scheme using the estimation method achieves two feasible costs as

$$\Omega_a = \max\left\{\frac{\beta}{-\kappa_0\ln(p^{\mathrm{d}}_{\max})}, \frac{1}{-\kappa_0\ln(p_{\max})}\right\}, \tag{43}$$

TABLE 3. The complexity of different schemes in frequency-flat fading channels with IUS.

Task	Scheme	Encoding Complexity	Decoding Complexity
Activity Recognition	Bloom Filter+NCOMP	$O(N \ln N)$	$O(N)$
	Bernouill+NCOMP [40]	$O(NL)$	$O(N)$
	Gaussian+AMP [13]	$O(NL)$	$O(NL)$
	Gaussian+vector-AMP [11], [12]	$O(NL)$	$O(N_r NL)$
Message Transmission	Bloom Filter+NCOMP	$O(N \ln N)$	$O(\max\{N, JN_a\})$
	Bernouill+M-AMP [23]	$O(NL)$	$O(JN_r NL)$

$$\Omega_1 = \max\left\{\frac{\beta}{-\kappa_1 \ln\left(p^{\mathrm{d}}_{\max}\right)}, \frac{1-\beta}{-\kappa_1 \ln(p_{\max})}\right\}, \tag{44}$$

where $\kappa_0 = \arg\min_{\kappa\in(0,1)} \Omega_a$, $\kappa_1 = \arg\min_{\kappa\in(0,1)} \Omega_1$ and $p^{\mathrm{d}}_{\max} = \max_{g,\zeta}\{\frac{p_{g,\zeta}}{p_g}\}$.

Proof: See Appendix F. ■

We remark that in frequency-flat fading MnACs, the estimation of propagation times of users do not increase the cost Ω_a, and when $\beta \leq 0.5$, the estimation does not increase the cost Ω_m.

VII. COMPLEXITY ANALYSIS AND NUMERICAL RESULTS

A. COMPLEXITY OF CODING SCHEMES

The complexity of our scheme is summarized as follows.

1) Encoding: the complexity of encoding schemes is characterized in terms of the number of hash functions required. Note that the Bloom filters in the encoding schemes are all of parameters $(O(N_a \ln N), O(\ln N))$, and thus the encoding complexity is $O(N \ln N)$.
2) Decoding: the complexity of decoding schemes is characterized in terms of the number of times to determine whether "1" in signature arrays or codewords is in the output array of one-bit quantizer **B**.
   - NCOMP: Note that for verifying an inactive user's Bloom filter signature array, as soon as $\lambda+1$ "1"s are not contained in **B** (λ is typically a modest constant), the receiver can discard this inactive user early, incurring only $O(N)$ times verification for all inactive users. Thus, the complexity of activity recognition is $O(N + \lambda N_a \ln N) = O(N)$. Similarly, the complexity of message transmission is $O(\max\{N, N_a J\})$, respectively.
   - MNCOMP: In this decoding scheme, the average number of checks per signature array/codeword is $\mathbb{E}[M]$ times that of NCOMP, and thus, the complexity of activity recognition and message transmission are $O(N\mathbb{E}[M])$ and $O(\max\{N\mathbb{E}[M], N_a J\mathbb{E}[M]\})$.
   - Sliding window method: Because sliding window method performs $\zeta_{\max}+1$ times of NCOMP or MNCOMP decodings, the complexity of sliding window method is $\zeta_{\max}+1$ times that of the complexity of NCOMP/MNCOMP decoding.

Table 3 provides a comparison of the complexity of different schemes, showing that our scheme is advantageous in terms of complexity.

B. SIMULATION PARAMETERS

In the remainder of this section, we evaluate the performance of the proposed scheme by Monte Carlo simulation. Set the number of potential users of the BS as $N = 5000$, and let each user be active with probability 0.0141, corresponding to $\beta = 0.5$. The number of messages of each user is $J = 71$, corresponding to $\gamma = 0.5$. According to our theoretical analysis, the hard-decision threshold η should increase with N. For this, η is selected as $\eta = \ln(N/2)$. We define the activity recognition cost ratio $\xi_a = \Omega_{a,s}/\Omega_a$ to characterize the activity recognition cost $\Omega_{a,s}$ in simulations based on the feasible activity recognition cost Ω_a. In other words, in the simulation of the activity recognition, the length of the transmitted array is $L_{0,s} = \xi_a \Omega_a N_a \log_2 N$. Similarly, the message transmission cost ratio $\xi_m = \Omega_{m,s}/\Omega_m$ characterizes the message transmission cost $\Omega_{m,s}$ based on the feasible message transmission cost Ω_m. In the simulation of message transmission, the length of the transmitted array is $L_{m,s} = \xi_m \Omega_m N_a \log_2 N$. In frequency-selective channels, we consider three non-zero taps with the same power, and the tap-delay vector is $\mathbf{l}_g = [0, 1, 2]$ in dense multipath and $\mathbf{l}_g = [0, 1, 3]$ in sparse multipath, for all $g = 1, \ldots, G$.

C. ERROR PROBABILITY: ACTIVITY RECOGNITION

First, we define average miss and false recognition probabilities respectively as

$$P^{\mathrm{MR}} = \mathbb{E}\left[N^{\mathrm{MR}}/N_A\right], \tag{45}$$

$$P^{\mathrm{FR}} = \mathbb{E}\left[N^{\mathrm{FR}}/(N - N_A)\right], \tag{46}$$

where N^{MR}, N^{FR} and N_A are the number of active users who are missed to be recognized, the number of inactive users who are falsely recognized and the number of active users in a communication block, respectively.

First we present the performance of the proposed scheme on frequency-flat fading channels in Fig. 3-6, where $\Omega_a = 1/\ln 2$. Fig. 3 shows P^{MR} and P^{FR} versus N, where $\mathsf{SNR} = 10$ dB, and $\xi_a = 1.5$, i.e., $L_{0,s} \approx 2.164 N^{0.5} \log_2 N$. It is observed that P^{MR} and P^{FR} decrease as N grows. Fig. 4 shows P^{MR} and P^{FR} versus SNR, with different λ, and $\xi_a = 1.5$ corresponding to $L_{0,s} = 1881$. It is observed that P^{MR} decreases with SNR, while P^{FR} gradually increases and

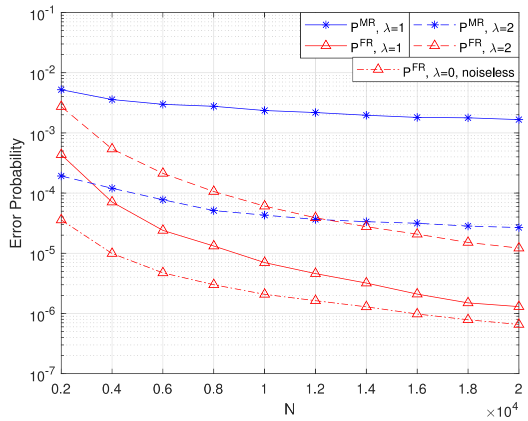


FIGURE 3. **Miss and false recognition probabilities versus the number of potential users N, when SNR = 10 dB, $\xi_a = 1.5$.**

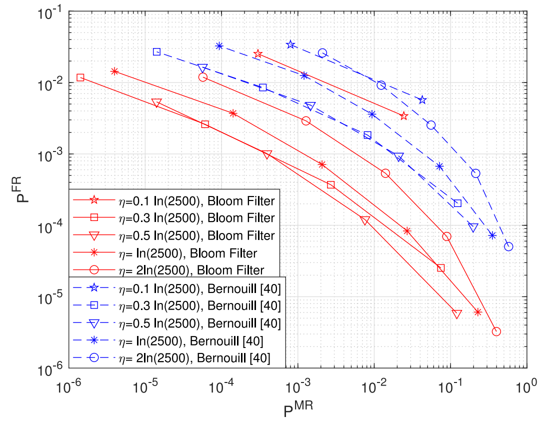


FIGURE 6. **Receiver operating characteristic curves (obtained by adjusting λ) with different η, when $N = 5000$, SNR = 5 dB, and $\xi_a = 1.5$, i.e., $L_{0,s} = 1881$.**

