IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 4 March 2024, accepted 28 March 2024, date of publication 8 April 2024, date of current version 15 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3386194

RESEARCH ARTICLE

Location-Based Unsourced Random Access

ROSHANAK SOLTANI[1], (Graduate Student Member, IEEE), DMITRI TRUHACHEV[1], (Member, IEEE), ALIREZA BAYESTEH[2], AND MONIROSHARIEH VAMEGHESTAHBANATI[2], (Member, IEEE)

[1]Department of Electrical and Computer Engineering, Dalhousie University, Halifax, NS B3H 4R2, Canada
[2]Huawei Canada, Ottawa, ON L3R 5A4, Canada

Corresponding author: Roshanak Soltani (roshanak.soltani@dal.ca)

ABSTRACT We propose a novel approach to Unsourced Random Access (URA) in which the encoding of each user's message is associated with the user's location. The presented reception algorithm is able to resolve message collisions and enhance the multiple user detection via estimation of the angles-of-arrival (AoA) of the users' signals. We demonstrate that the proposed technique allows to nearly double the number of supported active users in comparison to the state-of-the-art and reduces the required system's signal-to-noise ratio (SNR).

INDEX TERMS Multiple-input multiple-output (MIMO), unsourced random access (URA), location, angle-of-arrival (AoA).

I. INTRODUCTION

Grant-free multiple access [1], [2] that allows users to transmit data on demand, without obtaining a grant from the base station (BS), is becoming increasingly important for a variety of wireless applications. It is specifically useful for rapidly growing machine-generated traffic of short sporadic packets that needs to be organized in a low-latency and energy-efficient manner. Typically each packet contains a user identifier (ID) that associates a user with the transmitted information and often helps to decode the packet by identifying, for example, the permutation and spreading code used for payload encoding. For massive communications [3], on the other hand, the necessity to utilize and track user IDs complicates the receiver and takes away a fraction of the valuable data. Therefore, it is desirable that grant-free access is supplemented with a data transfer mechanism which does not require fixed user IDs.

Unsourced random access (URA) [4] has been proposed as an approach that enables the receiver to detect and decode grant-free packets without pre-assigned unique user IDs, where all users utilize identical codebooks. Unlike traditional multiple access methods, URA focuses on receiving the content of the messages and recovering a list of the packets that were transmitted rather than on identifying the transmitters that convey them. This makes URA an attractive approach for supporting massive low-latency communications with applications in the future sensor networks, machine-to-machine (M2M) communications, Internet-of-things (IoT) [5], etc.

In URA, the absence of pre-assigned unique IDs makes the detection of transmitted packets at the receiver more challenging compared to the conventional grant-free systems. A large variety of systems based on compressed sensing [6], [7], [8], [9], sub-slotted ALOHA style random access [10], preamble-payload techniques [11], [12], [13], [14] and chirp sequences [15] have been designed. The use of multiple receive and transmit antennas can improve the detection performance [16] by offering antenna diversity, combining gains, and beam-forming opportunities. Therefore, several classes of multiple-input multiple-output (MIMO) URA have been proposed.

A tensor-based approach to MIMO URA is presented in [17] and [18], where active users utilize rank-1 tensors to construct their messages. Employing tensors brings the opportunity to avoid transmitting pilots for channel estimation and allows the receiver to separate users' signals without multiple-user detection (MUD). However, this design is limited in terms of the number of supported active users and energy efficiency.

In the sub-block compressed-sensing (CS) approach proposed in [7] and extended to MIMO systems in [19], each active user splits its message into a few sub-blocks. The data bits of each sub-block are individually encoded

The associate editor coordinating the review of this manuscript and approving it for publication was Ronald Chang.

via a pseudo-random sensing matrix. An outer forward error-correction (FEC) code is used to establish a connection between the sub-blocks. At the receiver, the CS decoders retrieve the sets of sub-messages for each sub-block while FEC decoders stitch them together. A number of research works enhance the sub-block CS approach by exploiting wireless channel properties, where [20] proposes to utilize the channel information to assist the stitching decoder, while [21], [22] and their extension to near-field communications [23], remove the FEC code and the respective overhead and employ clustering of estimated channel information to stitch the sub-blocks for each user at the receiver. Although [20] and [21] have proposed more advanced decoders, they have only been tested in collision-free environments. Collision here means two or more users transmitting the same sequence from the sensing matrix as at least one of their message sub-blocks. More users can be accommodated in [22] which offers a collision resolution method. In [24], a two-phase sub-block URA system is presented, where the data of each user is initially split into two segments. The first segment serves as a preamble to assist the receiver in channel estimation, while in the second phase, the remaining data is transmitted following the conventional sub-block CS URA approach. Still, all the mentioned sub-block CS techniques have limitations in terms of the number of supported active users and energy efficiency, while their receiver performance relies significantly on the assumption of unequal channel gains.

Implementation of on-off division multiple access (ODMA) in URA [25], [26] has also received research attention recently. Each transmitter splits the message into several sub-blocks that are transmitted with a specific on-off pattern determined by a few bits selected from the message. While the receiver of [25] detects the on-off patterns solely from the received signal, [26] uses a preamble that defines the switching pattern. The ODMA URA solution performs well in medium to high signal-to-noise ratio (SNR) conditions. However, it requires longer transmission time slots compared to the conventional MIMO URA setups, which may not be ideal for low-latency communication.

The preamble-payload approach [27], [28] has, so far, been the most successful in supporting high numbers of active users. In this approach, each user forms a URA packet from a CS-encoded preamble and a payload, encoded for error correction. The preamble is chosen from a pool of preambles (that usually contains sequences constructed from randomly selected i.i.d. Gaussian or complex unitary elements) based on the first few bits of the data message. The remaining data bits are encoded by an error-correction encoder, permuted and spread to form the packet's payload. The permutation and spreading sequences are identified by the preamble [27]. The receiver performs CS-based preamble detection, to recover a set of active preambles, and estimates the respective users' channels. The obtained information is transferred to the MUD receiver that uses inverse permutation, de-spreading, and data estimation to decode the data payloads. In [29], a preamble-payload approach called slotted non-orthogonal pilot based unsourced random access (SNOP-URA) has been proposed. It breaks down the transmission frame into a few sub-slots and each active user randomly selects one sub-slot for transmission. At the receiver side, a generalized orthogonal matching pursuit algorithm is used for preamble detection and an iterative maximum-ratio combining-based MUD algorithm with channel estimation improvement is employed for payload decoding in each sub-slot.

A collision in preamble-payload URA occurs when two or more users randomly select the same preamble, which often results in a detection or decoding failure for the colliding users. In addition, significant numbers of collisions may hinder the detection and decoding process for all users' signals at the receiver. Therefore, a sufficiently large preamble pool is required to reduce the probability of collisions, which subsequently results in longer preambles for the CS decoder to function properly. This, in turn, leads to higher complexity of the preamble detection and reduces the length of the payload. Shorter payloads, on the other hand, can negatively impact the MUD performance and limit the ability of MIMO URA algorithm to support high numbers of active users.

Additionally, current MIMO URA approaches are often constrained by reliance on specific channel assumptions such as Rayleigh fading or presence of multiple paths with different path gains for user separation. Thus, designing a system which is more unpretentious in terms of channel environments and capable of supporting higher numbers of active users would be of interest and allow for efficient application to millimeter wave (mmWave) and Terahertz (THz) channels which are important for modern communication systems.

In this paper, we present a preamble-payload URA approach that exploits the user location geometry in order to support significantly higher numbers of active users compared to the state-of-the-art techniques. We link the preamble, permutation, and signature parameters of each user, which we call *transmission features*, to the user's angle-of-arrival (AoA) towards the receiving BS. This enables the multiple-antenna BS with the opportunity to recognize transmission features that were randomly selected by users based on both geography and the corresponding channel matrix. The proposed system is tested on a channel model with a single dominant path between each user and the BS, customary to low-scattering regions, and particularly useful in mmWave and THz environments [30].