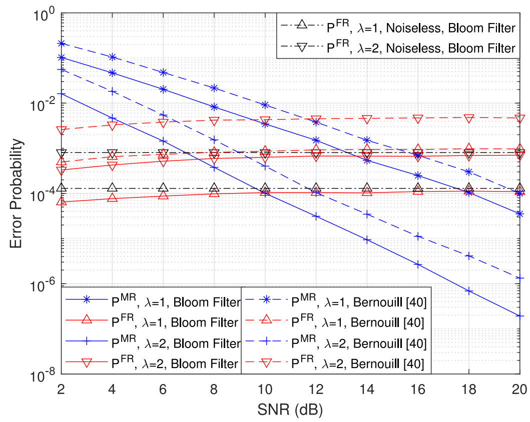


FIGURE 4. **Miss and false recognition probabilities versus SNR with different λ, when $N = 5000$ and $\xi_a = 1.5$, i.e., $L_{0,s} = 1881$.**

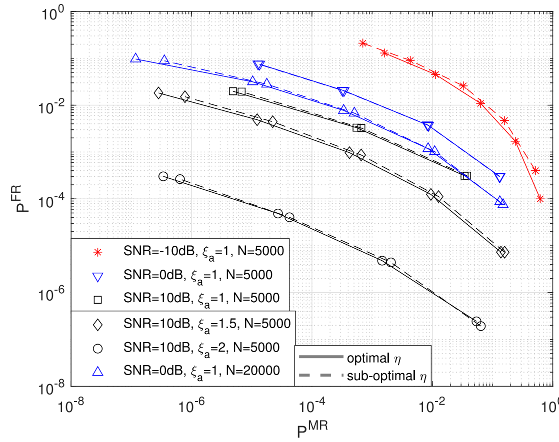


FIGURE 7. **Comparison of receiver operating characteristic curves with optimal η and sub-optimal η, $L_{0,s} = 1254\xi_a$.**

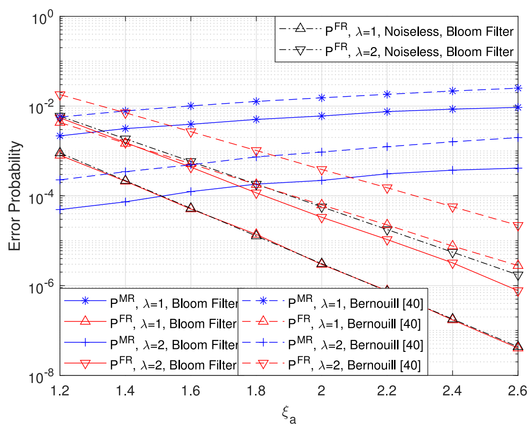


FIGURE 5. **Miss and false recognition probabilities versus ξ_a ($L_{0,s} = 1254\xi_a$) with different λ, when $N = 5000$ and SNR = 10 dB.**

converges to a constant as SNR increases. This is because noise helps avoid false detections, when choosing a relatively large η, e.g., $\eta = \ln(N/2)$ in our simulation. Fig. 5 shows P^{MR} and P^{FR} versus activity recognition cost ratio ξ_a ($L_{0,s} = 1254\xi_a$). It is observed that P^{FR} rapidly decreases, while P^{MR} increases slowly as ξ_a increases. This is because the relatively small ξ_a helps avoid miss detections, given a constant λ. Moreover, in Fig. 4 and 5, it is shown that considerable reductions of P^{MR} can be obtained by increasing λ at the cost of increasing P^{FR}. We can observe from Fig. 4-6 that both the miss and false recognition probabilities of the proposed scheme are significantly smaller than those of the scheme with Bernoulli distribution based coding in [40]. By receiver operating characteristic curves, Fig. 6 shows that compared to the scheme in [40], the miss recognition probability P^{MR} and the false recognition probability P^{FR} of our scheme exhibit a significantly improved tradeoff.

The optimal η for finite N can be found by an exhaustive search, e.g., Fig. 6 shows the optimal $\eta \approx 0.5\ln(2500)$ at 5 dB for $N = 5000$. In addition to the increase in optimal η with increasing number of users N as indicated by the theoretical analysis, we note that the optimal η also increases with SNR and decreases with coding blocklength L. This is because an increase in SNR and L will improve the performance of miss recognition probability P^{MR} and false recognition probability P^{FR} respectively, and thus a better trade-off can be obtained by increasing or decreasing η as SNR and L increase. When it is not practical to obtain the optimal η, we propose a heuristic method to choose the value of η, by making the probability of zero-to-one flip p_o equal

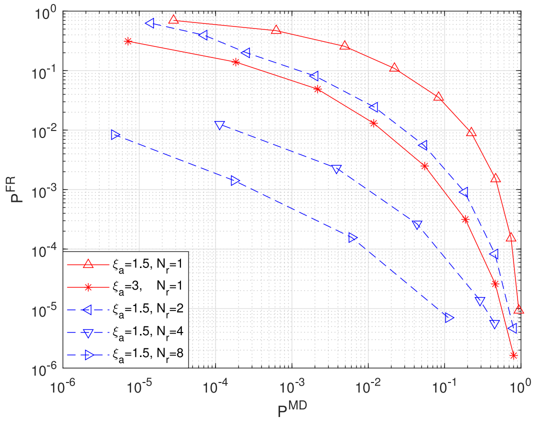


FIGURE 8. Receiver operating characteristic curves in the low SNR regime (SNR = −15dB), when $N = 5000$, $L_{0,s} = 1254\xi_a$.

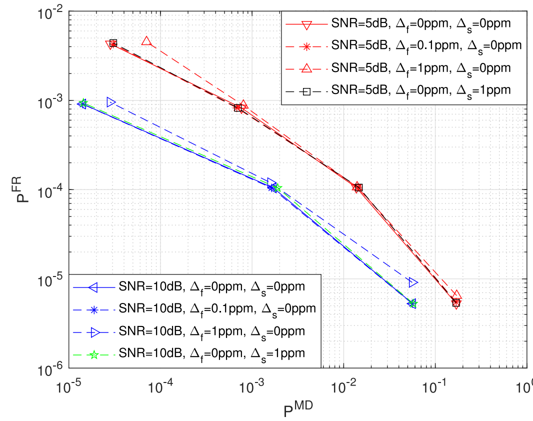


FIGURE 9. Receiver operating characteristic curves with different frequence offsets Δ_f and sampling offsets Δ_s, when $N = 5000$, and $\xi_a = 1.5$, i.e., $L_{0,s} = 1881$.

to the probability of one-to-zero flip p_v; that is, let

$$e^{-\eta} = 1 - e^{-\frac{\eta N_0}{A^2 T_s + N_0}}, \tag{47}$$

to obtain the value of η.[10] The sub-optimal η obtained by this heuristic method increases with the number of users and SNR, but does not change with the coding blocklength. We present the comparison of the performance of our scheme, using the optimal η and the sub-optimal η in Fig 7, and it demonstrates that there is no significant performance loss due to the heuristic method.

The mMTC usually requires low coverage and low cost. For this reason, it is necessary to demonstrate the performance of our scheme in the low SNR regime or in the presence of frequency offset and sampling offset. At low SNRs, we can either increase the coding blocklength L while keeping the number of hash functions K, i.e., $K = L\kappa/\xi_a N_a$ (lowering the code rate), or equip multiple antennas for the BS. Fig. 8 shows receiver operating characteristic curves with different L and number of antennas equiped at the BS N_r in the low SNR regime (parameter η of each curve in Fig. 8 and 9 is the optimal value obtained by exhaustive search), and indicates that our solution still has good performance by increasing L or using multiple antenna techniques, even at −15dB. In Fig. 9, we assume the carrier frequency is 2.4GHz and bandwidth is 125kHz, and show the performance of our scheme with different frequency offsets and sampling offsets. Fig. 9 shows that the detection performance of the scheme is not affected when the frequency offset $\Delta_f = 0.1$ppm and even when $\Delta_f = 1$ppm, the performance degradation is not significant. For the sampling offset Δ_s, even if $\Delta_s = 1$ppm, it still does not affect the performance of the scheme.

Lastly, we present the performance of the proposed scheme in frequency-selective fading channels in Fig. 10-11, taking three non-zero taps as an example. We choose $\Omega_a = 1.6688/\ln 2$, which is the feasible activity recognition cost of dense multipath MnACs with three non-zero taps in Prop. 4. Since the performance of NCOMP is almost

10. Here we use the lower bound of p_v in (12).

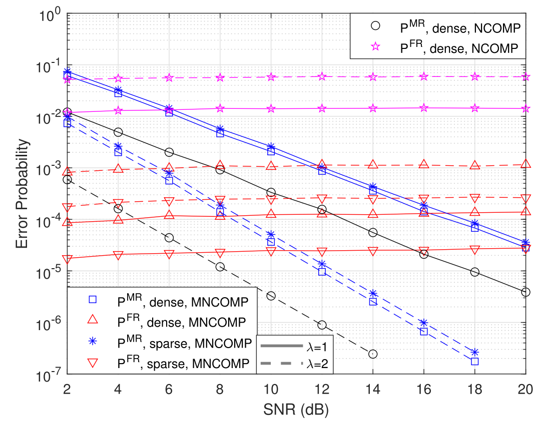


FIGURE 10. Miss and false recognition probabilities versus SNR with different λ and different decoding strategies, when $N = 5000$, $\lambda = 1$ and $\xi_a = 1.5$, i.e., $L_{0,s} = 3138$.

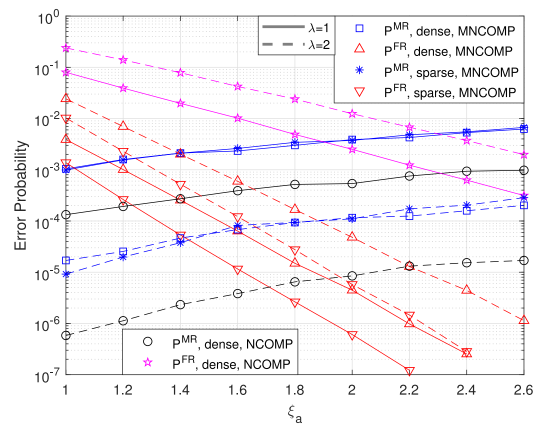


FIGURE 11. Miss and false recognition probabilities versus ξ_a ($L_{0,s} = 2092\xi_a$) with different λ and different decoding strategies, when $N = 5000$, SNR = 10 dB and $\lambda = 1$.

the same for both dense multipath and sparse multipath, we only show its performance for sparse multipath to keep the figure concise. Fig. 10 shows P^{MR} and P^{FR} versus SNR, with different λ, and $\xi_a = 1.5$ corresponding to $L_{0,s} = 3138$. It is observed that, P^{FR} is substantially decreased by the MNCOMP decoding strategy, despite increasing some P^{MR}. It can also be seen that MNCOMP

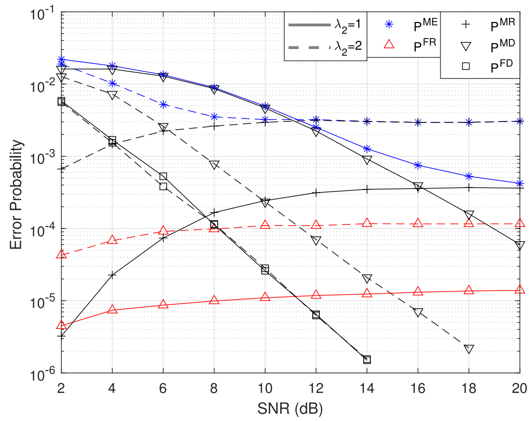


FIGURE 12. Message error and false recognition probabilities versus SNR with different λ_2, when $N = 5000$, $\lambda_1 = 2$ and $\xi_m = 2$, i.e., $L_{m,s} = 3761$.

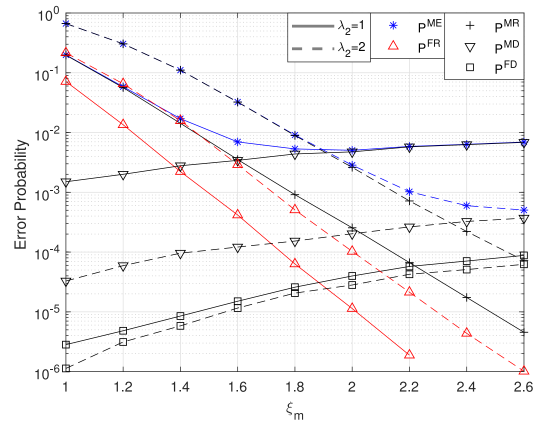


FIGURE 13. Message error and false recognition probabilities versus ξ_m ($L_{m,s} = 1881\xi_m$) with different λ_2, when $N = 5000$, SNR = 10 dB and $\lambda_1 = 2$.

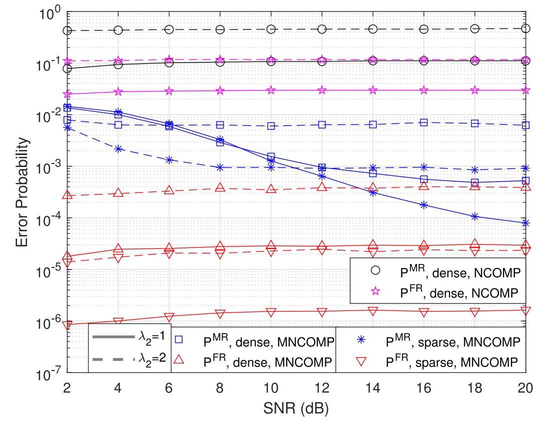


FIGURE 14. Message error and false recognition probabilities versus SNR with different λ_2 and different decoding strategies when $N = 5000$, $\lambda_1 = 2$ and $\xi_m = 2$, i.e., $L_{m,s} = 6276$.

can provide better performance for sparse multipath than that for dense multipath. Fig. 11 shows P^{MR} and P^{FR} versus ξ_a ($L_{0,s} = 2092\xi_a$), with different λ. It is observed that P^{FR} decreases significantly by the MNCOMP decoding strategy as ξ_a increases. Similar to frequency-flat fading channels, Fig. 10-11 also show that P^{MR} reductions can be obtained by increasing λ at the cost of increasing P^{FR}.

D. ERROR PROBABILITY: MESSAGE TRANSMISSION

In message transmission task, the error probability of active users is defined as

$$P^{\mathrm{ME}} = \mathbb{E}\left[N^{\mathrm{ME}}/N_A\right], \tag{48}$$

where N^{ME} is the number of active users who are recognized as inactive users or whose messages are incorrectly decoded. The message error probability contains three error probabilities, which are the miss recognition probability of active users in Phase 1, P^{MR}, the probability that transmitted messages of active users are missed, called miss decoding probability, P^{MD} and the false decoding probability of active users, P^{FD}. In order to focus on the message transmission phase (i.e., Phase 2), we choose $\lambda_1 = 2$ to ensure a small P^{MR}.