Our contributions are as follows:

- We propose a method to establishing the connection between the AoA and transmission features by partitioning the area surrounding the BS into several sectors based on the AoA of the users' signals. Each sector is assigned a sub-pool of transmission features. This helps to ensure that signals from users with highly correlated channels in terms of AoA can still be separated at the receiver. The sub-pools are effectively re-used. This reduces the size of the overall pool of the transmission

features, the respective storage and, more importantly, the preamble length, which, in turn, decreases the complexity of the user activity detection.
- To exploit the strength of the proposed transmission strategy, we present a receiver architecture that integrates the AoA detection of active users with CS preamble decoding. This boosts the activity detection which now separates the users based on both the AoA and nearly uncorrelated random columns of the sensing matrix.

As a result, we demonstrate that the proposed technique can resolve collisions, improves detection performance and channel estimation, allows for use of shorter preambles, accommodates longer payloads. Further, we show that the proposed design enables MIMO URA systems to support considerably more active users compared to the state-of-the-art systems.

The paper is organized as follows. In Section II, we present the proposed transmitter and receiver structure and operation. In Section III we demonstrate the results of the system performance and its merits over the state-of-the-art. Section IV concludes the paper. In the rest of the paper, $(.)^*$, $(.)^T$ and $(.)^H$ indicate conjugate, transpose, and conjugate-transpose, respectively, and $\mathbf{I}_n$ is the identity matrix of size $n \times n$, and i is the imaginary unit.

II. SYSTEM MODEL

We focus on a time-slotted URA scenario, where at each time slot K_a single-antenna active users from the overall user pool of size $K_\mathrm{tot} \gg K_\mathrm{a}$ transmit to a BS receiver equipped with a uniform (horizontal) linear array consisting of M antennas. We consider equal-power users whose signals propagate to the BS receiver via a single dominant path.

An active user $k, k = 1, \cdots, K_\mathrm{a}$ encodes its information bit vector $\mathbf{v}_k$ of length B into a message $\mathbf{x}_k$ of length N_t channel uses. The composite $N_\mathrm{t} \times M$ received signal matrix at the BS is given by

$$\mathbf{Y} = \sum_{k=1}^{K_\mathrm{a}} \mathbf{x}_k \mathbf{h}_k^T + \mathbf{W}, \tag{1}$$

where $\mathbf{x}_k \in \mathbb{C}^{N_\mathrm{t}\times 1}$ is the transmitted signal of the k-th user, and $\mathbf{W} \in \mathbb{C}^{N_\mathrm{t}\times M}$ is the matrix of the additive white Gaussian noise (AWGN). The complex vector of the channel coefficients between the k-th user and all the BS antennas is given by $\mathbf{h}_k = \beta_k \mathbf{a}(\theta_k) \in \mathbb{C}^{M\times 1}$, in which $\beta_k = \exp(i\alpha_k)$, where α_k is the initial signal phase, and θ_k is the AoA of the signal received from the k-th user. The vector $\mathbf{a}(\theta_k) = [a_1(\theta_k), \cdots, a_M(\theta_k)]^T$ of length M is the antenna array response vector at the BS receiver, where

$$a_m(\theta) = \exp(i2\pi(d/\lambda_w)(m-1)\sin(\theta))\ m = 1, \cdots, M, \tag{2}$$

d is the distance between the antenna array elements at the BS, and λ_w is the wavelength. This channel model can be regarded as a generalized representation of low-scattering environment models that are commonly employed for high-frequency technologies [30].

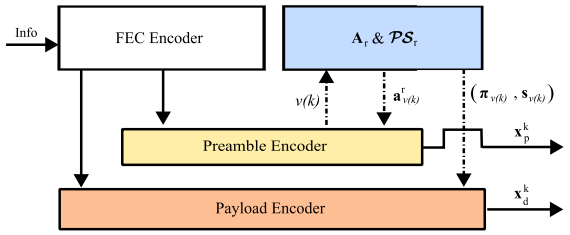

FIGURE 1. Block diagram of the k-th user's transmitter.

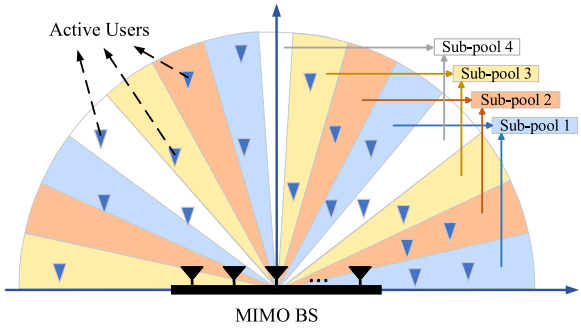

FIGURE 2. Space model and parameter sub-pools.

Let us denote the set of active user messages by $\mathcal{L} = \{\mathbf{v}_k\}_{k=1,\cdots,K_\mathrm{a}}$. Following the receiver processing the BS generates an estimated message set $\hat{\mathcal{L}}$. The system's performance is evaluated by measuring the per-user probability of error (PUPE) [4], which is defined by

$$P_\mathrm{e} = \frac{1}{K_\mathrm{a}} \sum_{k=1}^{K_\mathrm{a}} \Pr(\mathbf{v}_k \notin \hat{\mathcal{L}}). \tag{3}$$

Hence, the PUPE calculates the average fraction of messages that are decoded incorrectly. The focus of URA is on decoding the majority of the data messages correctly rather than targeting simultaneous low-error communications on all individual users' links. It is customary in URA literature to design systems capable of achieving a target PUPE in the order of $0.05 - 0.1$.

A. TRANSMITTER

Each active user forms its packet as illustrated in Fig. 1. The B-bit information message of the k-th active user is first encoded by a binary FEC of length n. The encoded vector is split into two parts. Vector $\mathbf{v}_k^p$ that consists of the first B_p bits is used to select the packet's preamble. The binary sequence $\mathbf{v}_k^\mathrm{p}$ is converted into a decimal index $\nu(k)$ pointing to a column $\mathbf{a}^r_{\nu(k)}$ of the $N_\mathrm{p} \times 2^{B_\mathrm{p}}$ sensing matrix $\mathbf{A}_r$. We consider sensing matrices with random, uniform, i.i.d. circularly-symmetric complex Gaussian entries. The index r of the sensing matrix is associated with the user's location as follows. The space around the receiving BS is split into S sectors based on the AoA as illustrated in Fig. 2. A preamble sub-pool is associated with each sector and the sub-pools are re-used after R sectors as illustrated in the figure (for $R = 4$). The sub-pool

of preambles associated with s-th sector, $s = 1, 2, \cdots, S$ is represented by the columns of the sensing matrix $\mathbf{A}_r$, where $r = s \bmod R$. This gives $N = R2^{B_\mathrm{p}}$ preambles overall. A set of active users (possibly from multiple sectors) that randomly select the same preamble is called a group.

Each user selects a preamble from the pool of its sector and, therefore, needs to know its location. This can be accomplished practically since most of the users are stationary and, in addition, a user and the BS have multiple means to learn and predict the user's location. The sectorization effectively reduces the probability of nearby users to collide by selecting the same preamble. On the contrary, the colliding users in one group are typically spread apart to distinct sectors and can be separated by means of angle detection. Finally, we note that the random choice of the preamble elements ensures that the preambles in the neighboring sectors are semi-orthogonal, thus, reducing the probability of erroneous angle detection.

Further Fig. 1 illustrates that the payload encoder is provided with a pair of signature and permutation patterns $(\pi_{\nu(k)}, \mathbf{s}_{\nu(k)})$ which are selected using the index $\nu(k)$ from the signature-permutation sub-pool $\mathcal{PS}_r$ associated with the user's sector s ($r = s \bmod R$). The payload encoder first applies quadrature phase-shift keying (QPSK) on the remaining $n - B_\mathrm{p} = B_\mathrm{d}$ bits of the binary message. Then each symbol is repeated Q times, and the symbols are permuted with $\pi_{\nu_r(k)}$. Following that, the sequence is scrambled via symbol-by-symbol multiplication with $\mathbf{s}_{\nu(k)}$, which is a vector of i.i.d. unitary complex numbers with uniform phases. This results in the encoded payload sequence $\mathbf{x}_k^\mathrm{d}$ of length N_d.

We note that the proposed design differs from the conventional URA, where the k-th user's packet is solely a function of its data ($\mathbf{x}_k = f(\mathbf{v}_k)$). In our case, the packet encoding is also dependent on the user's AoA θ_k, i.e., $\mathbf{x}_k = f(\mathbf{v}_k, \theta_k)$.

B. RECEIVER

The receiver schematic is presented in Fig. 3. Initially, the preamble detection block performs the detection and differentiation of the active preambles. While this task can be acomplished via different CS-based activity detection (AD) algorithms such as generalized subspace pursuit (GSP) [31], we choose to employ the multi-measurement vector - approximate message passing (MMV-AMP) algorithm [32] since it have been shown in [27] that MMV-AMP outperforms GSP algorithm and is of lower complexity. As a byproduct, the algorithm estimates the active preamble indices from the preamble pool and the corresponding channel coefficients. In cases where multiple users transmit the same preamble, the MMV-AMP estimates the sum of the colliding users' channels.

After detecting the indices of the active preambles and the corresponding channel coefficients related to each group, an angle differentiation and channel estimation algorithm is employed to estimate the AoAs and consequently channels

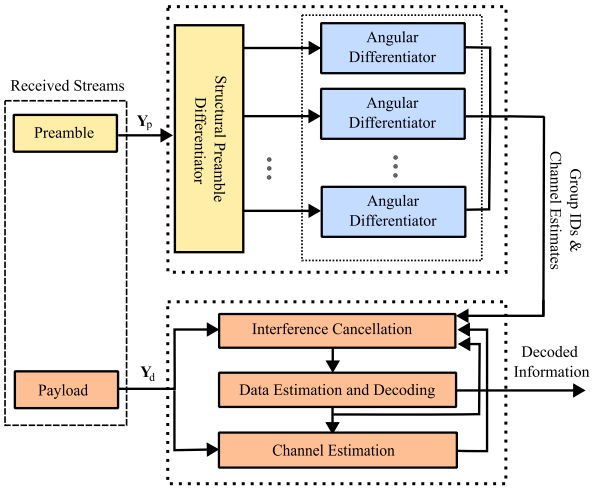

FIGURE 3. Block-diagram of the BS receiver.

of users who share the same preamble (users in one group). The designed algorithm is based on the multiple signal classification (MUSIC) [33], which is a subspace-based method to detect the angles of signal arrival. In general, MUSIC provides high-resolution AoA detection and is less complex than most of its counterparts. We note that other subspace-based approaches like the estimation of signal parameters via rotational invariance techniques (ESPRIT) [34] or expectation maximization (EM) methods [35] can also be applied for angle differentiation. Based on the estimated AoAs we estimate the initial phases of each user's channel coefficients by applying the linear minimum mean square error (LMMSE) estimator. By combining the estimated AoA and initial phase together we complete the channel estimation.

The estimated active indices and channel coefficients are passed to the MUD block of the receiver. At this stage, we use the estimated channel coefficients to apply LMMSE on the received signal formed by the payload portions of the message and obtain the initial estimates of the transmitted payload of each user. Those estimates are passed through an iterative process of signal demodulation, FEC decoding, detection of false users, re-encoding of the payloads, interference cancellation, and channel re-estimation. We have improved the MUD structure proposed in [27] by incorporating the LMMSE into our design. The detailed description of the receiver algorithm on both the preamble and payload parts is as follows.

1) ACTIVITY DETECTION, ANGLE & CHANNEL ESTIMATION

The received signal of the preamble parts of the packets is given by

$$\mathbf{Y}_\mathrm{p} = \mathbf{A}\boldsymbol{\Gamma}\mathbf{H}_\mathrm{p} + \mathbf{W}_\mathrm{p} \in \mathbb{C}^{N_\mathrm{p} \times M}, \tag{4}$$

where $\mathbf{A} = [\mathbf{A}_0, \cdots, \mathbf{A}_{R-1}]$ of size $N_\mathrm{p} \times N$ is the concatenation of all R sensing matrices representing the user sectors, accounting for reuse. $\boldsymbol{\Gamma} = \mathrm{diag}\{\gamma_1, \gamma_2, \cdots, \gamma_N\}$ is the diagonal group activity matrix being estimated.

The non-zero diagonal entries of $\mathbf{\Gamma}$ are the estimated indices of the preambles have been transmitted by users from the entire pool $\mathbf{A}$. Those satisfy $\gamma_g \in \{0, n_g\}$, where n_g is the number of users in group g, $g = 1, \cdots, N$, later estimated by Algorithm 1. Here we recall that a "group" represents the users which share the same preamble. By $\mathcal{K}_g$ we denote the indices of the users which fall into the g-th group. The overall channel matrix equals $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_{K_a}]^T \in \mathbb{C}^{K_a \times M}$ in which, $\mathbf{h}_k$ is the vector of channel coefficients of the k-th user, $k = 1, \cdots, K_a$. The rows of $\mathbf{H}$ form the sparse preamble channel coefficient matrix $\mathbf{H}_p = [\mathbf{h}_1^p, \cdots, \mathbf{h}_N^p]^T \in \mathbb{C}^{N \times M}$, where $\mathbf{h}_g^p = \sum_{\kappa \in \mathcal{K}_g} \mathbf{h}_\kappa^T$ (and $\mathbf{h}_g^p = 0$ in case $n_g = 0$, i.e., when group g is not active) and $\mathbf{W}_p \in \mathbb{C}^{N_p \times M}$ is the AWGN.

For activity detection, we first apply MMV-AMP [28], [32] to (4) to estimate $\mathcal{Y} = \mathbf{\Gamma}\mathbf{H}_p$. At the l-th iteration $l = 1, \cdots, L$ of the MMV-AMP we have

$$\tau_{l+1,m}^2 = \frac{||\mathbf{Z}_{:,m}^l||^2}{N_p} \quad m = 1, \cdots, M \,, \tag{5}$$

$$\mathcal{Y}^{l+1} = \eta(\mathbf{A}^H \mathbf{Z}^l + \mathcal{Y}^l, \hat{\gamma}^l, \tau_{l+1}) \,, \tag{6}$$

$$\mathbf{Z}^{l+1} = \mathbf{Y}_p - \mathbf{A}\mathcal{Y}^{l+1} + N/N_p \mathbf{Z}^l \langle \eta'(\mathbf{A}^H \mathbf{Z}^l + \mathcal{Y}^l, \hat{\gamma}^l, \tau_{l+1}) \rangle \,, \tag{7}$$