First, we present the performance of our scheme in frequency-flat fading channels in Fig. 12-13, where $\Omega_m = (1+\gamma)/\ln 2$. Fig. 12 shows P^{ME} and P^{FR} versus SNR, with different λ_2, and the message transmission cost ratio $\xi_m = 2$ (corresponding to $L_{m,s} = 3761$). It is observed that P^{ME} gradually decreases and converges to a constant as SNR increases. This is because, in the low SNR regime, P^{FD} is relatively small and P^{MD} dominates and as SNR increases, P^{MD} decreases rapidly and P^{FD} gradually dominates. Fig. 13 shows P^{ME} and P^{FR} versus ξ_m ($L_{m,s} = 1881\xi_m$), when SNR $= 10$ dB. Contrary to the curve of P^{ME} versus SNR, P^{FD} dominates in the low ξ_m regime, while P^{MD} dominates in the high ξ_m regime. Moreover, it is shown that considerable reductions of P^{ME} can be obtained by choosing $\lambda_2 = 2$ in the high ξ_m regime. On the contrary, in the higher SNR regime, choosing smaller λ_2 can obtain better performance of P^{ME}.

Then, we present the performance of the proposed scheme in frequency-selective fading channels in Fig. 14-15, taking the three non-zero taps as an example. We choose $\Omega_m = 1.6688(1+\gamma)/\ln 2$, which is the feasible message transmission cost of our scheme for dense multipath in Prop. 4. It is observed that P^{ME} and P^{FR} are reduced substantially by the MNCOMP decoding strategy and the proposed scheme with MNCOMP can provide better performance for sparse multipath than that for dense multipath.

E. ERROR PROBABILITY: FADING CHANNELS WITHOUT IUS

Taking activity recognition as an example, in Fig. 16-17, we provide the performance of the proposed scheme using the sliding window method without IUS with $\zeta_{\max} = 3$. It is observed that the performance approaches that of the scheme with IUS. In particular, the performance appears to be completely unaffected by user asynchrony in dense multipath fading channels. In addition, Fig. 16 also presents the probability that the propagation times of active users are incorrectly estimated, P^{EE}, with the proposed estimation method. It can be seen that the performance approaches that

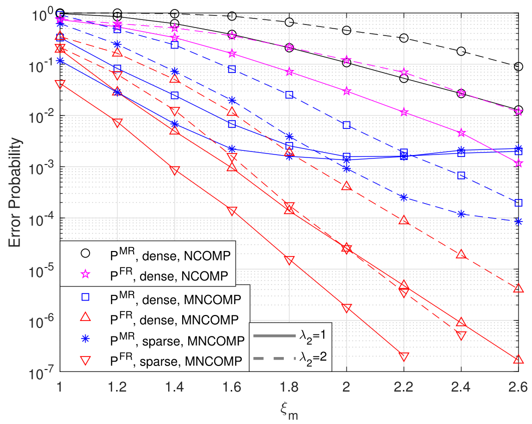


FIGURE 15. Message error and false recognition probabilities versus ξ_m ($L_{m,s} = 3138\xi_a$) with different λ_2 and decoding strategies when $N = 5000$, SNR = 10 dB and $\lambda_1 = 2$.

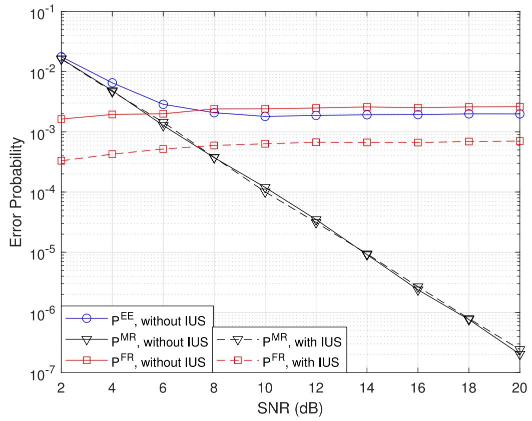


FIGURE 16. Various error probabilities versus SNR in frequency-flat fading channels with and without IUS, when $N = 5000$, $\lambda = 2$, and $\xi_a = 1.5$, i.e., $L_{0,s} = 1881$.

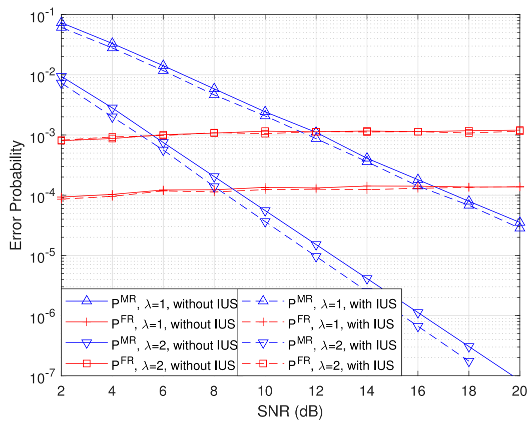


FIGURE 17. Miss and false recognition probabilities versus SNR in dense multipath channels with and without IUS when $N = 5000$ and $\xi_a = 1.5$, i.e., $L_{0,s} = 3138$.

of the scheme with IUS and the proposed estimation method can obtain a good estimation performance.

VIII. CONCLUSION

In this work, we propose a scheme using Bloom filter based coding and OOK modulation at transmitters and hard-decision envelope detection at the BS to realize a simple massive access system. We adopt the NCOMP decoding strategy in frequency-flat fading scenarios with IUS and propose the MNCOMP decoding strategy to improve the performance of activity recognition and message transmission in frequency-selective fading scenarios with IUS. In addition, a sliding window method is proposed to modify above decoding strategies to handle scenarios without IUS. By asymptotic analysis, we demonstrate that the proposed scheme using above decoding strategies can achieve asymptotically vanishing error probability as the number of potential users grows without bound, while guaranteeing a positive coding rate. Moreover, using the sliding window method, the performance of our scheme is not affected when IUS is absent. Numerical experiments for finite number of users verify analytical results and demonstrate that the proposed scheme provides good performance of activity recognition and message transmission with various modeling assumptions.

APPENDIX A
PROOF OF PROPOSITION 1

Note that the number of active users N_A is a binomial random variable of mean N_a. Thus, for any $\delta > 0$, the probability of erroneous recognition is derived as

$$\begin{aligned}
&\Pr\left[\mathsf{E}^{[r]}\right] \\
&\quad = \Pr\left[\mathsf{E}^{[r]}\big||N_A - N_a| \le \delta N_a\right]\Pr[|N_A - N_a| \le \delta N_a] + \\
&\quad \Pr\left[\mathsf{E}^{[r]}\big||N_A - N_a| > \delta N_a\right]\Pr[|N_A - N_a| > \delta N_a] \\
&\quad \le \max_{|a - N_a| \le \delta N_a} \Pr\left[\mathsf{E}^{[r]}|N_A = a\right] + \\
&\quad \Pr[|N_A - N_a| > \delta N_a]. \qquad (49)
\end{aligned}$$

Since $\Pr[|N_A - N_a| > \delta N_a] \to 0$ with $N \to \infty$ for any $\delta > 0$, we only need to ensure $\Pr[\mathsf{E}^{[r]}|N_A = a] \to 0$ for any $(1 - \delta)N_a \le a \le (1 + \delta)N_a$.