where $\mathbf{Z}^l \in \mathbb{C}^{N_p \times M}$ is the residual noise-and-interference at iteration l, and, initially, $\mathbf{Z}^0 = \mathbf{Y}_p$. The vector τ_l contains the estimates of the noise-and-interference power at the l-th iteration. The denoiser function $\eta(\tilde{\mathcal{Y}}^l, \hat{\gamma}, \tau)$ is applied in a row-by-row fashion to the rows of $\tilde{\mathcal{Y}}^l = \mathbf{A}^H \mathbf{Z}^l + \mathcal{Y}^l$ that are denoted by $\tilde{\mathfrak{y}}_{k'}^l$, $k' = 1, \cdots, N$ so that $\eta(\tilde{\mathcal{Y}}^l, \hat{\gamma}, \tau) = [\eta(\tilde{\mathfrak{y}}_1^l, \hat{\gamma}_1^l, \tau), \cdots, \eta(\tilde{\mathfrak{y}}_N^l, \hat{\gamma}_N^l, \tau)]^T$. Moreover, $\eta'(\cdot)$ refers to the Jacobi matrix, and $\langle \cdot \rangle$ is the averaging operator. Here we use the LMMSE denoiser function which is given by

$$\eta(\tilde{\mathfrak{y}}, \hat{\gamma}, \tau) = \phi(\tilde{\mathfrak{y}}, \hat{\gamma}, \tau)\hat{\gamma}(\hat{\gamma}\mathbf{I}_M + \mathrm{diag}(\tau^2))^{-1}\tilde{\mathfrak{y}}, \tag{8}$$

where $\phi(\tilde{\mathfrak{y}}_{k'}^l, \hat{\gamma}_{k'}^l, \tau_l) \in [0, 1]$ is the estimated probability that k'-th column of the matrix $\mathbf{A}$ was selected as a preamble by at least one active user. If the number of users transmitting a particular active preamble is large enough, the distribution of active rows of $\mathcal{Y}$ can be approximated by a Gaussian distribution according to the central limit theorem. Therefore, here for simplicity we consider a Gaussian distribution for the rows and as a result, $\phi(\tilde{\mathfrak{y}}, \hat{\gamma}, \tau)$ is computed via

$$\phi(\tilde{\mathfrak{y}}, \hat{\gamma}, \tau) = \left[1 + \frac{1-\rho}{\rho} \prod_{m=1}^{M} \mathrm{e}^{-\frac{\hat{\gamma}|\tilde{\mathfrak{y}}_m|^2}{\tau_m^2(\hat{\gamma}+\tau_m^2)}} \frac{(\hat{\gamma}+\tau_m^2)}{\tau_m^2}\right]^{-1}, \tag{9}$$

where ρ is the prior probability of the user activity, i.e., $\rho = K_a/N$. Although the algorithm is not sensitive to the knowledge of ρ, the number of active users K_a can be estimated using the technique presented in [12].

The maximum likelihood (ML) activity vector at the l-th iteration is estimated [28] via

$$\hat{\gamma}^l = \max\left(\frac{||\tilde{\mathfrak{y}}_{k'}^l||_2^2}{M} - \frac{\sum_{m=1}^M \tau_m}{M}, \tilde{\gamma}\right), \tag{10}$$

where the value of $\tilde{\gamma}$ is selected to prevent the algorithm's performance from being negatively affected by overly small estimates. Empirically we set $\tilde{\gamma} = 1$.

At the final L-th iteration, the probability of the k'-th column's activity, $k' = 1, 2, \cdots, N$ is set to

$$\hat{\gamma}_{k'} = \phi(\mathfrak{y}_{k'}^L, \hat{\gamma}_{k'}^L, \tau_L)\hat{\gamma}_{k'}^L \,. \tag{11}$$

The values in (11) are thresholded using $\beta = b \times \left(\frac{1}{M}\sum_{m=1}^M \tau_{l+1,m}\right)$, with b set empirically at 0.45 to choose the final active column indices. The output of the MMV-AMP algorithm is a set of detected transmitted preambles or, equivalently, the set of indices g of groups with at least one active user. The overall number of identified active groups is denoted by $\hat{G}$. In addition to the detected preambles, the algorithm also provides an estimate of the sum of the channel coefficients of each user group, represented by the vector $\mathfrak{y}_{g'}^L$, where $g' \in \{\hat{\mathcal{G}}\}$.

Following the group activity detection we apply an angle-differentiator algorithm presented by Algorithm 1 to separate the users within each group and obtain the estimates of their channels.

Algorithm 1 Angle Differentiation and Channel Estimation in Each Group

for $g' \in \{\hat{\mathcal{G}}\}$ **do**
  ▷ Begin MUSIC Algorithm:
  $\mathbf{R}_{g'} = \mathfrak{y}_{g'}^H \mathfrak{y}_{g'}$
  $\mathrm{SVD}(\mathbf{R}_{g'}) = \mathbf{U}_{g'}\mathbf{S}_{g'}\mathbf{V}_{g'}^T$
  $\tilde{\mathbf{U}}_{g'} = M - 1$ right-most columns of $\mathbf{U}_{g'}$
  for $\theta = 0, \Delta\theta, 2\Delta\theta, \cdots, (N_{\text{angle}} - 1)\Delta\theta$ **do**
    ▷ MUSIC spectrum:
    $P(\theta) = (\mathbf{a}(\theta)^H \tilde{\mathbf{U}}_{g'} \tilde{\mathbf{U}}_{g'}^H \mathbf{a}(\theta))^{-1}$.
  end for
  ▷ Potential AoAs of the users in group g':
  Find $\hat{\theta}_1, \cdots, \hat{\theta}_{\hat{n}_{g'}^0}$ related to $P(\theta)$'s local maximums.
  ▷ End MUSIC Algorithm.
  ▷ Beamforming matrix:
  $\mathbf{B}_{g'} = [\mathbf{a}(\hat{\theta}_1), \cdots, \mathbf{a}(\hat{\theta}_{\hat{n}_{g'}^0})]^T \in \mathbb{C}^{\hat{n}_{g'}^0 \times M}$
  ▷ Use LMMSE to estimate the potential initial phases:
  $\hat{\boldsymbol{\beta}}_{g'} = \mathfrak{y}_{g'}\mathbf{B}_{g'}^H(\mathbf{B}_{g'}\mathbf{B}_{g'}^H + \sigma^2\mathbf{I}_{\hat{n}_{g'}^0})^{-1} \in \mathbb{C}^{1 \times \hat{n}_{g'}^0}$
  ▷ Apply the threshold t to remove possible noise spikes:
  Select $k_{g'}$ such that $|\hat{\beta}_{k_{g'}}| > t$, $k_{g'} = 1, \cdots, \hat{n}_{g'}$, then compute $\hat{\alpha}_{k_{g'}} = \mathrm{angle}(\beta_{k_{g'}})$.
  ▷ Channel estimates for $\hat{n}_{g'}$ active detected users in group g':
  $\hat{\mathbf{h}}_{k_{g'}} = \exp(j\hat{\alpha}_{k_{g'}})\mathbf{a}(\hat{\theta}_{k_{g'}}), \quad k_{g'} = 1, \cdots, \hat{n}_{g'}$
  Output $\hat{n}_{g'}$ and $\hat{\mathbf{h}}_{k_{g'}}$.
end for

Algorithm 1 is applied to the estimated coefficients $\mathfrak{y}_{g'}^L$ of each group $g' \in \{\hat{\mathcal{G}}\}$. Steps 2-8 perform the MUSIC [33]

to find AoAs of the active users in group g'. Note that the SVD in Step 3 refers to the singular value decomposition. In Step 4, the left-most column of $\mathbf{U}_{g'}$ is the signal subspace represented by the active preamble common to all users in the group. The remaining $M-1$ columns of $\mathbf{U}_{g'}$ form the noise subspace, used to generate the MUSIC spectrum $P(\theta)$ at Step 6. Note that we compute it for the angles $\theta \in \{0, \Delta\theta, 2\Delta\theta, \cdots, 180^o\}$, where $\Delta\theta = 180^o/N_{\text{angle}}$ and N_{angle} is the angle resolution of the algorithm. The $\hat{n}^0_{g'}$ local maximums (peaks) of the spectrum indicate the AoAs of the potential active users in group g' are obtained in Step 8.

The initial AoA estimates obtained in Step 8 are inserted into the antenna array response vector $\mathbf{a}(\theta)$ to create a beamforming matrix $\mathbf{B}_{g'}$ in Step 9. In Step 10 we utilize $\mathbf{B}_{g'}$ to perform LMMSE to compute the initial channel phases $\hat{\alpha}_{k_{g'}}$ for each AoA estimate. In Step 11 we apply a threshold t, empirically set to 0.2, to identify and remove the peaks of the phase estimates that are likely to be noise spikes. The thresholding also helps to discard the groups that are falsely detected by MMV-AMP and enhance the activity detection performance. Ultimately, the output includes the estimated number of the active users in group g' and the complete channel estimates for each user in the group, combining the angle and the initial phase estimates.

2) MULTIPLE USER DETECTION OF THE PAYLOADS

The received signal of the payloads of the active user packets is given by

$$\mathbf{Y}_{\text{d}} = \sum_{k=1}^{K_{\text{a}}} \mathbf{x}^{\text{d}}_k \mathbf{h}^T_k + \mathbf{W}^k_{\text{d}} = \mathbf{X}_{\text{d}}\mathbf{H} + \mathbf{W}_{\text{d}} \quad \in \mathbb{C}^{N_{\text{d}}\times M}, \tag{12}$$

where $\mathbf{X}_{\text{d}} = [\mathbf{x}^{\text{d}}_1, \mathbf{x}^{\text{d}}_2, \cdots, \mathbf{x}^{\text{d}}_{K_{\text{a}}}] \in \mathbb{C}^{N_{\text{d}}\times K_{\text{a}}}$ in which $\mathbf{x}^k_{\text{d}}$ is the column data vector of the k-th active user and $\mathbf{W}_{\text{d}} \in \mathbb{C}^{N_{\text{d}}\times M}$ is the AWGN. For the k-th active user we have $\mathbf{Y}^{\text{d}}_k = \mathbf{x}^{\text{d}}_k \mathbf{h}^T_k + \mathbf{W}^{\text{d}}_k$. The LMMSE filter based on the channels estimated by the MMV-AMP for the preamble part is applied to (12) and results in

$$\mathbf{R}^0 = (\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \sigma^2\mathbf{I}_{K_{\text{a}}})^{-1}\hat{\mathbf{H}}^*\mathbf{Y}^T_{\text{d}} \in \mathbb{C}^{K_{\text{a}}\times N_{\text{d}}}. \tag{13}$$

The k-th row of $\mathbf{R}_0$ is denoted by $\mathbf{r}^0_k$, where $\mathbf{r}^{0^T}_k = \mathbf{x}^d_k + \boldsymbol{\zeta}_k$, and $\boldsymbol{\zeta}_k$ is the remaining interference from the other users' channels and the noise faced by the k-th active user. At the l-th MUD iteration, $l = 1, \cdots, L_{\text{MUD}}$, the estimated message $\mathbf{r}^{l-1}_k$ of the k-th active user consists of $r^{l-1}_{k,j,q}$, $q = 1, \cdots, Q$, $j = 1, \cdots, B_{\text{d}}/2$ in which

$$r^{l-1}_{k,j,q} = x_{k,j} + \zeta^{l-1}_{k,j,q}. \tag{14}$$

Since one QPSK symbol $x_{k,j}$, is made up of two data bits, the log-likelihood ratios (LLR)s can be obtained using the real and imaginary parts of $\lambda^l_{k,j,q} = 2r^{l-1}_{k,j,q}s^*_{k,j,q}/\mathbb{E}|\zeta^{l-1}_{k,j,q}|^2$. The LLRs of the QPSK symbols are combined over Q symbol replicas

$$\mu^t_{k,j} = \sum_{q'=1}^{Q} \lambda^l_{k,j,q'} \tag{15}$$

and passed to the FEC decoder which is activated only if the current estimated SINR is high enough to ensure decoding improvement. The FEC decoder outputs the estimated data bits $\tilde{x}^l_{k,j}$. When FEC code includes cyclic-redundancy check (CRC), the latter is used to exclude correctly decoded, i.e., CRC-checked, packets from the subsequent MUD iterations via the following interference cancellation. When the SINR is low and FEC decoding is not warranted, maximum likelihood estimates of the binary data are obtained

$$\begin{aligned}\tilde{x}^l_{k,j,q} = [&\tanh(\text{Re}(\mu^l_{k,j} - \lambda^l_{k,j,q})) \\ &+ i\tanh(\text{Im}(\mu^l_{k,j} - \lambda^l_{k,j,q}))]s_{k,j,q}.\end{aligned} \tag{16}$$

The data estimates $\tilde{x}^l_{k,j,q}$ undergo the permutation and scrambling process to reconstruct the original signals transmitted by the users, denoted by $\tilde{\mathbf{x}}^{\text{d}}_{l,k}$. We then calculate the interference cancellation signals via

$$\mathbf{r}^l_k = \hat{\mathbf{h}}^H_k \big(\mathbf{Y}_{\text{d}} - \sum_{k'\in\hat{K}_a, k'\neq k} \tilde{\mathbf{x}}^{\text{d}}_{l,k'}\hat{\mathbf{h}}_{k'}\big)^T \tag{17}$$

and estimate the remaining noise and interference power to re-compute the LLRs using

$$\mathbb{E}\big[|\zeta^l_{k,j,q}|^2\big] = \mathbb{E}\Big[|\hat{\mathbf{h}}^H_k \big(\mathbf{Y}_{\text{d}} - \sum_{k'\in\hat{K}_a, k'\neq k} \tilde{\mathbf{x}}^{\text{d}}_{l,k'}\hat{\mathbf{h}}_{k'}\big)^T|^2\Big]. \tag{18}$$

The system will then proceed to re-estimate the channel coefficients. Further, the canceled signal $\mathbf{r}^l_k$ is utilized to obtain the updated LLRs (15) at the $(l+1)$-th iteration and the iterative process continues. The entire MUD process is repeated for L_{MUD} iterations.

Utilization of the LMMSE filter before the interference-cancellation iterations leads to enhanced decoding performance, compared to the matched-filter type MUD approach of [27]. The LMMSE is used only once to reduce the complexity and for the consecutive iterations we use matched-filter instead of (13) (see (17)). However, it provides an important tool to resolve the collisions within the groups given the fact that the AoA-related preamble structure is capable of supporting large numbers of users and operating in heavily user-overloaded environments.

C. ALGORITHM COMPLEXITY

For the proposed algorithm, the complexity of the preamble detection, based on MMV-AMP (see (5)-(7)), in terms of multiplications, scales as $\mathcal{O}(NN_pM)$ operations per iteration. The complexity of the angle differentiation and channel estimation presented in Algorithm 1 is as follows. Based on the standard MUSIC algorithm, the complexity of the matrix multiplication in the covariance calculation (Step 2) scales as $\mathcal{O}(M^2)$, the SVD calculation (Step 3) scales as $\mathcal{O}(M^3)$, the MUSIC spectrum calculation (Steps 5-7) and search for the peaks of the spectrum (Step 8) scales as $\mathcal{O}(M^3 + N_{\text{angle}})$, where $N_{\text{angle}} = 180$ is the angle resolution. Finally, the complexity of MMSE calculation in Step 10 scale as $\mathcal{O}(n^0_g M^2 + n^{0^3}_g + n^0_g M)$. This results in the overall

complexity of $\mathcal{O}(\hat{G}(M^3 + N_{\text{angle}} + n_g^0 M^2))$ of the AoA detection and channel estimation. The complexity of LMMSE filtering at the initial phase of MUD scales as $\mathcal{O}(K_a^3)$. The MUD complexity per iteration, accounting for the interference cancellation and bit estimation, de-interleaving and de-spreading, polar and Hamming error-correction decoding (see Section II-B2) scales as $\mathcal{O}(K_a M N_d)$, $\mathcal{O}(K_a N_d)$, $\mathcal{O}(K_a N_c \log N_c)$ and $\mathcal{O}(K_a N_c)$, respectively, where N_c is the FEC length [27].