Since the event of erroneous recognition contains that active users are recognized as inactive users and inactive users are falsely recognized as active users, we have

$$\begin{aligned}
\Pr\left[\,\mathsf{E}^{[r]}\middle|N_A = a\right] &\le \Pr\left[\,\mathsf{E}^{MR}\middle|N_A = a\right] \\
&\quad + \Pr\left[\,\mathsf{E}^{FR}\middle|N_A = a\right]. \qquad (50)
\end{aligned}$$

This means that we need to ensure $\Pr[\mathsf{E}^{MR}|N_A = a] \to 0$ and $\Pr[\mathsf{E}^{FR}|N_A = a,] \to 0$ for any $(1 - \delta)N_a \le a \le (1 + \delta)N_a$.

Denote the Hamming weight of the signature array of user n by $W_n^{[s]}$, $n = 1, \ldots, N$. The probability that active user n is recognized as an inactive user is derived as

$$\begin{aligned}
&\Pr\left[\,\hat{S}_n = 0\middle|S_n = 1, N_A = a\right] \\
&\quad = \sum_{w=1}^{K_0} \Pr\left[\,\hat{S}_n = 0\middle|S_n = 1, W_n^{[s]} = w_n^{[s]}, N_A = a\right] \\
&\quad \Pr\left[\,W_n^{[s]} = w_n^{[s]}\middle|N_A = a\right] \qquad (51) \\
&\quad < \binom{K_0}{\lambda + 1}(p_v)^{\lambda+1} < (K_0 p_v)^{\lambda+1}, \qquad (52)
\end{aligned}$$

where (51) is because the Hamming weight of the signature array is at least one and at most K_0, and (52) is because an active user n being recognized as inactive corresponds to its signature array not satisfying (20). Then, we obtain an upper bound of $\Pr[\mathsf{E}^{\mathrm{MR}}|N_A = a]$ as

$$\Pr\Big[\mathsf{E}^{\mathrm{MR}}\Big|N_A = a\Big] \le a\Pr\Big[\hat{S}_n = 0\Big|S_n = 1, N_A = a\Big] < a(K_0 p_v)^{\lambda+1}. \tag{53}$$

From (53), noting that $\mathbb{E}[W] = (1-p)L_0$, $p = (1-\frac{1}{L_0})^{aK_0}$, we derive that

$$\lim_{N\to\infty}\max_{|a-N_a|\le\delta N_a}\Pr\Big[\mathsf{E}^{\mathrm{MR}}\Big|N_A = a\Big] = 0, \tag{54}$$

by letting

$$\eta = o\left(\frac{N^{\frac{\beta\lambda}{\lambda+1}}}{\log_2 N}\right), \quad \lambda > 0, \tag{55}$$

and choosing a sufficiently small δ.

Denote the Hamming weight of $\mathbf{B}^{[n]}$ as W and the false-number and the miss-number of $\mathbf{B}$ as O and V, respectively. We can derive $\Pr[\mathsf{E}^{\mathrm{FR}}|N_A = a]$ as

$$\begin{aligned}&\Pr\Big[\mathsf{E}^{\mathrm{FR}}\Big|N_A = a\Big]\\ &= 1 - \min_{\substack{|w-\mathbb{E}[W]|\le\epsilon L_0,\\ |o-\mathbb{E}[O|w]|\le\epsilon_1 L_0,\\ |v-\mathbb{E}[V|w]|\le\epsilon_2 L_0}} \Pr\Big[\bar{\mathsf{E}}^{\mathrm{FR}}\Big|V = v, O = o, W = w, N_A = a\Big]\\ &\Pr[|O-\mathbb{E}[O|w]| \le \epsilon_1 L_0|W = w, N_A = a]\\ &\Pr[|V-\mathbb{E}[V|w]| \le \epsilon_2 L_0|W = w, N_A = a]\\ &\Pr[|W-\mathbb{E}[W]| \le \epsilon L_0|N_A = a].\end{aligned} \tag{56}$$

Note that the Hamming weight of $\mathbf{B}$ is $(W+V-O)$. The probability that active user n is recognized as inactive is derived as

$$\Pr\Big[\hat{S}_n = 1\Big|S_n = 0, V = v, O = o, W = w, N_A = a\Big] < 2\left(\frac{w-v+o}{L_0}\right)^{K_0-\lambda}, \tag{57}$$

which is because an active user being recognized as inactive corresponds to its signature array satisfying (20). Thus, the probability of correctly recognizing all $(N-a)$ inactive users satisfies

$$\Pr\Big[\bar{\mathsf{E}}^{\mathrm{FR}}\Big|V = v, O = o, W = w, N_A = a\Big] > \left[1 - 2K_0^{\lambda}\left(\frac{w+o-v}{L_0}\right)^{K_0-\lambda}\right]^{N-a}. \tag{58}$$

From (56) and (58), utilizing Lemma 2 and Lemma 3, and noting that $\mathbb{E}[O|w] = (L-w)p_o$ and $\mathbb{E}[V|w] = wp_v$, we derive that

$$\lim_{N\to\infty}\max_{|a-N_a|\le\delta N_a}\Pr\Big[\mathsf{E}^{\mathrm{MR}}\Big|N_A = a\Big] = 0, \tag{59}$$

for any

$$\Omega_a > \frac{1}{\log_2\big(1-e^{-\kappa_0}(1-e^{\eta})\big)} \tag{60}$$

by letting $\lambda = o(\log_2 N/\log_2(\log_2 N))$ and choosing sufficiently small ϵ, ϵ_1, ϵ_2 and δ. In particular, when η satisfies $\lim_{N\to\infty} e^{-\eta} \to 0$, and $\kappa_0 = \ln 2$, we have $\Omega_a > 1/\ln 2$. Thus, we have completed the proof of Prop. 1.

APPENDIX B PROOF OF PROPOSITION 2

Denote the Hamming weight, the false-number and the miss-number of $\mathbf{B}$ of Phase i as W_i, O_i and V_i, $i = 1, 2$. Similar to the proof of Prop. 1, we need to ensure probabilities of E^{FR}, E^{MR}, E^{MD} and E^{FD} simultaneously tend to 0 conditioned on $N_A = a$, for any $(1-\delta)N_a \le a \le (1+\delta)N_a$.

For E^{MR} and E^{MD}, similar to (53), we have

$$\Pr\Big[\mathsf{E}^{\mathrm{MR}}\Big|N_A = a\Big] < a(K_1 p_v)^{\lambda_1+1}, \tag{61}$$

$$\Pr\Big[\mathsf{E}^{\mathrm{MD}}\Big|N_A = a\Big] < a(K_2 p_v)^{\lambda_2+1}. \tag{62}$$

Similar to (56) and (58), for E^{FD}, we have

$$\begin{aligned}&\Pr\Big[\mathsf{E}^{\mathrm{FD}}\Big|N_A = a\Big]\\ &< 1 - \min_{\substack{|w_2-\mathbb{E}[W_2]|\le\epsilon L_0,\\ |o_2-\mathbb{E}[O_2|w_2]|\le\epsilon_1 L_0,\\ |v-\mathbb{E}[V_2|w_2]|\le\epsilon_2 L_0}} \Pr\Big[\bar{\mathsf{E}}^{\mathrm{FD}}\Big|V_2 = v_2, O_2 = o_2, W_2 = w_2, N_A = a\Big]\\ &\Pr[|O_2-\mathbb{E}[O_2|w_2]| \le \epsilon_1 L_2|W_2 = w_2, N_A = a]\\ &\Pr[|V_2-\mathbb{E}[V_2|w_2]| \le \epsilon_2 L_2|W_2 = w_2, N_A = a]\\ &\Pr[|W_2-\mathbb{E}[W_2]| \le \epsilon L_2|N_A = a],\end{aligned} \tag{63}$$

where

$$\Pr\Big[\bar{\mathsf{E}}^{\mathrm{FD}}\Big|V_2 = v_2, O_2 = o_2, W_2 = w_2, N_A = a\Big] > \left[1 - 2K_2^{\lambda_2}\left(\frac{w_2+o_2-v_2}{L_2}\right)^{K_2-\lambda_2}\right]^{a(J-1)}. \tag{64}$$