We note that while the abovementioned complexity of the preamble and payload detection is common to that of the state-of-the-art URA systems [27], [28], [36], the extra complexity brought to the system by the angle resolution is just in the order of the complexity of a single MUD iteration (see Section III). Therefore, it adds just a small extra fraction to the overall complexity.

III. NUMERICAL RESULTS

We consider a system with the total frame length of $N_t = 3200$ channel uses, where each packet carries $B = 100$ information bits. These parameters are customary for comparing MIMO URA systems in the literature [17], [27], [28]. We assume that the AoAs of the users are uniformly distributed in $[-90^o, 90^o]$ and $d = \lambda_w/4$, $\lambda_w = c/f$, where c is the speed of light and $f = 28$ GHz (mmWave frequency). The number of iterations used for MMV-AMP and MUD in the proposed design are $L = 8$ and $L_{\text{MUD}} = 10$, respectively. In the presented result figures, the SNR E_b/N_0 required for achieving a certain PUPE target is plotted against the number of supported active users K_a. For each K_a the transmit energies per bit for the preamble and the payload parts are numerically optimized to minimize the overall E_b/N_0. Through numerous simulations conducted on the proposed design, it was determined that optimal system performance coupled with resource efficiency is achieved by reusing the pool after each 12^o segment of space.

The supported number of active users and the respective SNRs for the proposed model are presented in Fig. 4 and Fig. 5 for the target PUPE equal to 0.1 and 0.03, respectively. In each figure, the green and red bundles of curves correspond to the total pool size of $N = 2^{11}$ and $N = 2^{13}$, respectively. System performance in terms of minimum SNR needed to support various numbers of active users with $R = 2$ sub-pools (dashed curves with right triangles), $R = 4$ sub-pools (solid curves with squares), and $R = 8$ sub-pools (dot-dashed curves with triangles) has been evaluated.

For a system with R sub-pools each 12^o segment is divided into R sectors, resulting in the sector sizes of $\frac{12^o}{R}$. For instance, in $R = 4$ scenario, we have 60 sectors of size 3^o. The sub-pools are reused after each four consecutive sectors. For each scenario, we performed numerical experiments to select the preamble and payload lengths (N_p and N_d) that provide the best performance. The number of repetitions Q for encoding of the payload considering QPSK modulation is determined accordingly i.e., $Q \approx 2N_d/B_d$. The double-extended Hamming $(128, 119)$ code defined in [37] shortened to

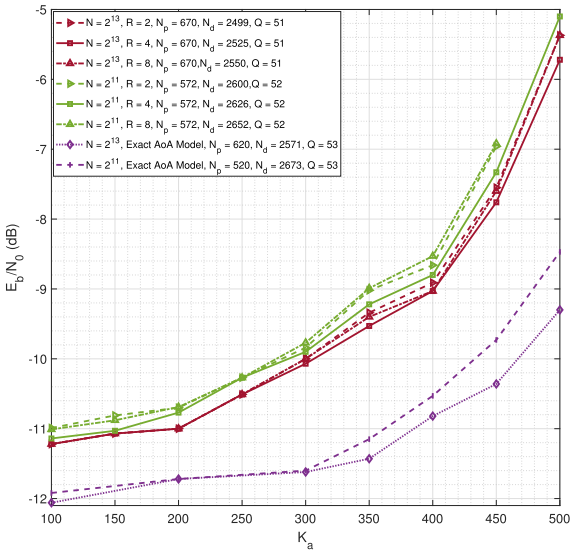

FIGURE 4. Minimum SNR required to support various numbers of active users. The target PUPE equals 0.1.

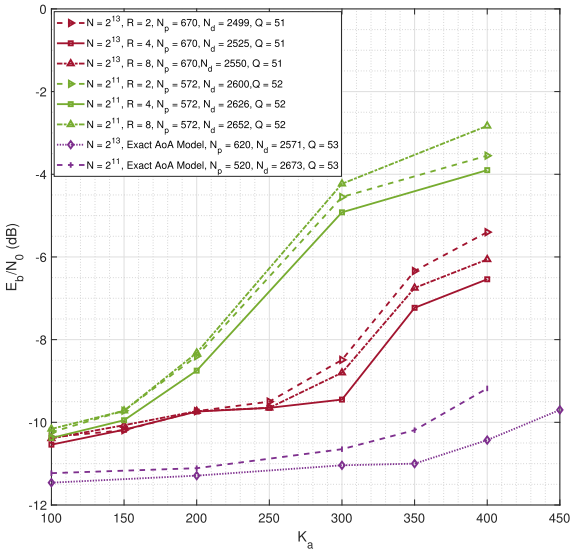

FIGURE 5. Minimum SNR required to support various numbers of active users. The target PUPE equals 0.03.

(110, 101) has been used as the FEC code. As we can see, the design with $R = 4$ sub-pools slightly outperforms the other two choices.

The magenta curves labeled "exact AoA model" in the legend corresponds to a theoretic model in which we assume that a specific preamble-signature-permutation is associated with each exact AoA towards BS. In other words, each active user knows its location exactly and is given a specific set of transmission features. The performance of such a model provides a lower bound on the performance achievable with the sectorized model. The performance for the "exact AoA model" is for the total pool size of $N = 2^{11}$ (dashed purple curve) and $N = 2^{13}$ (dotted purple curve) in Fig. 4 and 5. The optimal lengths for both the preamble and payload were selected for the simulations.

TABLE 1. Minimum SNR values (in dB) required to support K_a active users. Three different preamble/payload length compositions are considered. The overall pool size equals $N = 2^{13}$ and comprises $R = 4$ sub-pools.

Setting			K_a			
N_p	N_d	Q	100	200	300	400
570	2624	53	−11.00	−10.73	−9.79	−8.60
670	2525	51	−11.22	−11.00	−10.07	−9.03
770	2426	49	−11.15	−10.85	−10.00	−8.88

Table 1 provides a comparison of the system performance for several compositions of preamble lengths N_p and payload lengths N_d. The pool size of $N = 2^{13}$ with $R = 4$ sub-pools is considered. Each user maps its initial $B_p = N/R = 11$ bits to a preamble of size N_p, while the remaining packet symbols, $N_d = N_t - N_p$, are utilized for (110, 101) Hamming encoded payload transmission. The number of repetitions Q is chosen accordingly, i.e, $Q \approx 2N_d/B_d$. Three scenarios are compared: $N_p = 570$, $N_d = 2624$, $Q = 53$; $N_p = 670$, $N_d = 2525$, $Q = 51$; and $N_p = 770$, $N_d = 2426$, $Q = 49$. The table presents the smallest E_b/N_0 values (in dB) required to achieve PUPE equal to 0.1 and support a certain number of active users K_a. Although the differences are not substantial, it is noticeable that the middle scenario slightly outperforms the other two. It is worth mentioning that the first scenario with $N_p = 570$ does not accommodate $K_a = 500$ users. This observation arises from the fact that for the proper functioning of channel estimation and activity detection in the AMP algorithm, N_p needs to be sufficiently larger than K_a.

Fig. 6 presents a comparison between the proposed design and the state-of-the-art URA systems. Based on the previous results, we chose a pool size of $N = 2^{13}$ with $R = 4$ sub-pools. Each packet starts with a $N_p = 670$ symbol preamble, selected based on the first $B_p = N/R = 11$ information bits of the message as well as the user's AoA. The payload occupies the remaining $N_d = 2530$ channel uses and the number of symbol repetitions Q is determined accordingly.