For E^{FR}, we have

$$\begin{aligned}&\Pr\Big[\bar{E}^{\mathrm{FR}}\Big|N_A = a\Big]\\ &< 1 - \min_{\substack{|w_i-(1-p)L_i|\le\epsilon L_i,\\ |o_i-\mathbb{E}[O_i|w_i]|\le\epsilon_1 L_i,\\ |v_i-\mathbb{E}[V_i|w_i]|\le\epsilon_2 L_i,\\ i=1,2}} \Pr\Big[\mathsf{E}^{\mathrm{FR}}\Big|O_i = o_i, V_i = v_i, W_i = w_i,\\ &\qquad N_A = a : i = 1, 2\Big]\\ &\prod_{i=1}^{2}\Pr[|O_i-\mathbb{E}[O_i]| \le \epsilon_1 L_i|W_i = w_i, N_A = a]\\ &\Pr[|V_i-\mathbb{E}[V_i]| \le \epsilon_2 L_i|W_i = w_i, N_A = a]\\ &\Pr[|W_i-\mathbb{E}[W_i]| \le \epsilon L_i|N_A = a],\end{aligned} \tag{65}$$

and

$$\begin{aligned}&\Pr\Big[\bar{E}^{\mathrm{FR}}\Big|O_i = o_i, V_i = v_i, W_i = w_i, N_A = a : i = 1, 2\Big]\\ &> \left\{1 - 2K_1^{\lambda_1}\left(\frac{w_1+o_1-v_1}{L_1}\right)^{K_1-\lambda_1}\right.\\ &\qquad \left.\left[1 - \left[1 - 2K_2^{\lambda_2}\left(\frac{w_2+o_2-v_2}{L_2}\right)^{K_2-\lambda_2}\right]^{J}\right]\right\}^{N-a},\end{aligned} \tag{66}$$

where (66) is because an inactive user being recognized as active corresponds to its signature array satisfying (20) in

Phase 1 and one of its messages arrays satisfying (20) in Phase 2.

From (61), (62), (64) and (66), we derive that

$$\lim_{N\to\infty} \max_{|a-N_a|\le\delta N_a} \Pr\left[\mathsf{E}^{[\mathrm{m}]}\middle|N_A=a\right]=0 \tag{67}$$

for any

$$\Omega_2 \ln 2-\beta-\gamma>0, \qquad \Omega_m \ln 2-1-\gamma>0, \tag{68}$$

by letting decoding thresholds satisfy

$$\lambda_i>0, \quad \lambda_i=o\left(\frac{\log_2 N}{\log_2(\log_2 N)}\right), \quad \text{for } i=1,2, \tag{69}$$

and hard-decision threshold satisfy

$$\eta=o\left(N^{\frac{\beta\lambda_i}{\lambda_i+1}}/\log_2 N\right), \quad \lim_{N\to\infty} e^{-\eta}\to 0, \tag{70}$$

choosing sufficiently small ϵ, ϵ_i and δ, and setting $\kappa_i=\ln 2$, $i=1,2$.

From (68), we have $\Omega_2>(\beta+\gamma)/\ln 2$, $\Omega_m>(1+\gamma)/\ln 2$ (i.e., $\Omega_1>(1+\beta)/\ln 2$). Thus, we have completed the proof of Prop. 2.

APPENDIX C PROOF OF PROPOSITION 3

Consider the activity recognition task. Denote W as the weight of $\mathbf{B}^{[n]}$. Note that the derivation of the probability of each event is similar to the proof in Prop. 1, and only $\mathbb{E}[W]$ differs. Denote $\mathsf{Z}=L_0-W$. When $N_A=a$, the expectation of Z is derived as

$$\begin{aligned}\mathbb{E}[\mathsf{Z}]&=\prod_{g=1}^{G}\left[\left(1-\frac{M_g}{L_0}\right)^{K_0}\right]^{\alpha_g a} L_0\\ &=\exp\left(\frac{a\sum_{g=1}^{G} M_g\alpha_g\kappa_0}{N_a}\right)L_0.\end{aligned} \tag{71}$$

Thus, the expectation of W is obtained as

$$\mathbb{E}[W]=1-\mathbb{E}[\mathsf{Z}]=\left[1-\exp\left(\frac{a\mathbb{E}[M]\kappa_0}{N_a}\right)\right]L_0. \tag{72}$$

By choosing δ sufficiently small, we have

$$\mathbb{E}[W]=(1-\exp(\mathbb{E}[M]\kappa_0))L_0. \tag{73}$$

Substituting (73) into the formula of probability in the Appendix A, we can obtain that $\Omega_a>\mathbb{E}[M]/\ln 2$ by choosing $\kappa_0=1/\mathbb{E}[M]$, and η and λ are the same as the results of Appendix A. The proof of Ω_m is similar and thus omitted. Thus, we have completed the proof of Prop. 3.

APPENDIX D PROOF OF PROPOSITION 4

Consider the activity recognition task. Denote W_g as $\mathbf{l}_g$-weight of $\mathbf{B}^{[n]}$, and O_g as $\mathbf{l}_g$-false-number of $\mathbf{B}$. Similar to the proof of Prop. 1, we only need to ensure $\Pr[\mathsf{E}^{\mathrm{MR}}|N_A=a]\to 0$ and $\Pr[\mathsf{E}^{\mathrm{FR}}|N_A=a]\to 0$ for any $(1-\delta)N_a\le a\le(1+\delta)N_a$.

Similar to (53), we can obtain the upper bound of the probability of E^{MR} as

$$\begin{aligned}\Pr\left[\mathsf{E}^{\mathrm{MR}}\middle|N_A=a\right]&\le a\sum_{g=1}^{G}\alpha_g\Pr\left[\hat{S}_n=0\middle|S_n=1,N_A=a\right]\\ &<a\sum_{g=1}^{G}\alpha_g K_0^{\lambda+1}\left(M_g p_{\mathrm{v}}\right)^{\lambda+1}.\end{aligned} \tag{74}$$

Then considering E^{FR}, we have

$$\begin{aligned}&\Pr\left[\mathsf{E}^{\mathrm{FR}}\middle|N_A=a\right]\\ &<1-\min_{\substack{|W_g-\mathbb{E}[W_g]|\le\epsilon L_0,\\ |O_g-\mathbb{E}[O_g]|\le\epsilon_1 L_0,\\ |V_g-\mathbb{E}[V_g]|\le\epsilon_2 L_0,\\ g=1,\ldots,G}}\Pr\left[\bar{\mathsf{E}}^{\mathrm{FR}}\middle|O_g=o_g,V_g=0,W_g=w_g,\right.\\ &\qquad\left.N_A=a:g=1,\ldots,G\right]\\ &\Pr\left[\left|O_g-\mathbb{E}[O_g]\right|\le\epsilon_1 L_0\middle|W_g=w_g,N_A=a:g=1,\ldots,G\right]\\ &\Pr\left[\left|V_g-\mathbb{E}[V_g]\right|\le\epsilon_2 L_0\middle|W_g=w_g,N_A=a:g=1,\ldots,G\right]\\ &\Pr\left[\left|W_g-\mathbb{E}[W_g]\right|\le\epsilon L_0\middle|N_A=a:g=1,\ldots,G\right].\end{aligned} \tag{75}$$

Similar to (58), we have

$$\begin{aligned}&\Pr\left[\bar{\mathsf{E}}^{\mathrm{FR}}\middle|O_g=o_g,V_g=v_g,W_g=w_g,N_A=a:g=1,\ldots,G\right]\\ &>\prod_{g=1}^{G}\left[1-2K_0^{\lambda}\left(\frac{w_g+o_g-v_g}{L_0}\right)^{K_0-\lambda}\right]^{\alpha_g(N-a)}.\end{aligned} \tag{76}$$

From (74), (75) and (76), we derive that

$$\begin{aligned}&\lim_{N\to\infty}\max_{|a-N_a|\le\delta N_a}\Pr\left[\mathsf{E}^{\mathrm{MR}}\middle|N_A=a\right]=0\\ &\lim_{N\to\infty}\max_{|a-N_a|\le\delta N_a}\Pr\left[\mathsf{E}^{\mathrm{FR}}\middle|N_A=a\right]=0\end{aligned} \tag{77}$$

for any

$$\Omega_a>-1/\left[\kappa_0\log(p_{\max})\right], \tag{78}$$

by letting η and λ be the same as the results of Appendix A and choosing sufficiently small ϵ, ϵ_1, ϵ_2 and δ. where $\kappa_0=\arg\max_{\kappa\in(0,1)}-\kappa\log(p_{\max})$. Thus, we have proved Ω_a. The proof of Ω_m is similar and thus omitted.

For dense multipath, note that $p_{\max}=p_g$, for g satisfying $M_g=M_{\min}$. The m-th ($\mathsf{m}=1,\ldots,M_{\min}$) cumulative term to the right of the equal sign of (35) can be derived as

$$\sum_{m_1=1}^{M_{\min}-\mathsf{m}+1}\cdots\sum_{m_{\mathsf{m}}=m_{\mathsf{m}-1}+1}^{M_{\min}}\Pr\{B^{[n]}[m]=0:m=m_1,\ldots,m_{\mathsf{m}}\}. \tag{79}$$

Denote $\Pr\{B^{[n]}[\ell+m]=0:m=1,\ldots,\mathsf{m}\}$ as P_{m}. When $L\to\infty$, we have

$$\begin{aligned}P_{\mathsf{m}}&=\prod_{g=1}^{G}\left(1-\frac{\mathsf{m}+M_g-1}{L}\right)^{\alpha_g N_a K}\\ &=e^{-\kappa(\mathbb{E}[M]+\mathsf{m}-1)}.\end{aligned} \tag{80}$$

Note that, $\Pr\{B^{[n]}[\ell+1]=0, B^{[n]}[\ell+\mathsf{m}]=0\}=P_{\mathsf{m}}$. We can derive (79) as

$$\binom{\mathsf{m}-2}{\mathsf{m}-2}(M_{\min}-\mathsf{m}+1)p_{\mathsf{m}} + \binom{\mathsf{m}-1}{\mathsf{m}-2}(M_{\min}-\mathsf{m})p_{\mathsf{m}+1}+\cdots + \binom{M_{\min}-2}{\mathsf{m}-2}p_{M_{\min}}. \tag{81}$$

Substituting (79) and (81) into (35), we can derive (39). Thus, we have completed the proof of Prop. 4.

APPENDIX E
PROOF OF PROPOSITION 5

The proof takes activity recognition task in frequency-selective fading MnACs as an example. First, consider the probability that at least one inactive user is falsely recognized as an active user. For each active user, there exists one delayed window that can match the propagation time of it. Thus, the probability is derived as

$$\Pr\left[\mathsf{E}^{\mathrm{MR}}\middle|N_A=a\right] < a\sum_{g=1}^{G}\alpha_g {K_0}^{\lambda+1}\left(M_g p_{\mathrm{v}}\right)^{\lambda+1}. \tag{82}$$

Then, consider the probability that none of the inactive users is falsely recognized. When $W_g=w_g$, $V_g=v_g$, $O_g=o_g$, $g=1,\ldots,G$ and $N_A=a$, it can be derived as

$$\Pr\left[\bar{\mathsf{E}}^{\mathrm{FR}}\middle|V_g=v_g, O_g=o_g, W_g=w_g, N_A=a: g=1,\ldots,G\right] > \prod_{g=1}^{G}\left[\left(1-2{K_0}^{\lambda}\left(\frac{w_g+o_g-v_g}{L_0}\right)^{K_0-\lambda}\right)^{\zeta_{\max}+1}\right]^{\alpha_g(N-a)}. \tag{83}$$

This is because all inactive users being correctly recognized corresponds to signature arrays of all $N-a$ inactive users not satisfying (34) in all $\zeta_{\max}+1$ delayed windows. After a series of derivations similar to those in the proof of Prop. 4, we can obtain the same result as Prop. 4. Other proofs are similar and thus omitted. Thus, we have completed the proof of Prop. 5.

APPENDIX F
PROOF OF PROPOSITION 6

Let us take the activity recognition task in frequency-selective MnACs with the MNCOMP decoding strategy as an example. Without loss of generality, we assume that the propagation time of each active user is zero. The probability of correctly estimating the propagation time of an active user n with $\mathbf{l}_g$ is derived as

$$\Pr\left[\bar{\zeta}_n=\zeta_n\middle|S_n=1, N_A=a\right] > \left[1-(K_0p_{\mathrm{v}})^{\lambda+1}\right] \prod_{\zeta=1}^{\zeta_{\max}}\left(1-2K_0^{\lambda}\left(\frac{p_{g,\zeta}+2M_g(1-p_{g,\zeta})p_o}{p_g}\right)^{K-\lambda}\right). \tag{84}$$

This is because the propagation time of the active user being correctly estimated corresponds to its signature array satisfying (34) only in the first delayed window. Thus, a lower bound of the probability that propagation times of all active users are correctly estimated is written as

$$\begin{aligned}\Pr&\left[\bar{\mathsf{E}}^{\mathrm{EE}}\middle|N_A=a\right]\\ &> \prod_{g=1}^{G}\left[\left[1-(K_0p_{\mathrm{v}})^{\lambda+1}\right] \prod_{\zeta=1}^{\zeta_{\max}}\left(1-2{K_0}^{\lambda}\left(\frac{p_{g,\zeta}+2M_g(1-p_{g,\zeta})p_o}{p_g}\right)^{K_0-\lambda}\right)\right]^{a\alpha_g}\\ &> 1-\sum_{g=1}^{G}a\alpha_g(K_0p_{\mathrm{v}})^{\lambda+1} - \sum_{g=1}^{G}\sum_{\zeta=1}^{\zeta_{\max}}2a\alpha_g{K_0}^{\lambda}\left(\frac{p_{g,\zeta}+2M_g(1-p_{g,\zeta})p_o}{p_g}\right)^{K_0-\lambda}. \end{aligned} \tag{85}$$

The probability that none of the inactive users is falsely recognized is also (83). After manipulations of (83) and (85), we can obtain

$$\Omega_a > \max\left\{\frac{\beta}{-\kappa_0\ln\left(p_{\max}^{\mathrm{d}}\right)}, \frac{1}{-\kappa_0\ln p_{\max}}\right\}, \tag{86}$$

where $\kappa_0=\arg\min_{\kappa\in(0,1)}\Omega_a$.

For frequency-flat fading MnACs, we note that $p_{g,\zeta}=p^2$, $\zeta=1,\ldots,\zeta_{\max}$, and $p_g=p$, where $p=2^{a/N_a}$. Therefore,

$$\Pr\left[\bar{\mathsf{E}}^{\mathrm{EE}}\middle|N_A=a\right] > 1-(Kp_{\mathrm{v}})^{\lambda+1}-2K_0^{\lambda}\left[p+(1-p)p_o\right]^{K_0-\lambda}. \tag{87}$$

Finally, we can obtain that $\Omega_a>\max\{\beta/\ln 2, 1/\ln 2\}$. Other proofs are similar and are therefore omitted. Thus, we have completed the proof of Prop. 6.