The results are shown for two types of FEC codes, the double-extended Hamming $(128, 119)$ code shortened to $(110, 101)$ (solid red curve with squares) and the 5G $(512, 100)$ polar code with 11-bit CRC (dash-dotted red curve with squares).

We conducted a comparison of the proposed system with three state-of-the-art MIMO URA systems with preamble-payload structure. All results are obtained on the dominant-path channel model given in Section II and the target PUPE equals 0.1. In the system proposed in [27] (solid and dash-dotted yellow curves with stars for the Hamming and Polar codes, respectively), the payload is formed by repetition, scrambling, and permutation of the data symbols, similar to the design proposed in this work. At the receiver side, MMV-AMP is applied for preamble detection. It is followed by iterative data estimation, decoding, and channel estimation refinement throughout the MUD process. For simulations of the system according to [27] we chose pool size $N = 2^{16}$, preamble length $N_p = 1152$, payload length $N_d = 2021$ with $Q = 43$, for the encoding with

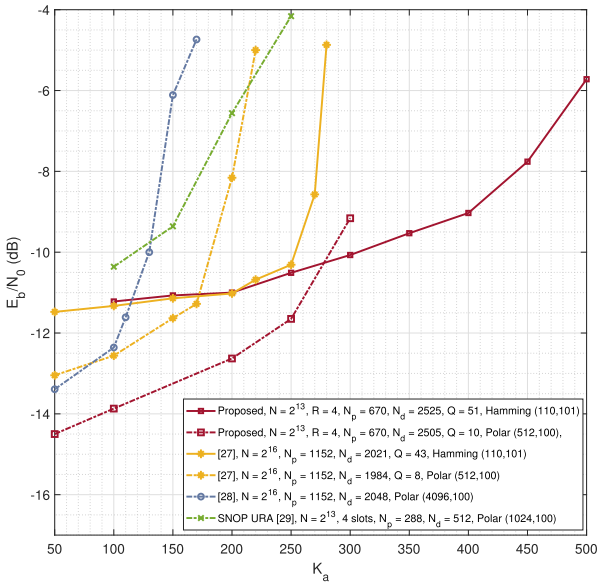

FIGURE 6. Minimum SNR required to support various numbers of active users, compared to the state-of-the-art. The target PUPE is set to 0.1.

the above-mentioned $(110, 101)$ Hamming code. The second case uses $(512, 100)$ Polar encoding for the payload size $N_d = 1984$ with $Q = 8$ repetitions. Compared to [27], the model proposed in [28] (dash-dotted blue-grey curve with circles) uses a lower rate Polar code $(4096, 100)$ for the payload, while repetition, signature, and permutation are not applied. Similarly, MMV-AMP is used for activity detection and channel estimation at the receiver. Maximum-ratio-combining and decoding are used for the MUD part. Based on [28] we chose $N = 2^{16}$, $N_p = 1152$, and $N_d = 2048$ as simulation parameters. Finally, the SNOP URA [29] (dash-dotted green curve with crosses) breaks down the transmission frame into a few sub-slots and each active user randomly selects one sub-slot for transmission. We chose four sub-slots for our simulations since SNOP URA with four slots achieves its best performance [29]. At the receiver side, a generalized orthogonal matching pursuit algorithm is used for preamble detection and an iterative maximum-ratio combining-based MUD algorithm with channel estimation improvement is employed for payload decoding in each sub-slot. For the simulations, we chose $N = 2^{13}$ $N_p = 288$, and $N_d = 512$ with $(1024, 100)$ Polar code as determined to be the most effective according to [29]. It should be highlighted that for the proposed URA model and SNOP URA [29], we considered the total pool of size $N = 2^{13}$ which has been demonstrated to be effective for these systems according to their shorter preambles and collision resolution mechanisms. For [27] and [28], we utilized their standard total pool size of $N = 2^{16}$, which is necessary for mitigating collisions in their respective systems.

The results demonstrate that the proposed system requires lower SNR to support the same numbers of users compared to the state-of-the-art approaches. Moreover, the proposed system can support almost twice as many active users as the existing schemes. Importantly, our design achieves

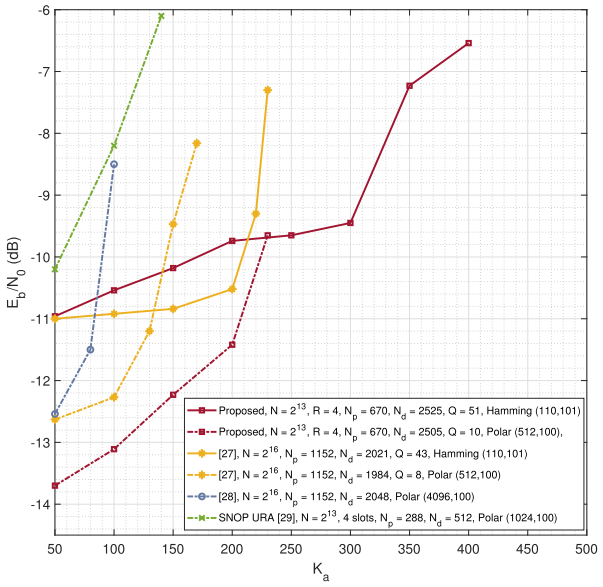

FIGURE 7. Minimum SNR required to support various numbers of active users, compared to the state-of-the-art. The target PUPE equals 0.03.

this with a significantly shorter preamble length (nearly half) and smaller pool size (reduced by a factor of 8) compared to [27] and [28]. This reduces the complexity of the iterative preamble detection significantly and allows for more repetitions in the payload improving both the detection and decoding. Moreover, as a result of AoA detection, the channel estimates provided to the MUD by the receiver's AD part are of significantly higher quality than the estimates in [27], [28], and [29]. This results in an enhanced MUD performance. The conducted simulations demonstrate that the average MSE of the channel estimation for the same SNRs is almost 4 dB better in the proposed scheme compared to [27], [28], and [29].

Fig. 7 illustrates a comparison with the state-of-the-art for the target PUPE of 0.03 and the same set of parameters as used for the systems in Fig. 6. The trend in the performance curves and the superior performance of the proposed scheme are also clearly observable. Slightly higher SNRs and lower maximum numbers of achievable users are consistent with the lower PUPE target.

A comparison to the performance results for the URA approaches reported in [17] and [19], tested on the Rayleigh channel, indicates that our proposed design can support almost twice as many active users and with higher power efficiency while operating on an equal-gain channel model which is known to be more challenging for the MUD. The proposed URA design can resolve collisions, which are not addressed in the aforementioned URA approaches. A similar comparison with [22] shows that the proposed design achieves a power reduction of approximately 25 dB and supports a higher number of active users. While [22] relies on the assumption that each user has a unique channel pattern created by different signal propagation paths, this is not required for the proposed design that performs well on a single-path model with equal channel gains. Contrary to [22], the proposed system which uses both preambles and AoAs to distinguish the users, can often differentiate users even if they are not well separated geographically and have strongly correlated channel coefficients.

IV. CONCLUSION

This paper proposes an approach to MIMO URA, where pseudo-random set of parameters used to modulate each active user's packet is tied to the user's location, specifically to the AoA towards the receiving BS. The proposed receiver algorithm is efficiently separating the colliding packets for the case of a dominant-path MIMO model which is well suited to describe mmWave and THz communications. The numerical results demonstrate that the proposed system significantly outperforms the existing MIMO URA methods by supporting more active users with lower operational SNR, utilizing considerably shorter preambles, and resolving the collisions.

REFERENCES

[1] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, "Grant-free non-orthogonal multiple access for IoT: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1805–1838, 3rd Quart., 2020.

[2] J. Choi, J. Ding, N.-P. Le, and Z. Ding, "Grant-free random access in machine-type communication: Approaches and challenges," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 151–158, Feb. 2022.

[3] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, Sep. 2016.

[4] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2523–2527.

[5] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1117–1174, 2nd Quart., 2022.

[6] O. Ordentlich and Y. Polyanskiy, "Low complexity schemes for the random access Gaussian channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2528–2532.

[7] 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.

[8] A. Fengler, P. Jung, and G. Caire, "SPARCs and AMP for unsourced random access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 2843–2847.

[9] V. K. Amalladinne, A. K. Pradhan, C. Rush, J.-F. Chamberland, and K. R. Narayanan, "Unsourced random access with coded compressed sensing: Integrating AMP and belief propagation," *IEEE Trans. Inf. Theory*, vol. 68, no. 4, pp. 2384–2409, Apr. 2022.

[10] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted Aloha," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.

[11] A. Glebov, N. Matveev, K. Andreev, A. Frolov, and A. Turlikov, "Achievability bounds for T-fold irregular repetition slotted Aloha scheme in the Gaussian MAC," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.

[12] A. Vem, K. R. Narayanan, J.-F. Chamberland, and J. Cheng, "A user-independent successive interference cancellation based coding scheme for the unsourced random access Gaussian channel," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8258–8272, Dec. 2019.

[13] 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)*, Dec. 2019, pp. 1–6.

[14] 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)*, Sep. 2019, pp. 1–5.

[15] 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, Dec. 2020.

[16] 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.

[17] A. Decurninge, I. Land, and M. Guillaud, "Tensor-based modulation for unsourced massive random access," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 552–556, Mar. 2021.

[18] A. Decurninge, I. Land, and M. Guillaud, "Tensor decomposition bounds for TBM-based massive access," in *Proc. IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2021, pp. 346–350.

[19] V. K. Amalladinne, J.-F. Chamberland, and K. R. Narayanan, "Coded compressed sensing with successive cancellation list decoding for unsourced random access with massive MIMO," 2021, *arXiv:2105.02185*.

[20] J. Che, Z. Zhang, Z. Yang, X. Chen, C. Zhong, and D. W. K. Ng, "Unsourced random massive access with beam-space tree decoding," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1146–1161, Apr. 2022.

[21] V. Shyianov, F. Bellili, A. Mezghani, and E. Hossain, "Massive unsourced random access based on uncoupled compressive sensing: Another blessing of massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 820–834, Mar. 2021.

[22] X. Xie, Y. Wu, J. An, J. Gao, W. Zhang, C. Xing, K.-K. Wong, and C. Xiao, "Massive unsourced random access: Exploiting angular domain sparsity," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2480–2498, Apr. 2022.

[23] X. Xie, Y. Wu, J. An, D. W. K. Ng, C. Xing, and W. Zhang, "Massive unsourced random access for near-field communications," *IEEE Trans. Commun.*, early access, Feb. 2, 2024, doi: 10.1109/TCOMM.2024.3361529.

[24] J.-C. Jiang and H.-M. Wang, "Two-phase unsourced random access in massive MIMO: Performance analysis and approximate message passing decoder," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1592–1607, Feb. 2024.

[25] J. Yan, G. Song, Y. Li, and J. Wang, "ODMA transmission and joint pattern and data recovery for unsourced multiple access," *IEEE Wireless Commun. Lett.*, vol. 12, no. 7, pp. 1224–1228, Jul. 2023.

[26] M. Ozates, M. Kazemi, and T. M. Duman, "Unsourced random access using ODMA and polar codes," *IEEE Wireless Commun. Lett.*, early access, Feb. 29, 2024, doi: 10.1109/LWC.2024.3359270.

[27] E. Nassaji, R. Soltani, M. Bashir, and D. Truhachev, "Spread unsourced random access with an iterative MIMO receiver," *IEEE Commun. Lett.*, vol. 26, no. 10, pp. 2495–2499, Oct. 2022.

[28] A. Fengler, O. Musa, P. Jung, and G. Caire, "Pilot-based unsourced random access with a massive MIMO receiver, interference cancellation, and power control," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 5, pp. 1522–1534, May 2022.

[29] M. Ozates, M. Kazemi, and T. M. Duman, "A slotted pilot-based unsourced random access scheme with a multiple-antenna receiver," *IEEE Trans. Wireless Commun.*, early access, Aug. 30, 2023, doi: 10.1109/TWC.2023.3308218.

[30] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, P. Popovski, and M. Debbah, "Seven defining features of terahertz (THz) wireless systems: A fellowship of communication and sensing," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 967–993, 2nd Quart., 2022.

[31] J.-M. Feng and C.-H. Lee, "Generalized subspace pursuit for signal recovery from multiple-measurement vectors," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2013, pp. 2874–2878.

[32] J. Kim, W. Chang, B. Jung, D. Baron, and J. C. Ye, "Belief propagation for joint sparse recovery," 2011, *arXiv:1102.3289*.

[33] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors. I. The MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1758–1774, Jul. 1992.

[34] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[35] P.-J. Chung and J. F. Bohme, "Recursive EM and SAGE-inspired algorithms with application to DOA estimation," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2664–2677, Aug. 2005.

[36] M. Gkagkos, K. R. Narayanan, J.-F. Chamberland, and C. N. Georghiades, "FASURA: A scheme for quasi-static massive MIMO unsourced random access channels," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2022, pp. 1–5.

[37] *Optical Internetworking Forum*, Optical Internetwork. Forum, Fremont, CA, USA, [Online]. Available: https://www.oiforum.com

ROSHANAK SOLTANI (Graduate Student Member, IEEE) received the B.Sc. degree from Babol Noshirvani University of Technology, Iran, in 2016, and the M.Sc. degree from Semnan University, Iran, in 2019, both in electrical engineering. She is currently pursuing the Ph.D. degree in electrical engineering with Dalhousie University, Canada. Her current research interests include wireless communication systems and multiuser communications.

DMITRI TRUHACHEV (Member, IEEE) received the Ph.D. degree in electrical engineering from Lund University, Sweden. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Dalhousie University, Canada. Before joining Dalhousie University, he held postdoctoral and research associate positions with the University of Alberta, Canada. His research interests include error-correction codes, spatial graph coupling, multiuser communications, fiber optical communications, and underwater acoustic communications.

ALIREZA BAYESTEH received the B.Sc. and M.Sc. degrees from the Sharif University of Technology, Iran, in 2000 and 2002, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2008. From 2008 to 2009, he was a Postdoctoral Fellow with the University of Waterloo. From 2009 to 2011, he was a Member of Technical Staff with Research In Motion (currently BlackBerry Ltd.). Since 2011, he has been with Huawei Canada, Ottawa, where he is currently a Senior Principal Engineer. His research interest includes various aspects of air interface design for 6G wireless communications.

MONIROSHARIEH VAMEGHESTAHBANATI (Member, IEEE) received the Ph.D. degree in electrical engineering from Carleton University, Ottawa, ON, Canada, in 2020. She is a Senior Wireless Research Engineer with the Huawei Technologies Canada Research Centre, Ottawa. She is also a co-inventor of more than 12 patents. Her research interest includes the advanced physical-layer technologies for future wireless systems. She is an Active Member of the IEEE Young Professionals. Her Ph.D. thesis was nominated for Senate Medal for outstanding academic achievement. She was a recipient of numerous awards, including the Ontario Trillium scholarship from 2013 to 2017 and the Outstanding Teaching Assistant Award at Carleton University from 2017 to 2018.