REFERENCES

[1] R. Deng, and W. Zhang, “Massive connectivity with bloom filter based coding and hard-decision envelope detection detection,” in *Proc. IEEE 23th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Oulu, Finland, Jul. 2022, pp. 1–5.
[2] *ITU-R M.[IMT-2020.TECH PERF REQ]-Minimum Requirements Related to Technical Performance for IMT-2020 Radio Interface(s)*, Rec. ITU-R M.2410-0, Int. Telecommun. Union, Geneva, Switzerland, Nov. 2017.
[3] 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.
[4] *Uplink Multiple Access Schemes for NR*, document R1-165174, TSG-RAN WG1 Meeting #85, 3GPP, Sophia Antipolis, France, NTT DOCOMO, Tokyo, Japan, May 2016.
[5] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, “Massive access for 5G and beyond,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Mar. 2021.
[6] X. Chen, T.-Y. Chen, and D. Guo, “Capacity of Gaussian many-access channels,” *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3516–3539, Jun. 2017.
[7] F. Wei, Y. Wu, W. Chen, W. Yang, and G. Caire, “On the fundamental limits of MIMO massive multiple access channels,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019, pp. 1–6.

[8] H. Zhu and G. B. Giannakis, "Exploiting sparse user activity in multiuser detection," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 454–465, Feb. 2011.
[9] R. Xie, H. Yin, X. Chen, and Z. Wang, "Many access for small packets based on precoding and sparsity-aware recovery," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4680–4694, Nov. 2016.
[10] 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.
[11] Z. Chen, F. Sohrabi, and W. Yu, "Sparse activity detection for massive connectivity," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 1890–1904, Apr. 2018.
[12] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic, and E. de Carvalho, "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 88–99, Sep. 2018.
[13] Z. Sun, Z. Wei, L. Yang, J. Yuan, X. Cheng, and L. Wan, "Exploiting transmission control for joint user identification and channel estimation in massive connectivity," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6311–6326, Sep. 2019.
[14] J. Wang, Z. Zhang, and L. Hanzo, "Joint active user detection and channel estimation in massive access systems exploiting Reed-Muller sequences," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 739–752, Jun. 2019.
[15] Y. Cheng, L. Liu, and P. Li, "Orthogonal AMP for massive access in channels with spatial and temporal correlations," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 726–740, Mar. 2021.
[16] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, "Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO," *IEEE Trans. Signal Process.*, vol. 68, pp. 764–779, 2020.
[17] X. Shao, X. Chen, C. Zhong, and Z. Zhang, "Joint activity detection and channel estimation for mmW/THz wideband massive access," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1–6.
[18] A. Fengler, S. Haghighatshoar, P. Jung, and G. Caire, "Non-Bayesian activity detection, large-scale fading coefficient estimation, and unsourced random access with a massive MIMO receiver," *IEEE Trans. Inf. Theory*, vol. 67, no. 5, pp. 2925–2951, May 2021.
[19] L. Cheng, B. Henty, F. Bai, and D. D. Stancil, "Doppler spread and coherence time of rural and highway vehicle-to-vehicle channels at 5.9 GHz," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, New Orleans, LO, USA, Dec. 2008, pp. 1–6.
[20] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2383–2399, Sep. 2001.
[21] M. Chowdhury, A. Manolakos, and A. Goldsmith, "Scaling laws for noncoherent energy-based communications in the SIMO MAC," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1980–1992, Apr. 2016.
[22] A. Manolakos, M. Chowdhury, and A. Goldsmith, "Energy-based modulation for noncoherent massive SIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7831–7846, Nov. 2016.
[23] K. Senel and E. G. Larsson, "Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6164–6175, Dec. 2018.
[24] Z. Tang, J. Wang, J. Wang, and J. Song, "Device activity detection and non-coherent information transmission for massive machine-type communications," *IEEE Access*, vol. 8, pp. 41452–41465, 2020.
[25] J. Huang, H. Zhang, C. Huang, and W. Zhang, "Compressed multiple random access with energy modulation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2020, pp. 1–6.
[26] Y. Polyanskiy, "A perspective on massive random access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2523–2527.
[27] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient random access for the quasi-static fading MAC," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 2768–2772.
[28] A. Fengler, G. Caire, P. Jung, and S. Haghighatshoar, "Massive MIMO unsourced random access," 2019, *arXiv:1901.00828*.
[29] A. Pradhan, V. Amalladinne, A. Vem, K. R. Narayanan, and J.-F. Chamberland, "A joint graph based coding scheme for the unsourced random access Gaussian channel," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
[30] E. Marshakov, G. Balitskiy, K. Andreev, and A. Frolov, "A polar code based unsourced random access for the Gaussian MAC," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, 2019, pp. 1–5.
[31] V. K. Amalladinne, J.-F. Chamberland, and K. R. Narayanan, "A coded compressed sensing scheme for unsourced multiple access," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6509–6533, Oct. 2020.
[32] R. Calderbank and A. Thompson, "CHIRRUP: A practical algorithm for unsourced multiple access," *Inf. Inference A, J. IMA*, vol. 9, no. 4, pp. 875–897, 2020.
[33] K. Andreev, P. Rybin, and A. Frolov, "Unsourced random access based on list recoverable codes correcting t errors," in *Proc. IEEE Inf. Theory Workshop (ITW)*, 2021, pp. 1–6.
[34] R. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 2, pp. 124–142, Mar. 1985.
[35] E. Onggosanusi, A. Gatherer, A. Dabak, and S. Hosur, "Performance analysis of closed-loop transmit diversity in the presence of feedback delay," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1618–1630, Sep. 2001.
[36] J. Y. N. Hui, "Fundamental issues of multiple accessing," Ph.D. dissertation, Lab. Inf. Decis. Syat., Mass. Inst. Technol., Cambridge, MA, USA, 1983.
[37] S. Verdu, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 35, no. 4, pp. 733–751, Jul. 1989.
[38] W. Zhang and L. Huang, "On OR many-access channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2638–2642.
[39] A. R. Cohen, J. A. Heller, and A. J. Viterbi, "A new coding technique for asynchronous multiple access communication," *IEEE Trans. Commun. Technol.*, vol. 19, no. 5, pp. 849–855, Oct. 1971.
[40] J. Robin and E. Erkip, "Capacity bounds and user identification costs in rayleigh-fading many-access channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Melbourne, VIC, Australia, Jul. 2021, pp. 2477–2482.
[41] B. Bloom, "Space/time tradeoffs in hash coding with allowable errors," *Commun. ACM*, vol. 13, no. 7, pp. 422–426, Jul. 1970.
[42] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[43] C. Chan, P. Che, S. Jaggi, and V. Saligrama, "Non-adaptive probabilistic group testing with noisy measurements: Near-optimal bounds with efficient algorithms," in *Proc. 49th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2011, pp. 1832–1839.

RUI DENG received the bachelor's degree in Internet of Things engineering from Anhui University, Hefei, China, in 2017. He is currently pursuing the Ph.D. degree with the University of Science and Technology of China, Hefei. His research interest includes random multiple access and nonlinear distortion in communication systems.

WENYI ZHANG (Senior Member, IEEE) received the bachelor's degree in automation from Tsinghua University, Beijing, China, in 2001, and the master's and Ph.D. degrees in electrical engineering from the University of Notre Dame, Notre Dame, IN, USA, in 2003 and 2006, respectively.

He is currently a Professor with the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China. His research interests include wireless communications, information theory, and signal detection/estimation. He was an Editor of IEEE COMMUNICATIONS LETTERS and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and is currently an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS.