Unsourced Random Access: A Comprehensive Survey

Mert Ozates, Mohammad Javad Ahmadi, Mohammad Kazemi, Deniz Gündüz and Tolga M. Duman

***Abstract*—Multiple access communication systems enable numerous users to share common communication resources, playing a crucial role in wireless networks. With the emergence of the sixth generation (6G) and beyond communication networks, supporting massive machine-type communications with sporadic activity patterns is expected to become a critical challenge. Unsourced random access (URA) has emerged as a promising paradigm to address this challenge by decoupling user identification from data transmission through the use of a common codebook. This survey offers a comprehensive overview of URA solutions, encompassing both theoretical foundations and practical applications. We present a systematic classification of URA solutions across three primary channel models: Gaussian multiple access channels (GMACs), single-antenna fading channels, and multiple-input multiple-output (MIMO) fading channels. For each category, we analyze and compare state-of-the-art solutions in terms of performance, complexity, and practical feasibility. Additionally, we discuss critical challenges such as interference management, computational complexity, and synchronization. The survey concludes with promising future research directions and potential methods to address existing limitations, providing a roadmap for researchers and practitioners in this rapidly evolving field.**

***Index Terms*—Multiple access, random access techniques, grant-free, unsourced random access, machine-type communications, Internet of things.**

I. Introduction

The rapid proliferation of Internet of Things (IoT) devices is driving a fundamental shift in wireless communication systems. Current estimates project over 38 billion IoT connections by 2029, representing a more than twofold increase compared to the 15 million devices in 2023 [1]. One example of such systems is machine-type communications (MTC) (also known as machine-to-machine (M2M) communications), which, as a key aspect of the IoT paradigm, enables machines, sensors, and devices to exchange information without direct human intervention. MTC plays a crucial role in applications ranging from industrial automation and smart cities to healthcare monitoring and environmental sensing [2]. Specifically, massive MTC is anticipated to accommodate a density of up to one million devices (per square kilometer), which typically operate with low computational and storage capabilities, low duty cycles, and small payloads [3]. This massive scale presents unprecedented challenges for communication systems as traditional wireless access paradigms have been designed for a relatively small number of high-data-rate users rather than an enormous number of devices with sporadic low-rate transmissions.

Multiple access (MA) techniques that enable different users to share a common communication medium, have been fundamental to wireless communication systems [4]. MA technologies can be broadly categorized into two categories: centralized access and random access (RA). In centralized access (also known as *coordinated access*), a central authority controls and manages access to the shared communication channel by allocating distinct resources to each user for data transmission. This approach has been the predominant MA mode in the fifth generation (5G) and previous cellular technologies. However, centralized coordination becomes impractical for massive IoT scenarios due to prohibitive signaling overhead and coordination complexity. In contrast, in RA (also known as *uncoordinated access*), users transmit their messages without requiring explicit coordination with a central authority. This approach reduces signaling overhead and, in some cases, latency. These advantages make RA solutions particularly suitable for next-generation (6G and beyond) communication networks, especially when serving a large number of users [5], [6].

Traditional RA solutions, such as ALOHA [7] and its variants, suffer from an excessive number of collisions as the number of users increases. More sophisticated grant-free RA schemes [8] employ unique signatures or preambles for each potential user to facilitate collision reduction at the physical layer. Grant-free RA solutions can accommodate more users compared to ALOHA, as they can accomplish user scheduling through a limited amount of feedback [9], or by exploiting device-specific patterns such as codebook structures and spreading sequences [10]. On the other hand, with the number of devices within a certain coverage area reaching millions [15], it is not feasible to assign a distinct pilot/preamble to each device, as this would result in prohibitive receiver complexity. Therefore, existing grant-free RA solutions cannot address the multiple access problem in emerging massive access networks.

To mitigate the challenge of identifying and decoding

Mert Ozates is with the System Architectures Department, IHP - Leibniz Institute for High-Performance Microelectronics, 15236 Frankfurt (Oder), Germany (e-mail: oezates@ihp-microelectronics.com).

Mohammad Javad Ahmadi is with the Chair of Information Theory and Machine Learning, Technische Universität Dresden, 01062 Dresden, German (E-mail: mohammad_javad.ahmadi@tu-dresden.de)

Mohammad Kazemi and Deniz Gündüz are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, U.K. (e-mail: {mohammad.kazemi,d.gunduz}@imperial.ac.uk)

Tolga M. Duman is with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, 06800, Turkey (email: duman@ee.bilkent.edu.tr)

This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 119E589. Mert Ozates' work is supported by the Federal Ministry of Education and Research of Germany in the programme of "Souverän. Digital. Vernetzt" joint Project 6G-RIC, project identification number: 16KISK026. Mohammad Kazemi's work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 101103430].

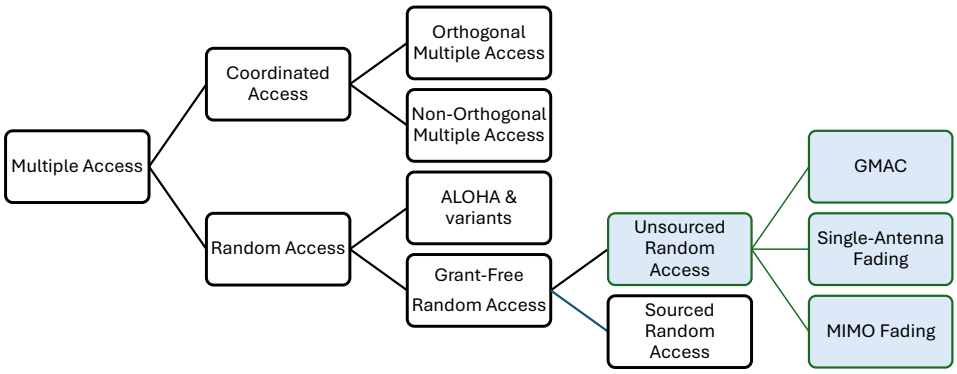


Fig. 1: Taxonomy of different MA techniques.

messages from such a large number of potential devices, Polyanskiy introduced the *unsourced random access* (URA) paradigm in [11]. URA relies on the observation that, despite their increasing numbers, MTC and IoT devices are activated only sporadically; that is, at any time instant, only a handful of devices are active simultaneously. URA fundamentally reimagines the massive access problem by decoupling user identification from data transmission. Instead of assigning unique preambles or codebooks, users employ a common codebook, and the receiver focuses solely on recovering the list of transmitted messages. Since a common codebook is employed, the user identity is eliminated and the system can operate irrespective of the total number of devices, where only the number of active devices becomes important. This paradigm shift drastically reduces system complexity, and enables practical implementations for truly massive numbers of devices. Following Polyanskiy's pioneering work, several low-complexity coding solutions have been developed in the subsequent literature, where successive interference cancellation (SIC) techniques are employed to resolve collisions up to a certain level, depending on the user's transmit powers. Many techniques, such as random spreading [12], multi-stage orthogonal pilots [13], and time frame slotting [14] have been employed to mitigate collisions and interference. While a single-antenna receiver in a Gaussian MA channel (GMAC) setup is considered in the earlier works, the literature has shifted towards fading channels with multiple-antenna receivers, especially massive multiple-input multiple-output (MIMO), to support more practical channel models and an even greater number of users. A block diagram of different MA techniques is depicted in Fig. 1 alongside the three main setups in URA research, namely, GMAC, and single-antenna and MIMO fading channels.

A. Objectives and Contributions

Over the years since Polyanskiy's initial work, a large number of papers have appeared studying URA from both theoretical and practical perspectives. Although prior efforts have been made to review existing works on URA, our survey fills an important gap by providing a systematic classification of the literature and covering more recent works on practical channel models and device characteristics. The first survey on massive access was [15], in which the authors review many-user MA, where each user is assigned an individual codebook (i.e., the coordinated approach), as well as URA, which they call uncoordinated massive access. However, since URA was in its early stages at that time, the survey in [15] covers mainly the earlier works over GMAC. In [16], the focus is again on URA over GMAC, and the authors review the existing design paradigms and coding schemes for URA over GMAC, while URA over fading channels and in MIMO systems are presented as open problems. In a less URA-oriented survey [17], an overview of compressive sensing-based grant-free massive access is provided, which also includes some compressed sensing-based URA schemes. In a recent survey [18], through an overview of the existing URA schemes studying mostly GMAC, the authors present URA as a suitable candidate for next-generation communication systems and highlight that it can support the 3rd Generation Partnership Project (3GPP) standardization efforts.

Table I: Summary of the surveys on URA.

Paper	Scope
Wu et al., 2020 [15]	Review of many-user MAC and URA with a focus on GMAC.
Li et al., 2022 [16]	Review of basics of URA and coding schemes on URA over GMAC.
Gao et al., 2024 [17]	Review of grant-free massive access with a focus on compressed sensing-based schemes.
Liva et al., 2024 [18]	Review of URA over GMAC and showing that it can support 3GPP standardization.
This survey	A comprehensive review of the URA literature, covering all existing solutions in GMAC and single-antenna and MIMO fading channels.

While the existing surveys focus on the GMAC setup, we also cover the URA solutions that take into account channel fading, which comprise the majority of recent works in the URA literature. More specifically, we conduct a comprehensive review of the existing URA solutions in the literature across three main setups, namely, URA over GMAC, and single-antenna and MIMO fading channels, accompanied by a discussion on the main challenges, possible solutions, and promising research directions. We study works that propose coding schemes, as well as those that provide fundamental information-theoretical analysis for URA. To the best of our knowledge, this is the first comprehensive survey on URA, reviewing this crucial emerging paradigm for 6G and beyond communication networks in all aspects. A summary of the

Section	Subsection/Subsubsection	
I. Introduction	A. Objectives and Contributions	
	B. Structure of the Survey	
II. Background: Multiple Access	A. Coordinated Access	
	B. Random Access	1) ALOHA & Variants
		2) Grant-Free Random Access
III. URA: A Recent Paradigm	A. General System Model	
	B. Challenges	
IV. URA over GMAC	A. Performance Bounds	
	B. URA Solutions over GMAC	1) Slotting-based schemes
		2) CCS-based schemes
		3) Random spreading-based schemes
		4) IDMA-based schemes
		5) ODMA-based schemes
	C. Summary and Comparison of GMAC Solutions	1) Performance Comparison
		2) Computational Complexity Comparison
V. URA over Fading Channels	A. Performance Bounds	
	B. Single-Antenna Fading Channels	1) CCS-based schemes
		2) Joint decoding-based schemes
		3) TIN-SIC-based schemes
		4) Alternative solutions
	C. Summary and Comparison of Single-Antenna Fading Solutions	1) Performance Comparison
		2) Computational Complexity Comparison
	D. MIMO Fading Channels	1) CCS-based schemes
		2) Tensor-based schemes
		3) Pilot-based schemes
		4) ODMA-based schemes
		5) Bayesian approach
	E. Summary and Comparison of MIMO Fading Solutions	1) Performance Comparison
		2) Computational Complexity Comparison
	F. Existing URA Solutions on Alternative Setups	
VI. Conclusions	A. Lessons Learned	
	B. Future Research Directions	1) Asynchronous URA
		2) Robust URA
		3) Machine-Learning Aided URA
		4) Semantic and Task-Oriented URA
		5) ISAC in URA
		6) URA with Feedback
		7) RIS-Aided URA
		8) Security
		9) Standardization

Fig. 2: Structure of the survey.

existing surveys and our contribution can be found in Table I.

In particular, the paper contains the following key contributions:

- Comprehensive overview: Delivering an extensive survey of URA solutions, examining both theoretical and practical perspectives.
- Classification: Systematically categorizing URA solutions according to the fundamental channel models: GMAC, single-antenna fading channels, and MIMO fading channels.
- Performance and feasibility analysis: Evaluating and comparing state-of-the-art URA solutions in terms of energy efficiency, computational complexity, and practical viability.
- Identifying challenges: Addressing essential challenges in URA in terms of interference, collisions, channel estimation, and computational complexity.
- Guiding future research: Proposing promising future research directions and potential solutions to overcome existing limitations in URA systems.

Throughout the survey, we adopt a unifying narrative: to demonstrate how URA techniques progress from theoretical achievability results on the GMAC to increasingly practical frameworks under fading and MIMO channels. This progression underscores the significance of URA as an enabler of massive connectivity in emerging 6G and IoT systems.

B. Structure of the Survey

The remainder of the survey is organized as follows. In Section II, a general discussion of MA techniques is provided. We begin with coordinated MA techniques and progress to random access. In Section III, we focus on URA as a new paradigm for uncoordinated MA communications with a massive number of potential devices, enumerate its advantages, and hint at some solutions developed in the literature to address the related challenges. Sections IV and V are dedicated to in-depth overviews of the existing URA solutions, with Section IV focusing on URA over GMAC and Section V on URA over fading channels. Regarding communication models over practical wireless channels, we first investigate the works that take into account channel fading with a single antenna receiver and describe the general approaches; then, we discuss those with a MIMO assumption. Finally, we conclude the survey in Section VI with remarks on URA solutions and provide future research directions. The structure and organization of the paper are illustrated in Fig. 2. Unlike prior surveys that focus primarily on GMAC, our structure integrates both single-antenna and MIMO fading URA solutions, thereby offering a comprehensive view that highlights practical deployment challenges. While our survey spans the full URA literature, the unifying application context is 6G and IoT, where sporadic traffic and massive device density make URA particularly relevant.

The acronyms used in this paper are alphabetically summarized in Table II.

II. Background: Multiple Access

Multiple access refers to the scenario in which two or more users simultaneously share the system resources. In this section, we provide a concise overview of the main MA approaches and outline their pros and cons to better understand the position of URA solutions in next-generation communication systems.

Table II: List of acronyms

Acronym	Definition
5G	fifth generation
5G-NR	5G new radio
6G	sixth generation
AD	activity detection
AMP	approximate message passing
AOA	angle of arrival
AWGN	additive white Gaussian noise
BCH	Bose–Chaudhuri–Hocquenghem
BP	belief propagation
BPSK	binary phase shift keying
BS	base station
CB-MLD	covariance-based MLD
CCS	coded/coupled compressive sensing
CDMA	code division MA
CP	cyclic prefix
CRC	cyclic redundancy check
CS	compressed sensing
CSA	coded slotted ALOHA
CSI	channel state information
DFT	discrete Fourier transform
DOF	degree of freedom
FBL	finite blocklength
FDMA	frequency division MA
FEC	forward error correction
FFT	fast Fourier transform
FL	federated learning
GB-CR	graph-based channel reconstruction and collision resolution
GMAC	Gaussian MA channel
gOMP	generalized orthogonal matching pursuit
HWI	hardware impairments
HyGAMP	hybrid generalized AMP
i.i.d.	independent and identically distributed
IDMA	interleave-division multiple access
IISD	iterative inter-symbol decoder
IoT	Internet-of-things
IRSA	irregular repetition slotted ALOHA
ISAC	integrated sensing and communication
LASSO	least absolute shrinkage and selection operator
LDPC	low-density parity check
LLR	log-likelihood ratio
LMMSE	linear minimum mean square error
LS	least squares

Acronym	Definition
MA	multiple access
MAC	MA channel
MIMO	multiple-input multiple-output
ML	machine learning
MLD	maximum likelihood detection
MMSE	minimum mean square error
mMTC	massive machine-type communications
MMV-AMP	multiple measurement vector AMP
MRC	maximal ratio combining
MSE	mean square error
MS-MRA	multi-stage setup with multiple receive antennas
MS-MRA-WOPBE	MS-MRA without pilot bits encoding
MSUG	multi-stage setup with user grouping
MSUG-MRA	MSUG for multiple receive antennas
MTC	machine-type communications
MUSIC	multiple signal classification
NNLS	non-negative least squares
NOMA	non-orthogonal MA
NP	Neyman-Pearson
ODMA	on-off division MA
OFDM	orthogonal frequency division multiplexing
OMA	orthogonal MA
OMP	orthogonal matching pursuit
PMF	probability mass function
PUPE	per-user probability of error
QPSK	quadrature phase shift keying
RA	repeat accumulate
RIS	reconfigurable intelligent surface
RS	Reed-Solomon
SCLD	successive cancellation list decoder
SCMA	sparse code MA
SDMA	space division MA
SIC	successive interference cancellation
SINR	signal-to-interference-plus-noise ratio
SISO	soft-input soft-output
SKP	sparse Kronecker product
SNR	signal-to-noise ratio
SPARC	sparse regression code
TDMA	time division MA
TIN	treating interference as noise
UE	user equipment
UNISAC	unsourced ISAC
URA	unsourced RA

A. Coordinated Access

One way to serve multiple users is to assign dedicated resources (e.g., time, frequency, code, or their combinations) to them by a central coordination unit through a prior handshaking procedure. The conventional techniques in this framework include time-division MA (TDMA), frequency-division MA (FDMA), orthogonal frequency-division MA (OFDMA), code-division MA (CDMA), and space-division MA (SDMA). In TDMA and FDMA, non-overlapping time slots and frequency sub-channels are assigned to the different users, respectively, and each user utilizes its allocated time frame or frequency band for transmission [4]. On the other hand, in CDMA, all users can simultaneously use all time-frequency resources through different (nearly) orthogonal code sequences, coupled with a low-complexity decorrelation-based receiver. OFDMA provides frequency division multiplexing by assigning different sets of subchannels to the individual users in an orthogonal frequency division multiplexing (OFDM) system, while SDMA relies on the principle of creating (nearly) orthogonal spatial channels for different users to minimize the inter-user interference by employing beamforming and spatial multiplexing [19].

All the approaches described in the previous paragraph fall under the category of OMA, where the number of users is strictly limited due to orthogonal resource assignment. To increase the number of supported users, non-orthogonal multiple access (NOMA) techniques have also been investigated in the literature, where multiple users can utilize the same resources to transmit their data, thereby allowing the system to support a greater number of users. However, this imposes inter-user interference on the system, and more sophisticated interference-canceling receivers are needed with an increased complexity compared to OMA.

NOMA schemes are divided into two main categories: power-domain NOMA and code-domain NOMA. In power-domain NOMA [20], the users employ different power levels to transmit their data using the same resources, and the receiver exploits the received power differences for detection through SIC. On the other hand, code-domain NOMA is similar to classical CDMA in that users employ user-specific code

sequences to transmit their data. The code sequences are sparse or non-orthogonal sequences with low cross-correlations [21]. Here, the receiver can employ message-passing algorithms that utilize the code structure to differentiate between users. Some examples of code-domain NOMA schemes in the literature are low-density spreading code division MA [22], low-density spreading aided orthogonal frequency division multiplexing [23], sparse code MA (SCMA) [24], and multiuser shared access [25].

B. Random Access

All of the schemes described in the previous subsection are coordinated. That is, they require scheduling grants and coordination with the base station (BS) before transmission, and it is assumed that the number of users and their activity patterns are available at the BS. However, when the number of users becomes large, they become infeasible due to the excessive delays and signaling overhead. For example, the typical number of users per signal dimension in the time-frequency domain (per degree of freedom) is between 1.5 and 3 according to 3GPP studies, as the system performance significantly degrades beyond this threshold [15]. The common solution addressing these drawbacks is random access, where the users perform their transmission without any coordination by the BS whenever they have data. ALOHA and grant-free random access are the two common RA approaches in the literature.

1) ***ALOHA & Variants****:* An early RA approach is the ALOHA protocol [7] (also known as *pure ALOHA*) proposed in the 1970s, where a device sends the entire packet at a randomly selected time. If no acknowledgment is received within a certain period of time, it is assumed that a collision has occurred with a packet sent by another device, and the packet is retransmitted after an additional random waiting time to avoid repeated collisions. This process is repeated until a successful transmission and its subsequent confirmation occur or until the process is terminated by the user console. Slotted ALOHA (SA) is a variant of ALOHA, where packets are transmitted in designated time slots, thereby doubling the maximum throughput compared to pure ALOHA [26]. However, collision is still a significant challenge in SA, as several active users may share the same slot to send their data packets, leading to packets colliding within that slot and the discarding of data. The transmission structures in pure ALOHA and slotted ALOHA are illustrated in Fig. 3.

With the motivation to increase the throughput, several variants of SA have been proposed in recent literature, where the main idea is to transmit multiple copies of a packet in different time slots. If the packet is successfully recovered in one slot, the effect of its replicas in the other slots is removed. Based on this idea, two copies of the same packet are transmitted in contention resolution diversity slotted ALOHA (CRDSA) [27], and the packet repetitions are picked from a probability mass function optimized for higher throughput in irregular repetition slotted ALOHA (IRSA), proposed in [28]. An extension of IRSA, exploiting forward error correction, is called coded slotted ALOHA (CSA) [29]. This approach

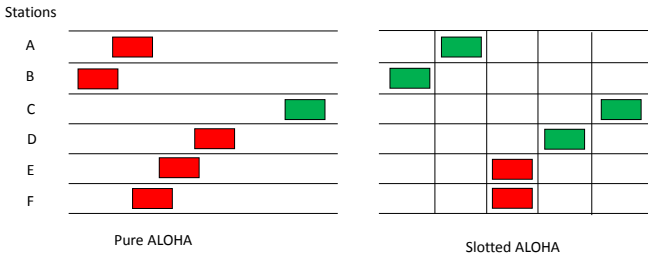


Fig. 3: An illustration of pure and slotted ALOHA. Red boxes show the collided packets, and green boxes are the successfully transmitted ones.

achieves a throughput of almost one when sufficiently low coding rates are employed. IRSA and CSA are further studied in [30]–[34]. Moreover, the resolution of partial collisions in random access setups is discussed in [35].

2) ***Grant-Free Random Access****:* Grant-free RA is another RA candidate that enables communication among a large number of users, where users transmit their data by utilizing shared system resources without requiring a prior grant from the BS, thereby reducing latency and signaling overhead. Nevertheless, the users can have unique signatures/preambles or they can utilize different codebooks, which means that there is user identity during the data transmission phase, and the receiver can use this information for activity detection or user separation [36]. This is also called *sourced random access*. Among the grant-free RA solutions, grant-free NOMA is a prominent solution for future communication systems due to its low latency, improved spectral efficiency, and potential to support massive connectivity. Note that the main difference between grant-based and grant-free NOMA is that the BS does not have information on the set of active users in the latter, which brings significant challenges requiring different solutions such as blind activity detection, channel estimation, data recovery, and synchronization with minimal overhead [37].

Among the two main categories of NOMA, power-domain NOMA is challenging to extend to the grant-free scenario, as the successful recovery of transmitted packets at the receiver heavily depends on the power difference between users, which may not be maintained due to the grant-free nature of the transmissions. On the other hand, signature-based schemes in code-domain NOMA can be adapted to the grant-free case with proper activity detection before the data detection or joint activity detection and data recovery [38]. For instance, a grant-free version of SCMA is proposed in [39]. Furthermore, the inherent sparsity of the user activity in grant-free MA is utilized in [40]–[43] through compressive sensing algorithms for multiuser detection.

III. URA: A Recent Paradigm

In some next-generation communications applications, such as massive MTC and massive IoT, the total number of devices can be in the order of millions, and they are only active very sporadically, as illustrated in Fig. 4. This makes achieving any level of coordination infeasible. As a result, none of the

MA techniques summarized in the previous section is suitable for enabling the communication of these devices. Specifically, the complexity of the grant-free MA is prohibitive due to the codebook size when each potential user is assigned a unique preamble. On the other hand, ALOHA leads to excessive collisions, significantly degrading communication efficiency.

To address this problem, Polyanskiy proposed the URA paradigm in [11] where the data transmission and user identification problems are decoupled via the utilization of a common codebook by all users. Therefore, the receiver is not interested in user identities during data transmission, with decoding being performed only up to a permutation of the transmitted messages.

A. General System Model

For brevity, we present the general system model for a URA system with a multiple-antenna receiver and complex channel coefficient, since conversions to real channel coefficients and a single-antenna receiver can easily be obtained as special cases. We also assume a quasi-static channel, where the channel coefficients remain constant over the transmission frame [1], as the setup can easily be extended to cases where the channel coefficients vary over the frame by replacing the frame with a coherence time block or a time slot over which the channel remains constant. Note that these assumptions are made for the sake of brevity in presenting the general URA system model; however, we have investigated all existing URA solutions in the literature, regardless of their specific system models. Considering a URA system with an M_r-antenna receiver and single-antenna users, in which K_a out of K_T users are active at a given frame of length n [11], the received signal, $\mathbf{Y} \in \mathbb{C}^{M_r \times n}$, in the absence of synchronization errors can be written in the following general form [13]

$$\mathbf{Y} = \sum_{i=1}^{K_a} \mathbf{h}_i \mathbf{x}_i(\mathbf{u}_i) + \mathbf{Z} \in \mathbb{C}^{M_r \times n}, \tag{1}$$

where $\mathbf{x}_i(\mathbf{u}_i)$ is the $1 \times n$ transmitted signal of the i-th user corresponding to its message bit sequence $\mathbf{u}_i \in \{0,1\}^B$ of length B, $\mathbf{Z} \in \mathbb{C}^{M_r \times n}$ is the circularly-symmetric additive white Gaussian noise (AWGN) with independent and identically distributed (i.i.d.) elements drawn from $\mathcal{CN}(0, \sigma_z^2)$, with σ_z^2 being the noise power, and $\mathbf{h}_i \in \mathbb{C}^{M_r \times 1}$ is the vector of channel coefficients. Note that in the URA literature over GMAC, a single-antenna receiver and unit channel coefficients are assumed. However, the default assumption for fading channels is quasi-static fading; i.e., the channel coefficients remain constant across a frame and change from one frame to the next.

In the URA literature, it is assumed that each user selects its message index uniformly from the set $\{1, 2, ..., 2^B\}$. The objective in URA is to determine the list of transmitted messages, $\mathbf{u}_i$'s, $\forall i = 1, 2, ..., K_a$, by processing the received signal $\mathbf{Y}$ so that the required energy-per-bit for a target per-user probability of error (PUPE), as the performance metric, is minimized

[1]The transmission frame is the collection of available resources in different dimensions, such as time and frequency, to be used for transmission. We call it *frame* for brevity hereafter.

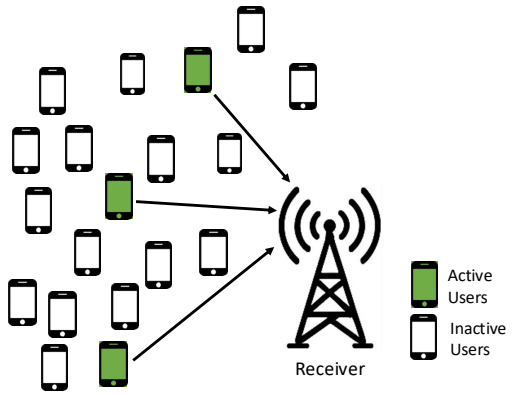


Fig. 4: An illustration of a system where a large number of potential users (devices) try to communicate with a receiver (access point), with only a small subset of them being active at each time frame.

[12], [57], [102]. The PUPE captures both missed detections and false alarms. Specifically, it is defined as the sum of probabilities that an active user's message is not correctly included in the decoded list (probability of misdetection, p_{md}), and that an invalid message appears in the list (probability of false-alarm, p_{fa}); i.e., [14]

$$P_e = p_{fa} + p_{md}, \tag{2}$$

where $p_{md} = \frac{1}{K_a} \sum_{i \in \mathcal{K}_a} \Pr(\mathbf{u}_i \notin \mathcal{L}_d)$, where $\mathcal{K}_a$ and $\mathcal{L}_d$ are the set of active user indices and the list of decoded messages, respectively, and $p_{fa} = \mathbb{E}\left\{\frac{n_{fa}}{|\mathcal{L}_d|}\right\}$, with n_{fa} being the number of decoded messages that were indeed not sent.

The energy per bit can be written as

$$\frac{E_b}{N_0} = \frac{nP}{c\sigma_z^2 B}, \tag{3}$$

where P denotes the average power of each user per channel use, i.e., $\|\mathbf{x}_i\|^2 \leq nP$, and we have $c = 2$ for real-valued and $c = 1$ for complex-valued channels.

B. Challenges

In the following sections, we review existing URA solutions in the literature. However, to gain a better understanding of the URA setup and the reasoning behind the solutions provided in the literature, we first need to identify the main challenges faced in a URA setup, which will be discussed in the remainder of this section.

Interference: Interference stands out as the primary challenge in MA channels, especially in random access setups such as URA, due to the lack of coordination between the receiver and the users. Failure to incorporate strategies to reduce or cancel it can lead to significant performance degradation. To demonstrate the detrimental impact of interference, we analyze the treating interference as noise (TIN) scheme, where single-user decoding is carried out by taking interference from other users as noise, without employing any special technique to mitigate it. The signal-to-interference-plus-noise ratio (SINR) of the i-th user, α_i, in the TIN strategy for a GMAC scenario

can be written as

$$\alpha_i = \frac{P}{\sigma_z^2 + (K_a - 1)P}, \tag{4}$$

where P is the transmit power of each user. Note that the SINR in a fading scenario is approximately the same as in (1) since we can approximate $\frac{1}{M_r}\|\mathbf{h}_i\|^2 \approx 1, \ \forall i$, where $\|\cdot\|$ denotes the Euclidean norm. For a sufficiently large active user load K_a, the SINR can be written as

$$\alpha_i \approx \frac{1}{K_a}. \tag{5}$$

This equation illustrates that when the number of users exceeds a specific threshold (more specifically, when it is much greater than the inverse of the signal-to-noise ratio (SNR), i.e. $K \gg \frac{\sigma_z^2}{P}$, which is not a very large value in many envisioned URA applications), the SINR becomes inversely proportional to K_a. Therefore, boosting the transmit power does not result in an enhanced SINR, or equivalently, improved system performance. Consequently, in systems such as URA, the relatively large number of active users has an adverse impact on system performance.

There are two main techniques in the URA literature to alleviate the effects of interference, with the most common being SIC, which removes the effects of successfully detected messages from the received signal to decrease the interference before another round of decoding. Especially in URA, SIC is a crucial step in achieving low error rates due to heavy multiuser interference. Other techniques for interference management involve randomly separating users into different groups, such as slotting the time frame, employing different transmit power levels, and multiplying or concatenating a known sequence to the data segment.

1) Time-domain separation: Dividing the frame into multiple slots and letting each user randomly select a slot to transmit its data can lead to improved system performance. In a slotted structure, the level of interference is reduced due to the lower number of interfering users in each slot. At the same time, the gain of channel coding is also decreased because fewer channel uses are available to each user. Thus, optimizing the number of slots based on system parameters, such as active user load, is crucial for maximizing the advantages of slotting.
2) Code-domain separation: In some URA schemes, each user randomly selects a preamble/pilot/pattern and uses it to prepare its data for transmission. One way is to transmit a concatenation of a preamble or pilot and the main signal. At the receiver end, the channel coefficients for different users can be estimated using the detected pilots. Then, minimum mean square error (MMSE) or maximum ratio combining (MRC) solutions can be employed to separate the user packets, before decoding them using single-user channel decoders. Another way is to multiply a randomly selected pattern by the main signal, which is used in random-spreading-based solutions. However, employing these techniques usually decreases the amount of power and time resources dedicated to the data (main signal) part. Therefore, for interference management to be effective, the choices of type, length, and power of the pilots and pattern sequences are crucial.
3) Power-domain separation: An alternative approach for user separation, especially effective in GMAC scenarios, is the so-called *power diversity* technique [44], where each user randomly picks a transmit power level from a limited set with a specific probability distribution, satisfying the overall power constraint. Since the users are now separated by different power levels, the receiver, with a high probability, decodes the users with the highest power level first, and after each SIC step, it moves to decode the lower power levels. The appropriate design of the power diversity parameters, namely, the number of power levels, their values, and their probability distribution, is crucial for optimizing the system performance.

Collisions: In many classical random access setups, such as ALOHA, collision is used interchangeably with interference, i.e., when user packets are fully or partially received at the same time at the receiver. However, in the URA literature, collision by default refers to the collision of either user messages or user pilots, which occurs when two or more users have the same messages or pilot sequences, respectively.

1) Message collision: In the case of user message collision, with the help of a randomized encoding process independent of user messages, the receiver can recover collided user messages to some extent. Note that in a conventional URA setup, since the receiver is only interested in the list of transmitted messages, being able to recover at least one of the collided messages does not increase the misdetection rate. However, this matter becomes more important in some URA systems where the receiver is also interested in the frequency of user messages.
2) Pilot collision: Many existing works taking into account channel fading employ preambles/pilots to estimate the channel coefficients used in the decoding phase. However, it is impossible to assign dedicated preambles/pilots due to the massive number of potential users. Therefore, as a common approach, the users independently pick a preamble/pilot from a common codebook, either randomly or based on a part of their message bits. Hence, if the pilot bits of multiple users are the same (called pilot contamination in massive MIMO literature [45]), a collision occurs, and the system struggles to decode the transmitted packets. Therefore, it is crucial to perform a preliminary analysis to optimize the design parameters to keep the pilot collision probability lower than a desired threshold or develop techniques for its mitigation.

Computational Complexity: In order to increase the number of active users supported by a URA system, larger values of parameters such as frame length and the number of antenna elements (in MIMO systems) are necessary. However, the computational complexity of URA solutions typically increases superlinearly with these parameters, as well as the number of active users. As a result, solutions for a URA sys-

tem with high dimensions, including temporal (frame length), spatial (number of receiver antennas), and/or the number of active users, are expected to experience high computational complexity. To mitigate this complexity, various strategies are employed in the literature, such as slotted transmissions and coded compressed sensing (CCS). The idea of CCS is to divide the bit sequence into smaller segments, map each segment to a column of a common codebook, and stitch them together using a tree code, rather than mapping the entire bit sequence to the columns of a $2^B \times n$ codebook, which would lead to prohibitive receiver complexity. This reduction in dimensionality helps lower the computational load at the decoder, albeit at the cost of reduced performance.

In this section, we summarized the main challenges in a URA setup along with a high-level discussion on the proposed solutions in the literature. In the two subsequent sections, we provide a more detailed examination of these solutions, with Section IV dedicated to GMAC and Section V covering single-antenna (V.A) and MIMO (V.B) fading setups. We take a systematic approach in presenting the URA solutions in these sections: different contributions are grouped based on their solution approach. In each group, the underlying approach is explained with an emphasis on the main structure of its encoding and decoding approaches, and the relevant works are presented in chronological order. To provide an overall view, at the end of each main subsection, a summary of the primary contributions is presented, and a comparative discussion of their performance is given, accompanied by some numerical results.

IV. URA over GMAC

In this section, we provide an overview of URA works over GMAC. To this end, we first describe an achievability bound on the performance of the URA over GMAC, as developed in Polyanskiy's original paper [11], and then review the existing URA over GMAC solutions to achieve it.

A. Performance Bounds

Polyanskiy provides a random coding achievability bound on the performance of URA over GMAC in [11], as a combination of a Chernoff-type bound and an information density-based one, which has been used as a benchmark for the energy efficiency in the URA over GMAC literature. The random coding bound is an information-theoretic performance limit, derived without any complexity constraints. It assumes that all active users are jointly decoded, which has a prohibitive complexity. So, practical coding schemes for URA with reasonable complexity are needed to achieve the results predicted by the bound.

In [11], the performances of common RA solutions are evaluated in a URA setup, assuming that there are K_a active users transmitting $B = 100$ bits of information through a transmission frame of length $n = 30000$, with a target PUPE of $\epsilon = 0.1$. Note that the same set of parameters has been adopted by subsequent URA works over GMAC for performance comparison, for obvious reasons. However, a more comprehensive analysis of the existing competing solutions, especially those nominated for inclusion in future protocols, is needed that covers different sets of parameters representing different future URA applications.

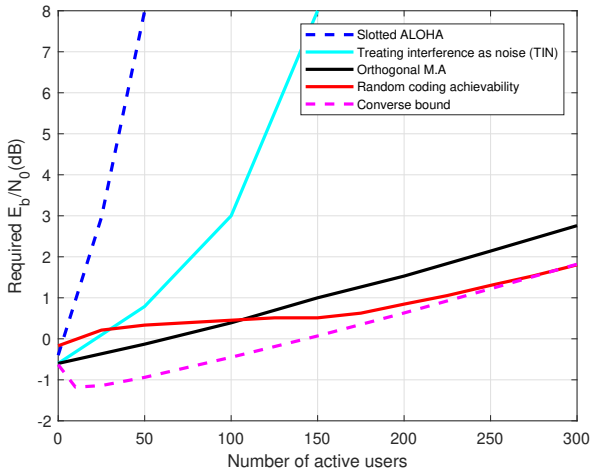


Fig. 5: Performance comparison of conventional random access techniques with Polyanskiy's URA (random coding) achievability bound.

The curves for the required E_b/N_0 versus the number of active users for slotted ALOHA, TIN, and orthogonal MA are plotted in Fig. 5 along with the random coding bound derived in [11] and a converse bound for K_a-GMAC from [47]. The results in Fig. 5 show that the conventional RA techniques (i.e., slotted ALOHA and TIN) are energy inefficient in a URA setup, as there is a huge performance gap between these and the random coding bound, and the development of low-complexity coding schemes is needed. For slotted ALOHA, it is assumed that a user packet is successfully recovered if there is no collision and the single-user decoding is successful, while the finite blocklength capacity of TIN is estimated using the normal approximation in [46]. The performance of the grant-based TDMA scheme is also plotted in the figure (referred to as *Orthogonal M.A*) as a benchmark, in which the frame is divided into sub-frames of length $\frac{n}{K_a}$ and each active user transmits in its dedicated sub-frame, with its required E_b/N_0 being calculated using the normal approximation in [46]. However, note that the TDMA scheme is not feasible in a URA setup, mainly due to the lack of coordination. Leaving the coordination issue aside, the system works in the regime of $n < K_a B$ in some URA applications, rendering orthogonal solutions such as TDMA infeasible.

The random coding bound in [11] is developed under the assumption that the receiver knows the number of active users K_a. The authors in [48] extend this bound to the case where the number of active users is random and unknown at the receiver. Furthermore, an achievability bound for the binary-input signaling for URA over GMAC is derived in [49]. A random coding bound for the scenario that the users send a common alarm message in addition to standard messages is developed in [50], where the alarm message needs to be decoded with a lower error rate than the standard ones. An achievability bound for the case of asynchronous GMAC is obtained in [51], where the delay in transmission is assumed

to be up to a small portion of the time frame. However, the provided achievability bound is loose for the asynchronous case as the authors have employed a fixed frame decoder in which all the packets that have some overlaps with the end of the frame are deemed undecodable, leading to huge performance degradation, which is not valid in a fully asynchronous setup, where the time horizon is unbounded. The drawbacks in [51] are alleviated in [52], where the authors study a wrap-around version of the model in [51], that is, the last D_m symbols are added to the first D_m symbols, where D_m is the maximum delay. As a result, the receiver observes the superposition of K_a cyclically shifted codewords with Gaussian noise, where the noise power in the first D_m symbols is doubled. In this way, the energy efficiency of the bound in [51] is significantly improved. The problem of joint delay and activity detection in the asynchronous setup is investigated in [53] using methods of tropical geometry, which is the study of the geometric properties of polynomials when addition is replaced by minimization and multiplication by ordinary addition. Moreover, the conditions for path delays for zero-error detection are identified. Furthermore, the authors in [54] examine a two-user binary adder channel in the asynchronous scenario, and propose an algorithm to avoid the small weight stopping sets in the joint factor graph, which mitigates the error floor due to these sets.

B. URA Solutions for GMAC

As shown in the previous subsection, the conventional RA techniques are energy-inefficient for the URA framework. As a result, there have been substantial efforts to develop low-complexity and energy-efficient schemes that can be used in practical systems. The main approaches for this purpose in the literature can be categorized as slotting the transmission frame, coded compressed sensing (CCS), random spreading, interleave-division MA (IDMA), and on-off division MA (ODMA). In the following, we review these schemes.

*1) **Slotting-based schemes**:* One of the design strategies is slotting the transmission frame, where each active user randomly picks one or a few slots to transmit its data, depending on the exact setting. In this direction, Ordentlich and Polyanskiy propose the first practical scheme for URA in [55], where a concatenated coding scheme combined with the T-fold ALOHA protocol is considered. T-fold ALOHA is an approach similar to slotted ALOHA in the sense that the users choose a random sub-block (slot) to transmit their packets, but in slotted ALOHA, packets are lost in the case of a collision. However, in T-fold ALOHA, T or fewer users can be simultaneously decoded in each slot. In [55], each user randomly chooses one slot to transmit its message. This concatenated coding scheme consists of an inner binary linear code, which is used to decode the modulo-2 sum of the codewords transmitted within a slot, and an outer code to recover the individual messages. Many off-the-shelf codes can be used for the inner code. The authors construct the outer code from the columns of a T-error correcting Bose–Chaudhuri–Hocquenghem (BCH) code. This scheme suffers in terms of energy efficiency as the performance gap with the random coding bound is about 20 dB when the number of active users is 300. Its performance is significantly improved in [56] (by up to 18 dB), where the authors employ a concatenated coding scheme with polar and BCH codes as inner and outer codes, respectively, and propose a decoding approach that exchanges information between inner and outer decoders. Specifically, all the candidate codewords from the inner successive cancellation list decoding of the polar code, rather than just one, are passed to the outer BCH decoder to facilitate its error checking. In addition, indicators for the decoding status of the outer BCH decoder are used to enhance the inner polar decoding process.

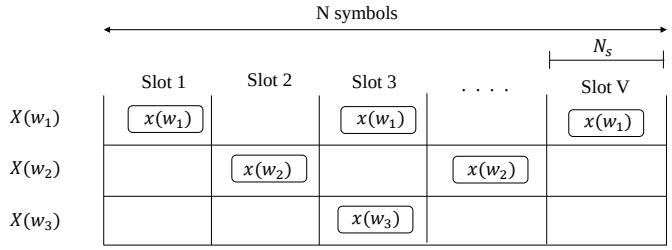


Fig. 6: The slotted transmission structure in [57] based on T-fold IRSA.

The authors in [57] consider a slotted structure where the users can repeat their codewords across sub-blocks, which is referred to as T-fold IRSA. They propose splitting the message into two parts and employing a coding scheme based on a combination of compressed sensing (CS) [58] and LDPC coding at the slot level. Namely, the users select a preamble determined by the first part of the message from a common codebook, and the second part is encoded using an LDPC code and interleaved before transmission. The preamble carries information about the interleaver and the utilized slots for a specific user. At the receiver side, the preamble part is recovered by a non-negative least squares (NNLS) algorithm, and the coded part is decoded by joint belief propagation (BP). The replicas of the recovered codewords are peeled from the other slots via SIC. This scheme improves the performance of [55] by up to 13 dB, and the difference with the random coding bound is less than 8 dB. The general transmission structure in T-fold IRSA based URA solutions is depicted in Fig. 6, where N is the length of the transmission frame, N_s is the slot length, V is the number of slots, w_i is the message of the i-th user, and $x(w_i)$ denotes the transmitted signal of the i-th user.

T-fold IRSA is also utilized in [59] by replacing the LDPC codes with polar codes and utilizing joint successive cancellation decoding [60], which provides a significant performance gain over [57] due to the superiority of the polar codes to the LDPC codes in the short blocklength regime. In [61], the authors employ random signatures to convert the equal-gain channel into a more favorable i.i.d. fading scenario along with trellis-based codes in a T-fold ALOHA structure. Their proposed approach provides a low-complexity solution with moderate performance. Moreover, the authors in [62] employ finite-field coding in a slotted structure. More specifically, the users encode their data using a non-binary LDPC code and transmit it after repetition, zero-padding, and interleaving. At

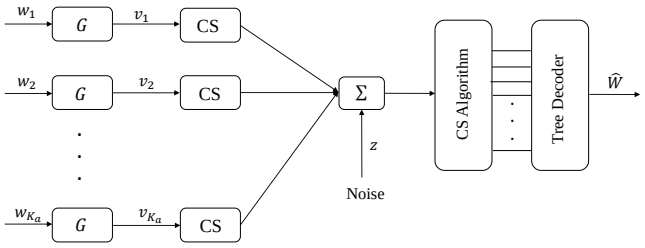


Fig. 7: Encoding-decoding structure in a coded compressed sensing (CCS) approach.

the receiver, intra-slot decoding is performed using a joint belief propagation (BP) algorithm over a non-binary factor graph. Their proposed approach improves the performance of the previously proposed slotted URA schemes for high active user loads.

2) ***CCS-based schemes****:* Another main approach in URA over GMAC is called coded compressed sensing (CCS) [63]–[69], where the messages are divided into sub-blocks and each sub-block is first encoded by an outer code by appending parity bits before being mapped to a CS codebook for inner encoding. The encoded sub-blocks are then transmitted in different slots. At the receiver side, each sub-block is recovered via compressed sensing algorithms, and the recovered segments are stitched together by an outer code, where a tree code is typically used. The general encoding-decoding structure in coded compressed sensing is illustrated in Fig. 7, where G is a random binary matrix performing tree encoding, v_i is the output signal of the tree encoder for the i-th user, and $\hat{W}$ is the set of decoded messages. More specifically, the encoding process is performed in two phases: 1) In the tree encoder, B bits of information are divided into S parts, and all of these segments, except the initial one, are appended additional parity-check bits; 2) in the CS encoder, each S sub-message is mapped to a CS codebook to select a sequence for transmission. Hence, the received signal is comprised of S slots, with each slot being a superposition of the sequences from K_a users. Note that no parity-check bits are appended to the initial segment of the tree encoder since this segment is used as a root node in the tree decoding process [63]. The decoder also functions in two stages: 1) In the initial stage, K_a sequences associated with the K_a active users are identified within each slot. 2) In the second stage, called outer decoder or tree decoder, with the assistance of the unique parity bits, the sub-messages corresponding to the detected sequences are combined to create a message of length B.

Several variations of the CCS approach are examined in the URA literature. Specifically, in [63], the sub-blocks are first encoded by an outer tree code and then mapped to the columns of a common sensing matrix for inner encoding. In [64], sparse regression codes (SPARCs), a class of channel codes that can achieve channel capacity of a point-to-point AWGN channel under approximate message passing (AMP) decoding, are exploited as inner codes to improve the performance of the scheme in [63], while the outer code is again a tree code. The authors in [65] further improve the performance of the approach in [64] by introducing a dynamic interaction between the inner AMP and the outer tree decoder. SPARCs are also employed in [66], where the sub-blocks are CRC-encoded and connected via block Markov superposition transmission. At the receiver, a hybrid decoder combining SIC and AMP is used for inner decoding, with the addition of tree decoding. Furthermore, the authors in [67] employ constant-weight codes as outer codes with the use of Gabor frames to design the inner CS coding matrix, and a concatenation of Raptor codes and tree codes is used as the outer code in [68]. Finally, in a recent work [69], the authors propose a CCS-based scheme eliminating the tree code by using the encoded sub-message in the first slot as a user identifier. This reduces the parity bits and the resulting rate loss, offering the best performance among the CCS-based schemes on GMAC and outperforming the other solutions for $K_a \geq 225$.

3) ***Random spreading-based schemes****:* In another design strategy for URA over GMAC, there is no slotting, and the active users utilize the entire transmission frame. In some works that follow this strategy, the idea is to spread the channel-coded bits throughout the transmission frame using random signatures. The random spreading idea was first introduced in [70], where the message is divided into two parts: the first part determines the signature sequence, while the second part is encoded by a polar code. At the receiver, first, the selected signature sequences are detected by a correlation-based energy detector. Then, an iterative decoding algorithm consisting of MMSE filtering followed by single-user polar decoding is employed to recover the channel-coded bits. Moreover, SIC is applied at the end of each iteration. The general decoding process of random spreading-based solutions is illustrated in Fig. 8, where $\mathbf{Y}$ is the received signal, $\mathbf{Y}_t$ is the residual of the received signal at the t-th iteration, and D^t, β_i^t, and $\hat{v}_i^t$ are the set of detected signature indices, LLR vector, and the message estimate of the i-th user at the t-th iteration, respectively.

In [71], the authors modify the scheme in [70] by replacing the polar codes with polarization-adjusted convolutional (PAC) codes, with a slight performance improvement and reduced implementation complexity. The performance of the scheme in [70] is improved in [44] by allowing users to randomly select a power level from a predefined set. This approach, called *power diversity*, creates an artificial fading-like scenario, where the users with higher power levels are decoded first with a high probability, and those with lower power levels are decoded next after SIC. This results in a performance improvement over [70] for $K_a \geq 150$. In [72], the authors propose to encode the data as the Kronecker product of two-component codewords, and employ an iterative decoder based on bilinear generalized AMP to decompose the Kronecker product and a soft-in-soft-out decoder for the individual components. The authors in [73] replace the polar codes in [70] with LDPC codes to exploit the soft information produced by the belief propagation decoder to perform soft-input-soft-output (SISO) MMSE filtering. This approach enhances decoding performance by leveraging information from partially decoded codewords, resulting in the best performance among schemes based on random spreading. Moreover, enhanced spread spectrum ALOHA (E-SSA) [74], an asynchronous RA protocol that combines ALOHA with

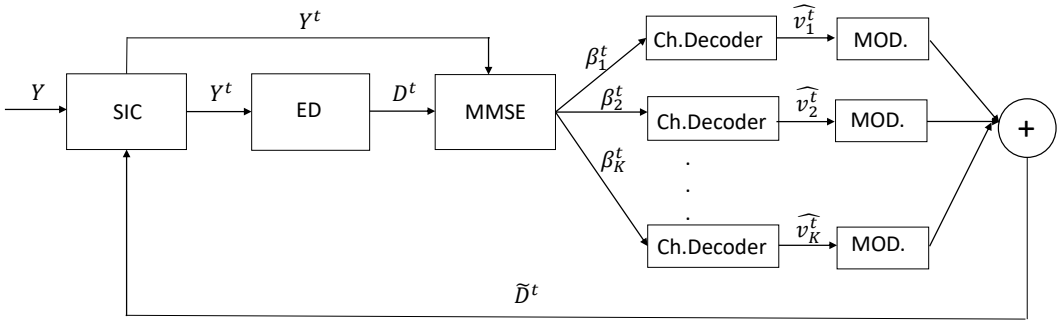


Fig. 8: A typical decoding structure in a random spreading approach.

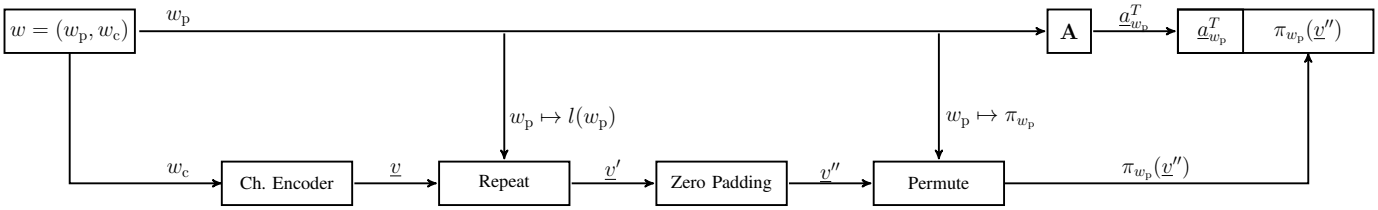


Fig. 9: Encoding process of sparse IDMA [77].

direct sequence spread spectrum, is adapted for URA over GMAC in [75], where a wrap-around version of E-SSA is employed to adapt it for a synchronous URA setup. In a recent work, random spreading with higher order pulse amplitude modulation (PAM) such as 8-PAM or 16-PAM is studied in [76]. The authors propose an AMP-based solution combined with expectation maximization and multiple measurement vectors, demonstrating that their proposed solution can support up to 1100 active users. This is the highest number of supported users reported in the literature for URA over GMAC for the practical PUPE value of $\epsilon = 0.05$.

*4) **IDMA-based schemes**:* In random spreading-based solutions, each active user occupies the entire frame by random spreading of channel-coded bits for data transmission. In some other works, active users distribute their codeword bits across the transmission frame in various ways. For instance, in the sparse interleave-division MA (IDMA) approach [77], the message bits and the frame are divided into different parts. The users transmit a preamble sequence dictated by the first part of their message in the preamble part of the frame. The remaining bits are encoded using an LDPC code, zero-padded, and permuted by an interleaver determined by the preamble to be transmitted over the rest of the channel frame. This encoding process is illustrated in Fig. 9. At the receiver, after decoding the preamble part using the least absolute shrinkage and selection operator (LASSO), a single joint Tanner graph is employed for message recovery, leveraging the information on the location of transmitted symbols obtained from the identified preamble. In a recent sparse IDMA-based scheme proposed in [78], the authors employ AMP for the recovery of the preamble part and convolutional coding in the data part, which improves the performance of [77] in high active user loads.

A similar approach to the sparse IDMA in [77] is adopted in [79], employing polar codes; however, different code lengths are used among the users, and TIN is utilized as the decoding algorithm. Moreover, in [80], the message bits are first encoded by a high-rate error correction code. The encoded bits are divided into two parts, with the first part being mapped to a column of a common sensing matrix, which determines the interleaver and signature to be utilized in the second part after its repetition. The receiver utilizes AMP to recover the first part, followed by a multiuser decoding algorithm for the second part.

While the works mentioned above consider the synchronous transmission, a fully asynchronous (i.e., with an unbounded time horizon) URA setup is studied in [81], where each user transmits a preamble (from a pool of available ones) for timing acquisition followed by its payload that is interleaved and multiplied by a signature similar to [80]. The receiver employs a sliding window decoding approach where correlation-based arrival time detection, multiuser decoding, and SIC operations are iteratively performed.

*5) **ODMA-based schemes**:* Another way to distribute the user codeword bits onto the transmission frame is proposed in [82], where the authors employ on-off division MA (ODMA) [83] in the context of URA. In this approach, the active users exploit a small disjoint portion of the transmission frame to transmit their data encoded by a repeat-accumulate code according to a randomly selected transmission pattern, with the remaining time instances being idle. At the receiver, the BS first recovers the transmission patterns and then employs multiuser detection and decoding over a sparse factor graph. The authors in [84] investigate the ODMA technique in URA further, namely, they employ polar codes to improve the performance in the low active user load regimes combined with a

Table III: A summary of URA solutions for GMAC

Paper	Approach	Description	Channel Code	Synchron. vs. Asynch.
Polyanskiy, 2017 [11]	-	Information-theoretic formulation of URA, and Random coding achievability and converse bounds.	-	-
Ordentlich et al., 2017 [55]	slotting	T-fold ALOHA-based concatenated coding scheme.	BCH	Synchron.
Vem et al., 2019 [57]	slotting	T-fold IRSA-based scheme employing preamble transmission, channel coding, joint BP decoding, and SIC.	LDPC	Synchron.
Marshakov et al., 2019 [59]	slotting	T-fold IRSA-based scheme employing preamble transmission, channel coding, joint successive cancellation decoding, and SIC.	Polar	Synchron.
Amalladinne et al., 2020 [63]	CCS	Concatenated coding scheme employing inner CS and outer tree coding.	-	Synchron.
Amalladinne et al., 2022 [65]	CCS	Concatenated coding scheme employing inner SPARC and outer tree codes. Interaction between inner AMP and outer BP decoding is proposed.	-	Synchron.
Nassaji et al., 2024 [69]	CCS	Concatenated coding scheme employing SPARCs and CRC encoding. Eliminating the outer tree code is proposed to improve the performance.	-	Synchron.
Pradhan et al., 2020 [70]	random spreading	Spreading-based scheme combining correlation-based signature detection, MMSE symbol estimation, single-user decoding, and SIC.	Polar	Synchron.
Ahmadi et al., 2020 [44]	random spreading	Spreading-based scheme employing similar operations with [70] and improve its performance with power diversity.	Polar	Synchron.
Han et al., 2021 [72]	random spreading	Spreading-based scheme combining Bayesian matrix factorization, BCJR decoding, and SIC at the receiver side.	Convolutional	Synchron.
Pradhan et al., 2021 [73]	random spreading	Spreading-based scheme combining energy detection, soft MMSE estimation, soft cancellation, single-user decoding, and SIC.	LDPC	Synchron.
Pradhan et al., 2022 [77]	IDMA	Spreading-based scheme employing preamble transmission, joint Tanner graph decoding, and soft cancellation.	LDPC	Synchron.
Karami et al., 2024 [81]	IDMA	Spreading-based scheme employing correlation-based timing detection, multiuser decoding and SIC.	-	Asynch.
Yan et al., 2023 [82]	ODMA	ODMA-based scheme employing blind pattern detection, joint graph decoding, and soft cancellation.	RA	Synchron.
Ozates et al., 2024 [84]	ODMA	ODMA-based scheme employing energy test for pattern detection, single-user decoding, and SIC.	Polar	Synchron.
Yan et al., 2024 [85]	ODMA	ODMA-based scheme employing blind pattern detection, joint graph decoding, collision resolution, and soft cancellation.	Non-binary LDPC	Synchron.
Ozates et al., 2025 [87]	ODMA	ODMA-based scheme employing nested sliding windows, timing detection by ODMA patterns, single-user decoding and SIC.	Polar	Asynch.

simpler pattern detection algorithm. Another high-performing solution is proposed in [85], which improves the scheme in [82] by employing a non-binary LDPC code. Moreover, an analytical power division method that is formulated based on TIN-SIC is proposed in [86], which achieves the same performance as that of [85] using 5G-NR LDPC codes instead of non-binary LDPC coding. Furthermore, the authors in [87] examine the fully asynchronous setup introduced in [81], and propose a preamble-free ODMA-based solution. The receiver utilizes a sliding-window decoding approach with two nested windows. In the inner sliding window, an iterative decoder combining arrival time detection using the ODMA patterns, single-user polar decoding, and SIC is employed, while the outer window enhances the interference cancellation.

C. Summary and Comparison of GMAC Solutions

A summary of the URA solutions for GMAC is presented in Table III, with comparisons among them in different aspects.

*1) **Performance Comparison**:* The performance of the GMAC URA schemes is compared in terms of energy efficiency, the standard performance criterion in URA, in Fig. 10 for $n = 30000$, $B = 100$, and $\epsilon = 0.05$, which are the common set of parameters in the literature. Note that due to the vast number of papers, we have included only the state-of-the-art papers, at least in some regime, and those

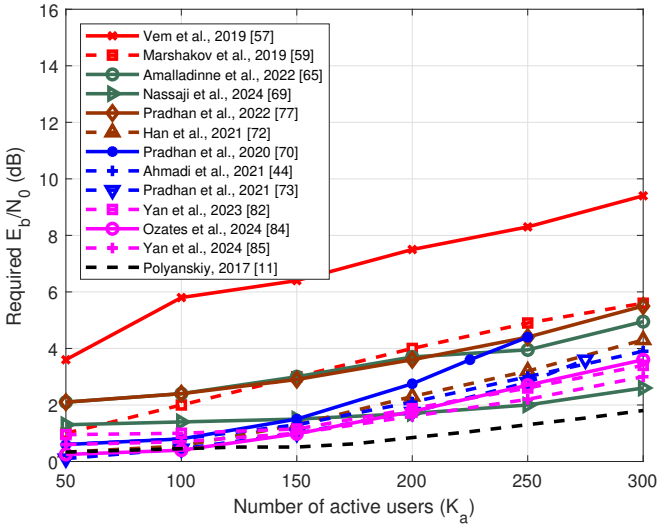


Fig. 10: Performance comparison of the coding schemes in URA over GMAC for $n = 30000$, $B = 100$, and $\epsilon = 0.05$, which are the common set of parameters in the literature.

introducing a new approach when published. The results in Fig. 10 demonstrate that random spreading-based schemes in [44], [70], [73] offer good performance for a low number of active users; however, their performance degrades with increasing the number of active users since each active user occupies the entire transmission frame. On the other hand, ODMA-based approaches in [82], [84], [85] outperform the others, except for the solution proposed in [69], in the high multiuser interference regime due to their sparsity. Meanwhile, they perform well in the low multiuser regime when combined with polar coding. Moreover, the CCS-based approaches can exhibit very good performance for a high number of active users when the rate loss due to the parity bits is alleviated, as shown in [69]. Compared to the achievability bound, the proposed schemes can perform similarly or slightly better for $K_a \leq 100$, and the gap is at most about 0.8 dB for $K_a \leq 300$.

In summary, the performance trends over GMAC reveal a clear trade-off: While random spreading-based schemes are attractive at low user loads, their scalability is limited due to full-frame occupancy and high complexity. Conversely, ODMA-based approaches and recent CCS variants achieve robust performance in high-user regimes, benefiting from sparsity and efficient decoding. These results underscore that practical URA deployment must carefully balance energy efficiency, complexity, and user density, with CCS and ODMA schemes emerging as the most scalable options.

*2) **Computational Complexity Comparison**:* In terms of complexity, slotting is the most beneficial approach among the existing solutions for URA over GMAC, as the complexity of the operations, depending on the user count, increases super-linearly by the number of active users in the slot rather than K_a. In addition, slotting leads to smaller preamble/pilot lengths and a shorter channel code length n_c, thereby decreasing the complexity of the operations. Coded compressed sensing schemes also offer low complexity since slot-wise compressed sensing and low-complexity tree codes are employed in conjunction with each other. Specifically, its complexity order is linear or quadratic in K_a, depending on the regime [63]. ODMA is another low-complexity approach, as only a small fraction of the frame is occupied by each user, which simplifies the iterative multiuser decoding process [82]. More specifically, its complexity order depends linearly on n_c and K_a. On the other hand, random spreading-based schemes where each user occupies the entire frame have the highest complexity, as the complexity order is cubic in K_a due to the MMSE filtering [70].

V. URA over Fading Channels

GMAC is an ideal channel model that does not capture many practical effects of the communication channel. With this motivation, more realistic fading MAC models, where the transmitted signals of the users are attenuated with random fading coefficients, are considered in several works in the URA literature.

There are some challenges in extending and applying the URA over GMAC solutions to fading scenarios. The primary challenge in fading MAC is estimating the channel coefficients at the receiver side and designing the decoder to account for fading. In the literature, two approaches are proposed to address this problem; the first is to embed the channel estimation into the decoding algorithm without requiring a separate channel estimation step (e.g., [88], [93], [115]). The second approach is to exploit a part of the frame for pilot transmission, perform explicit channel estimation using the received signal in the pilot part, and utilize the estimated coefficients for detection in the data part (e.g., [12], [102], [120]). While most schemes developed for single-antenna fading channels, and the CCS-based and tensor-based schemes in URA over MIMO fading channels opt for the former, there is also a vast literature on pilot-based schemes. The main algorithms applied for channel estimation in pilot-based schemes are MMSE [12], [14], [84], [119], [120], [124] and AMP [102], [121], [122]. Note that in most URA works over fading channels, quasi-static fading is considered.

In this section, we first review the bounds on the performance of URA solutions taking into account channel fading that have been provided in the literature. We then cover the URA solutions that present low-complexity coding schemes to achieve those bounds. To this end, in order to explain the URA approaches that take into account channel fading in a simpler setup for better understanding, we first cover the single-antenna fading solutions and then turn our attention to the MIMO fading ones.

A. Performance Bounds

In [88], [91]–[93], the authors derive some performance bounds for the URA framework over a Rayleigh fading channel. Specifically, the authors in [88] present a converse bound based on the results in [89] and the meta-converse bound in [90]. These results numerically demonstrate that the E_b/N_0 requirements for a fading channel are necessarily higher than those for the AWGN channel. They also argue that, for typical system parameters in this setup, a Fano-type converse bound is worse than the proposed one.

In [91] and [92], achievability and converse bounds are proposed for the cases of known and unknown CSI in an asymptotic setting, where the number of users grows linearly with the blocklength. The authors in [88] provide both asymptotic and non-asymptotic achievability and converse bounds for the case of unknown CSI. In [93], authors derive an achievability bound for the URA scheme over the Rayleigh fading channel by considering a slotted structure and perfect SIC in their analysis, which is tighter than the other bounds. However, since the channel coefficients used in an SIC process are estimated, hence imperfect, the assumption of a perfect SIC is not realistic, which makes the result in [93] an approximation rather than an achievability one.

Some achievability and converse bounds for URA over MIMO Rayleigh fading channels are obtained in [94], [95]. Among the developed results, the ML-based achievability bound is obtained by employing a projection decoder developed in [96]. Moreover, the results in [94] are extended to the case of random and unknown number of active users in [97].

B. Single-Antenna Fading Channels

We first examine the URA solutions over fading channels that assume a single-antenna receiver. The works in [77], [88], [93], [98]–[104] introduce practical coding schemes tailored for this setup. Among these works, some assume that the received signals from all users are synchronized [77], [88], [93], [98], [102], [104], while others take into account asynchronous transmission [99]–[101], [103]. To estimate the delay of asynchronous signals, [99], [100], [103] use OFDM-based transmission to convert the time delay to a phase shift in the frequency domain, while the authors in [101] consider time-domain transmission where the delay of a signal is detected via a correlation-based approach.

Focusing on Rayleigh fading channels, we present existing solutions for URA over fading channels, categorized based on their general approach.

*1) **CCS-based schemes**:* As explained in the previous section, CCS was initially proposed for URA over GMAC. In subsequent literature, it has also been applied to the fading scenario, see [98], [101], [104]. In [98], the authors replace the tree code of the GMAC CCS scheme in [63] by list recoverable codes that are capable of correcting t errors. They propose two outer code constructions: The first is a modified tree code with the ability to recover t errors (called the t-tree), and the second is a Reed-Solomon (RS)-based code. In [104], the outer encoded message bits are partitioned into sub-blocks, and each sub-block is transmitted in a different slot after CS-type encoding, with AMP at the receiver. Moreover, in [101], the authors utilize CCS for the asynchronous URA model. More specifically, during the encoding stage, they append T_z additional zeros to the end of each CS sequence, where T_z represents the maximum possible delay of the system. At the receiving end, the decoder takes into account all possible delays, ranging from one symbol to a maximum of T_z-symbol delays during the decoding process.

*2) **Joint decoding-based schemes**:* In [88], a joint decoding approach is employed. For encoding, this scheme partitions the frame into distinct slots. Each user transmits its LDPC codeword through a randomly chosen slot. As previously mentioned, slot allocation offers the significant benefit of mitigating the per-slot interference by distributing users across slots. This scheme utilizes an iterative belief propagation (BP) decoder, as illustrated in Fig. 11. The graph incorporates four distinct node types: variable nodes (in red); check nodes (in blue), which constitute the traditional LDPC decoder; *functional* nodes (in green) corresponding to the symbols of the received signal; and a fourth type (in magenta) corresponding to the fading coefficients. This graph allows them to update the estimated values of both the channel coefficients and the symbols simultaneously for each user.

*3) **TIN-SIC-based schemes**:* A lower-complexity alternative to joint decoding is the TIN-based decoding. Similar to GMAC, the TIN solution for fading channels operates under the assumption that a single user is present and treats the signals of other interfering users as noise. Consequently, this approach results in an exceptionally low average SNR because it encounters interference-plus-noise as the system's overall noise level, leading to inferior performance. However, when combined with an appropriate channel code and SIC, the performance can be improved.

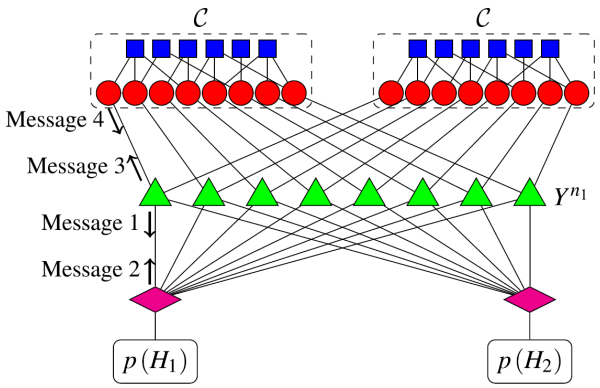


Fig. 11: Iterative BP decoder for a two-user LDPC-based scheme with joint decoding [88].

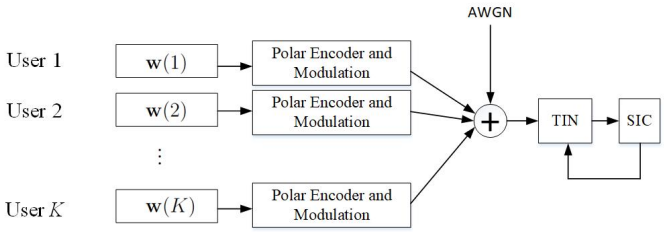


Fig. 12: Encoding and decoding processes of TIN-SIC scheme with polar codes in [93] in a slot with K users.

TIN-SIC with LDPC coding: In [99], building upon the synchronous schemes presented in [88], the authors present a low-complexity coding scheme which incorporates OFDM techniques to mitigate the effects of the asynchronous transmission by converting the time delays into phase shifts, and employing SIC to decode the user packets. Further numerical results for this scheme are presented in [100].

TIN-SIC with Polar Coding: In [93], a TIN-SIC solution in conjunction with polar coding is presented, where the frame in this approach is divided into S slots. Each user randomly selects a slot to transmit its polar-coded and modulated signal. For decoding, the scheme considers a TIN strategy followed by SIC. Since different users in a slot experience different fading coefficients, some of them (those with larger fading coefficient amplitudes) are more likely to be decoded by TIN. After decoding the strongest user's signal while treating the others as noise, the channel coefficient is estimated using orthogonal matching pursuit (OMP). Given the estimated channel coefficient and the decoded signal of a user, its contribution can be removed from the received signal using SIC. This strategy continues until no signal is successfully detected during an iteration. It is important to emphasize that the determination of whether a decoded message is successful relies on the use of a cyclic redundancy check (CRC) message sequence. The encoding and decoding process of the scheme with polar coding with SIC in a slot with K users is illustrated in Fig. 12.

*4) **Alternative solutions**:* Some other solutions also focus on URA over fading channels with a single-antenna receiver. For instance, in [102], the active users utilize the entire trans-

Table IV: A summary of URA solutions for fading channels with a single-antenna receiver.

Paper	Approach	Description	Synchron. vs. Asynch.	Channel Estimation	Channel Code
Kowshik et al., 2019&2021 [91], [92]	-	Achievable and converse bounds for known and unknown CSI in an asymptotic setting + comparing with CDMA, TDMA, and FDMA.	Synchron.	-	-
Kowshik et al., 2019 [100]	TIN-SIC	T-fold ALOHA-based scheme employing OFDM, channel coding, and SIC.	Asynch.	Embedded	LDPC
Andreev et al., 2020 [93]	TIN-SIC	T-fold ALOHA-based scheme employing channel coding, single-user decoding, and SIC.	Synchron.	OMP	Polar
Kowshik et al., 2020 [88]	joint decoding	Achievability and converse bounds for known and unknown CSI + a scheme employing slotting, channel coding, and joint BP decoding.	Synchron.	Embedded	LDPC
Amalladinne et al., 2019 [101]	CCS	Extension of CCS in [63] to accommodate asynchronous scenarios.	Asynch.	Embedded	-
Andreev et al., 2022 [98]	CCS	Improving CCS in [63] by replacing the outer tree code with 1) a code capable of correcting t errors, and 2) a Reed–Solomon code.	Synchron.	Embedded	RS
Yun et al., 2024 [104]	CCS	A slotted scheme employing CCS and AMP in a block fading model.	Synchron.	Embedded	-
Nassaji et al., 2022 [102]	Spreading	A non-slotted scheme employing preamble transmission, channel coding, repetition, permutation, scrambling, and SIC.	Synchron.	AMP	Polar
Ozates et al., 2024 [84]	ODMA	ODMA-based scheme employing gOMP for pilot and pattern detection, single-user decoding and SIC.	Synchron.	LMMSE	Polar
Ozates et al., 2023 [103]	OFDM+ SIC	A slotted scheme in the frequency-selective channel employing OFDM, pilot transmission, channel coding, and SIC.	Asynch.	OMP	Polar

mission frame to transmit their data. Specifically, the frame is divided into pilot and data parts. The active users select a pilot sequence from a common codebook for transmission in the pilot part, and they repeat their codeword bits for the remainder of the transmission time frame, with permutation and scrambling applied in the data part. At the receiver, AMP is employed for joint pilot detection and channel estimation, and the data part is recovered by iterative data estimation and interference cancellation using the repeated codeword bits.

The ODMA approach is also applied to the single-antenna fading scenario in [84] by allocating a portion of the frame for pilot transmission while a polar codeword is transmitted in the rest of the frame by ODMA. At the receiver side, the selected pilots and patterns are detected by a generalized OMP (gOMP) algorithm, followed by linear MMSE (LMMSE) channel estimation, single-user decoding, and SIC to recover the transmitted messages. Furthermore, the authors in [103] study URA over frequency-selective channels and propose OFDM to mitigate adverse channel effects combined with CS-based activity detection and TIN-SIC for message recovery at the receiver side.

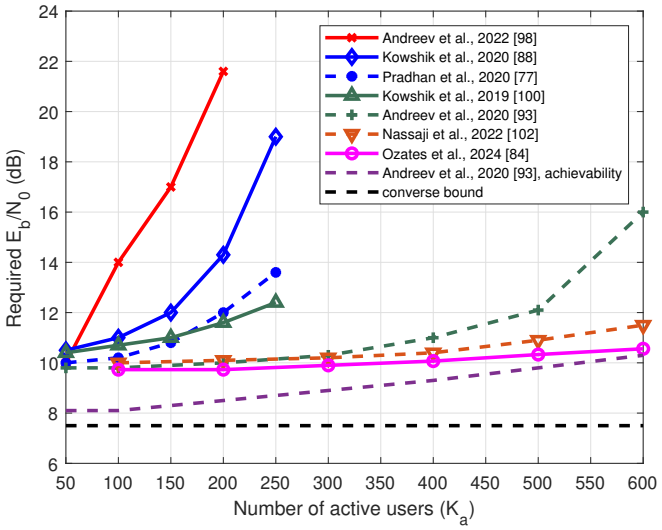


Fig. 13: Performance comparison of the coding schemes for URA over fading channel with a single-antenna receiver, with a target PUPE of 0.1, $n = 30000$, and $B = 100$ (the common set of parameters in the literature).

C. *Summary and Comparison of Single-Antenna Fading Solutions*

A summary of URA studies over fading channels involving a single-antenna receiver is given in Table IV, where *Embedded* under channel estimation column refers to the case that estimation of the fading coefficient is included as a part of the decoding algorithm.

*1) **Performance Comparison:**:* The performance of different coding schemes developed for URA over single-antenna fading channels is compared in Fig. 13 assuming $n = 30000$, $B = 100$, and $\epsilon = 0.1$. The reason that some diagrams seem not to be fully plotted is that those solutions do not work after a certain active user load, i.e., they cannot achieve the target PUPE, no matter how much the E_b/N_0 is increased. As can be observed from Fig. 13, the CCS-based scheme in [98] exhibits lower energy efficiency compared to the other alternatives, as it employs a suboptimal coding scheme, whereas the others use strong channel codes, such as LDPC or polar codes. The performance of the LDPC-based scheme with joint decoding in [88] is inferior to that of the ones in [77], [100], and [93], which can be attributed to the absence of an efficient SIC block in this LDPC-based approach. On the other hand, the TIN-SIC approach in [100] and [93] offers a better performance as it allows the receiver to recover the messages of users with a stronger fading coefficient, subtract their effects, and decode the remaining users in subsequent

iterations, especially when combined with polar coding as in [93]. We also note that the TIN-SIC scheme in [93] is outperformed by the spreading-based solution in [102], and the ODMA-based scheme employing polar codes in [84], which offers the best performance among the proposed schemes, due to the sparsity of the transmissions.

Overall, the results in fading channels show that CCS-based solutions lack a strong channel code to provide a high energy efficiency. Simple TIN-SIC solutions offer a good performance in low user loads, however, they demonstrate the limitations of weak interference management as device density grows. The comparative analysis suggests that practical single-antenna fading URA systems will favor ODMA and spreading-based approaches, which can be flexibly adapted to varying user loads without incurring excessive complexity.

*2) **Computational Complexity Comparison**:* We now compare the computational complexity of the URA solutions over single-antenna fading channels. The tree-code-based CCS approach has a low complexity, which is linear in the size of the list of messages at each decoding step and exponential in the number of message bits in each slot [98]. In addition, the decoding operations in [98] can be carried out by additions, which are easier to implement on hardware compared to multiplications. The complexity of the T-fold ALOHA-based schemes employing LDPC with joint decoding [88] or polar coding with TIN-SIC [93] is mainly determined by the parameter T, that is, the number of users that the decoder tries to recover in each slot. Note that TIN-SIC is more preferable than joint decoding since it utilizes single-user decoding, which makes it generally less complex and more parallelizable. Spreading-based and ODMA-based schemes, proposed in [102] and [84], respectively, have higher complexities compared to their counterparts in exchange for their superior performance. For instance, the complexity of MMSE filtering for fading coefficient estimation in [84] is cubic in K_a.

D. MIMO Fading Channels

We now turn our focus to the URA solutions developed for MIMO fading channels, which can be categorized into five groups: CCS-based, tensor-based, pilot-based, ODMA-based, and Bayesian approaches, with inevitable overlaps among them.

*1) **CCS-based schemes**:* In CCS-based algorithms for URA over MIMO fading channels [105], [106], the bit sequence of each user is divided into S sub-messages, which are then mapped to codewords in pilot codebooks with lower dimensions to obtain S different pilots, transmitted through S different slots. At the receiving end, K_a sub-messages are detected within each slot using the ML [105] or hybrid generalized approximate message passing (HyGAMP) [106] technique, and they are then combined to form the complete length-B bit sequence for each user. To establish connections among these sub-messages, the authors in [105] incorporate parity-check bits into each sub-message, while [106] relies on the correlation between channel coefficients of each user in different slots (assuming a quasi-fading channel model). Eliminating the need for additional parity bits enhances the performance of [106] compared to [105].

Eliminating the parity bits in [105] turns the problem into an uncoupled compressed sensing one, which is also considered in some recent works [107]–[110]. To stitch the decoded sub-messages, the authors in [107] employ a joint Bayesian decoding approach that exploits the channel characteristics, while geometric similarities between the codeword covariance matrices are utilized in the decoding algorithm in [108], and the similarity of the estimated channel vectors is used in [109]. The structure of the inner SPARC code allows parallel single-user outer decoding of user messages in [110], which mitigates the need for the parity check bits. An uncoupled structure is employed in [111] to design a CCS-based URA scheme for near-field communications. Furthermore, a two-phase CS coding scheme is proposed in [112], where a long sub-block recovered in the first phase is used for channel estimation, and very short-length sub-blocks are recovered in the second phase by employing a mutually orthogonal codebook and a hybrid AMP in an uncoupled structure. Moreover, the heterogeneous mobility of the massive number of users is considered in [113] via a two-codebook aided URA scheme with covariance-based maximum likelihood detection (CB-MLD) at the receiver, where the static users employ a larger codebook while the mobile users utilize a smaller one in both the codeword length and the number of codewords. Moreover, an uncoupled CS scheme employing polar coding and power diversity in combination with AMP-based decoding and SIC at the receiver side is proposed in [114].

*2) **Tensor-based schemes**:* Another trend in MIMO URA is the tensor-based schemes [115]–[117], where each user sends a rank-1 tensor of order d. Therefore, the received signal is a rank-K_a tensor of order $d+1$ (considering the channel coefficient vector as an extra dimension of the tensor). For decoding, they conduct tensor decomposition on the received signal, followed by de-quantization of the tensors. In particular, in [115], by tensor decomposition, the rank-K_a tensor of order $d+1$ is decomposed into K_a rank-1 tensors of order $d+1$, which is equivalent to separating the signals of K_a users, where the transmitted tensors can be identified by the first d dimensions, and the $(d+1)$th dimension gives the channel vector of the user. Then, by de-quantizing the tensor of each user, its transmitted bit sequence is detected. By employing OFDM transmission, this scheme is extended to the case with timing offsets in [116]. On the other hand, the authors in [117] propose a block-term decomposition for the decomposition of 3-dimensional tensors, where two low-rank factor matrices and a factor vector are utilized in each block term. Moreover, the authors in [118] study tensor-based modulation with soft decoding, and propose a polar-coded tensor-based scheme with iterative Bayesian decoding with feedback at the receiver side. The overall coding structure of tensor-based methods is depicted in Fig. 14.

*3) **Pilot-based schemes**:* Pilot-based techniques constitute another category of algorithms in the context of MIMO URA systems, where the idea is to divide the user bit sequences into pilot and data parts. In the existing solutions, the transmitted signal of each user also consists of two parts, as depicted

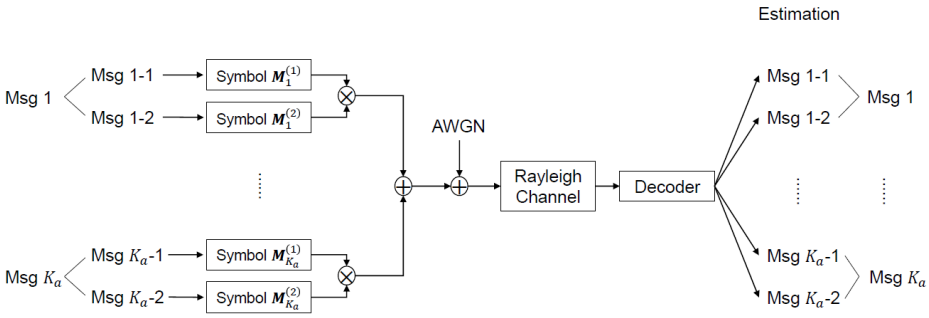


Fig. 14: Encoding and decoding structures of tensor-based schemes for $d = 2$ [117].

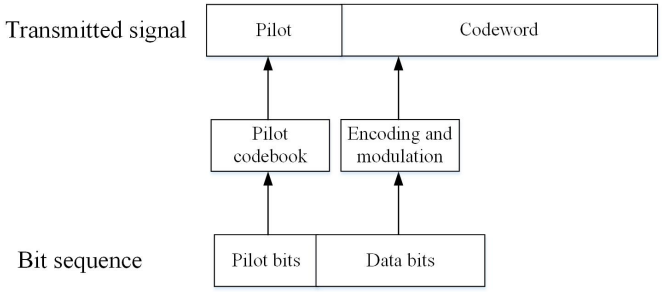


Fig. 15: Encoding structure of pilot-based schemes.

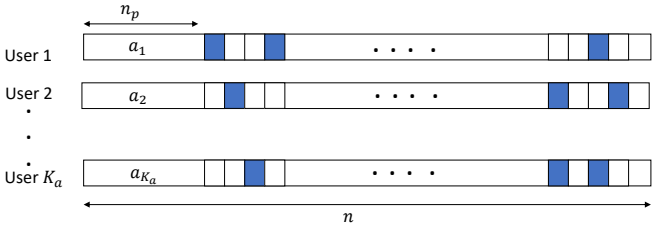


Fig. 16: An illustration of the ODMA idea for the fading scenario with the pilot and data parts. Colored boxes show the utilized time instances in the data part.

in Fig. 15: the pilot part which is obtained by mapping the pilot bits to a codeword in a pilot codebook (a non-orthogonal codebook [12], [14], [119]–[122], or a codebook consisting of multiple stages of orthogonal pilots [13]), and the data part which is encoded by a channel code (a polar [12]–[14], [120] or an LDPC code [121], [122]). For decoding, as shown in Fig. 17, these schemes iteratively employ the following steps:

- At the initial stage, the receiver detects the active pilots using a user activity detection technique, such as gOMP [14], [119], AMP [121], [124], multiple-measurement vector AMP (MMV-AMP) [120], energy detector [12], NP hypothesis testing [13], or CB-MLD [122]. Moreover, an improved version of the NNLS algorithm exploiting multiple measurements for activity detection is proposed in [123].
- The channel vectors corresponding to each detected pilot are estimated by techniques such as MMSE [12], [14], [119], [120], [124], MRC [13] or AMP [121], [122].
- Using the estimated channel coefficients, soft estimates of the transmitted codewords are obtained via decoupling techniques, such as MRC [13], [14], [119], [120], MMSE [12] or AMP [122], which are then used to generate log-likelihood ratio (LLR) values to be fed to single-user decoders [12]–[14], [120], [122], [124] or a joint factor graph [121].
- In most pilot-based schemes, following the SIC approach, the contribution of the successfully decoded codewords is removed from the received signal to decrease the interference level at the subsequent iterations. Note that in [12]–[14], [119], the channel vectors are re-estimated using the successfully decoded codewords for improved interference cancellation at the SIC stage.

*4) **ODMA-based schemes**:* ODMA is a powerful technique that was originally proposed for URA over GMAC. However, it is also extended to the fading scenario for the case of a single-antenna receiver in [84] and a massive MIMO receiver in [119], [125]–[127], [129]–[131]. The main idea in [119] is to allocate a part of the frame for pilot transmission, which facilitates the channel estimation at the receiver side, while the rest of the frame is utilized for data transmission in an ODMA manner as illustrated in Fig. 16. At the receiver side, the gOMP algorithm is employed to detect the active pilots, followed by LMMSE channel estimation, MRC symbol estimation, single-user polar decoding, and SIC to recover the data part. Note that if the channel is a GMAC, the entire frame is used for data transmission. On the other hand, the authors in [125] consider superimposed pilot and data transmission, where the pilot sequence occupies the entire frame and the data signal is transmitted using ODMA. The receiver operation begins with energy detection-aided non-Bayesian pilot detection, and the same steps as in [119] are used for decoding the data part. A slotted ODMA-based scheme combining preamble transmission and convolutional LDPC coding is proposed in [126], with AMP decoding for the preamble and turbo probabilistic data association for multiuser detection of the data part at the receiver side. In another slotted ODMA-based scheme [127], the authors consider uncoupled ODMA where the pilot sequences and transmission patterns are selected by different bits of the message sequence. They utilize low-rank matrix factorization and simultaneous OMP [128] for activity detection and channel estimation, followed by a joint pattern and data detection based on AMP. Moreover, two slotted pilot-free ODMA schemes are proposed in [129] and [130]. While an iterative TIN strategy is employed for joint pattern and data detection in [129], random dictionary

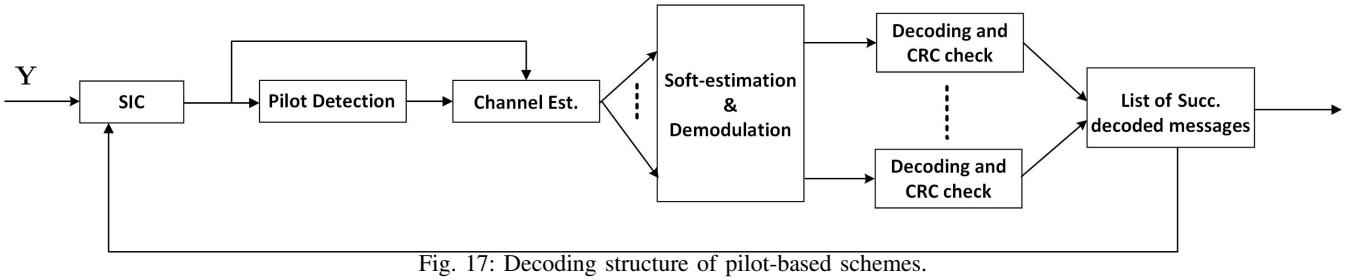


Fig. 17: Decoding structure of pilot-based schemes.

learning is studied in [130] in conjunction with clustering-aided activity detection. Finally, in a recent work [131], the authors present a strategy that is called *Less is More* in a slotted ODMA framework. They propose to transmit a small portion of data using a polar or LDPC code, while embedding the rest of the data into the proposed transmission structure. The transmitted signal consists of a pilot sequence followed by a polar or LDPC codeword. At the receiver side, they employ simultaneous OMP for activity detection followed by zero-forcing symbol estimation, single-user decoding, and SIC. This scheme outperforms the others for fewer than 500 active users, and its performance is slightly better than the achievability bound in [94] when the number of active users is less than 400.

*5) **Bayesian approach**:* In pilot-based algorithms, the channel estimation and data detection tasks are performed independently. Hence, the error from one block is propagated to the next one. To resolve this issue, a new set of MIMO URA schemes has been developed that employ a multi-layer iterative Bayesian decoder to jointly estimate each user's transmitted signal and their corresponding channel coefficients [132], [133]. Specifically, in [132], the message bits are divided into two parts; one part is encoded by a convolutional code, and the other is encoded by index modulation. The transmitted signal is obtained by the Kronecker product of the two components. Additionally, the authors in [133] propose encoding the user payload by SPARCs, where all segments of the message bits of a user are mapped onto one codeword. This eliminates the need for an outer tree code, increasing spectral efficiency.

E. Summary and Comparison of MIMO Fading Solutions

A summary of MIMO URA schemes is provided in Table V, accompanied by the decoding steps of the main approaches summarized in Fig. 18.

*1) **Performance Comparison**:* In Fig. 19, a performance comparison of the proposed coding schemes for URA with a massive MIMO receiver is provided for $n = 3200$, $B = 100$, $\epsilon = 0.05$, and $M_r = 50$, assuming that the channel vectors consist of i.i.d. Rayleigh fading coefficients across the different users and different antennas. The results in Fig. 19 show that, due to the adaptation of strong channel coding algorithms such as LDPC and polar codes, pilot-based schemes in [12]–[14], [119], [120] usually outperform the tensor-based scheme in [115] for which a relatively weak channel code is utilized, while the SKP coding-based scheme in [132], following the Bayesian approach, offers similar performance to the best-performing pilot-based solutions. Note that, to benefit from

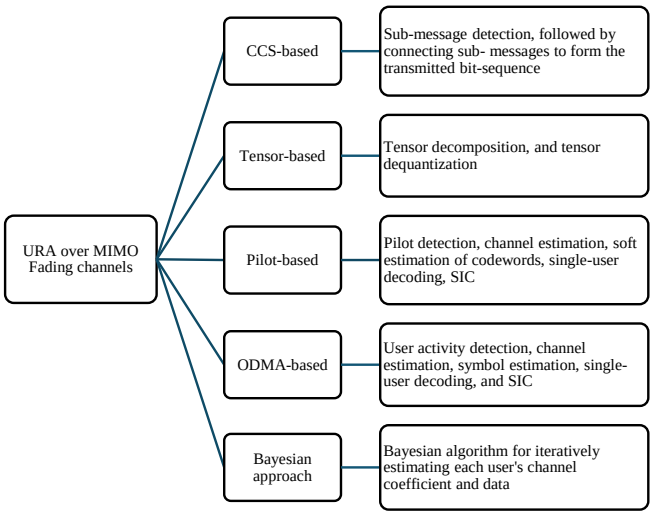


Fig. 18: The decoding steps of the main URA approaches for MIMO fading channels.

increased spatial diversity, the existing URA works on MIMO fading channels assume that the receiver is equipped with a large number of antennas (i.e., the massive MIMO regime), and only a few provide results for a small number of receive antennas [14], [132].

Taken together, the MIMO comparisons highlight that leveraging spatial dimensions is critical for scaling URA to 6G-class scenarios, which indicates that MIMO URA is among the most promising paths toward practical deployment, particularly when combined with structured coding and advanced interference management.

*2) **Computational Complexity Comparison**:* Comparing the computational complexity of the coding strategies for MIMO URA, we observe that CCS exhibits the lowest complexity among them, as its complexity is linear in M_r and the number of common codebook elements [105]. Tensor-based [115] and non-slotted pilot-based [12] schemes have a similar complexity (cubic in K_a). This complexity results from the canonical polyadic decomposition in the former and MMSE filtering for channel estimation in the latter. On the other hand, the complexity of slotted pilot-based schemes [13], [14] is cubic in the number of active users in the slot, which demonstrates the complexity advantage of slotting. Furthermore, the complexity of the Bayesian approach in [132] is linear in K_a, M_r, and n. To summarize, for practical system parameters, CCS has the lowest complexity, while the complexities of slotted pilot-based schemes and the Bayesian approach are similar and smaller than those of tensor-based

Table V: A summary of URA solutions for MIMO fading channels.

Paper	Approach	Description	Synchron. vs. Asynch.	Channel Estimation	Channel Code
Gao et al., 2023 [94]	-	Provides achievability and converse bounds.	-	-	-
Fengler et al., 2021 [105]	CCS	Concatenated coding scheme where the inner pilots and channel vectors are jointly recovered by ML detection and the recovered segments are stitched using a tree code.	Synchron.	ML	-
Shyianov et al., 2021 [106]	CCS	CCS-based scheme where a variation of AMP is employed for inner CS decoding and channel estimation. The recovered sub-messages are stitched exploiting correlation, which improves the energy efficiency.	Synchron.	HyGAMP	-
Jiang et al., 2021 [112]	CCS	A two-stage scheme with a long sub-message in the first stage is used for channel estimation, and very short sub-messages are recovered using orthogonal codebooks in the second stage.	Synchron.	AMP	-
Wang et al., 2022 [151]	CCS	A concatenated coding scheme employing SPARCs, tree coding and OFDM.	Asynch.	Embedded	-
Decurninge et al., 2021 [115]	tensor-based	Users' signal separation and detection is done by tensor decomposition.	Synchron.	Embedded	BCH
Luan et al., 2022 [117]	tensor-based	A new tensor-based modulation scheme in two versions: one employing QPSK encoding, and the other implementing Grassmann encoding. It shows improved energy efficiency over the tensor-based scheme in [115].	Synchron.	Embedded	BCH
Fengler et al., 2022 [120]	pilot-based	A non-slotted scheme employing MMV-AMP pilot detection, LMMSE channel estimation, and single-user decoding. Transmit signal structure: non-orthogonal pilot + polar codeword.	Synchron.	LMMSE	Polar
Ahmadi et al., 2024 [13]	pilot-based	A slotted scheme employing NP hypothesis testing pilot detection, MRC channel estimation, single-user decoding, and SIC. Transmit signal structure: Multiple stages of orthogonal pilots + polar codeword.	Synchron.	MRC	Polar
Gkagkos et al., 2023 [12]	pilot-based	A non-slotted scheme employing energy test-based pilot detection, LMMSE channel estimation, single-user decoding, and SIC. Transmit signal structure: non-orthogonal pilot + randomly spread polar codeword.	Synchron.	LMMSE	Polar
Ozates et al., 2024 [14]	pilot-based	A slotted scheme employing gOMP pilot detection, LMMSE channel estimation, single-user decoding, and SIC. Transmit signal structure: non-orthogonal pilot + polar codeword.	Synchron.	LMMSE	Polar
Ozates et al., 2024 [119]	ODMA	A non-slotted scheme employing gOMP pilot detection, LMMSE channel estimation, single-user decoding, and SIC. Transmit signal structure: non-orthogonal pilot + polar codeword spread by ODMA.	Synchron.	LMMSE	Polar
Han et al., 2023 [132]	Bayesian	SKP coding-based scheme employing BiG-AMP for outer matrix factorization and BCJR for decoding of the convolutional code.	Synchron.	Embedded	Convolutional
Jiang et al., 2023 [133]	Bayesian	A Bayesian scheme employing SPARCs for transmission and a three-layer message passing algorithm for decoding.	Synchron.	Embedded	-
Li et al., 2025 [152]	Bayesian	A Bayesian scheme employing SPARCs, LDPC, and OFDM, where both timing and frequency offsets are considered.	Asynch.	GB-CR	LDPC

and non-slotted pilot-based approaches.

F. Existing URA Solutions on Alternative Setups

Most URA schemes in the literature consider direct links between all the users and the BS; however, in certain environments, the direct link between some users and the BS may be blocked or significantly attenuated. In these cases, one solution is to employ RIS, a promising technology that can furnish the URA system with high spectral efficiency and energy savings. Specifically, it has been shown that a passive RIS equipped with many low-cost passive elements, which can intelligently tune the phase-shift of the incident electromagnetic waves, and reflect them in a desired direction without any amplification, can improve the efficiency of a URA system by enabling line-of-sight paths between the transmitters and the receiver in problematic environments with many blocking obstacles [134], [135]. More specifically, the RIS-based schemes in [136], [137] operate in two phases. In the first phase, known as the RIS configuration phase, the BS detects the active pilots and estimates their respective channel coefficients, using which the BS suitably selects the phase shifts of the RIS elements. In the second phase, known as the data phase, the transmitted messages of active users are decoded using SIC-based decoders. The RIS-based URA solution in [136] outperforms the one in [137], with a power savings of up to 10 dB. This improvement is due to the employment of a slotted transmission structure, which reduces the interference, as well as a more suitable metric in the RIS design in [136], where the overall phase-shift matrix is obtained by minimizing the decoding error probability, in contrast to maximizing the minimum channel gain among active users for the RIS design in [137]. The authors enhance the RIS design metrics presented in [136] and further improve upon them in [138], obtaining bounds on the system performance.

Conventionally, in MIMO URA systems, ideal hardware units are assumed at both the BS and the UE sides. However, due to the massive number of users and antenna elements, the hardware needs to be inexpensive, and these elements are prone to hardware impairments (HWI) such as I/Q imbalance, phase noise, and quantization errors. The problem of URA with HWI is considered in [139], where the solution developed in [14] is adapted to the presence of HWI.

While most URA works consider transmission over i.i.d. Rayleigh block-fading channels, some investigate other chan-

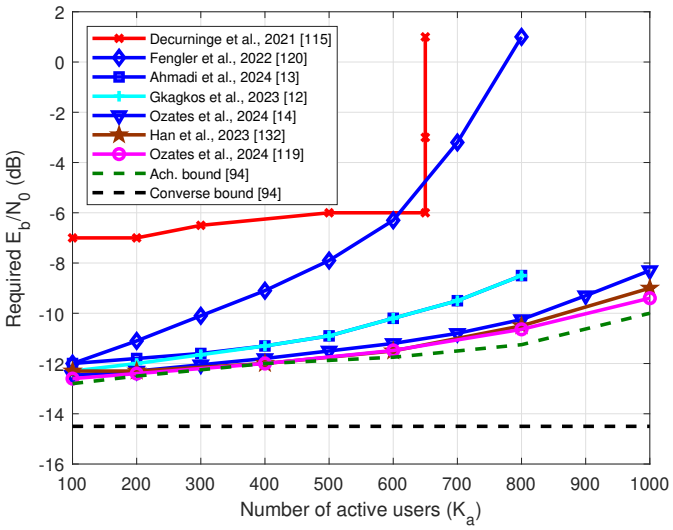


Fig. 19: Comparison of URA schemes for MIMO fading channels with a target PUPE of 0.05, $n = 3200$, and $B = 100$ (the common set of parameters in the literature).

nel models. A channel model with one single dominant path between a user and the BS is considered in [140], and an approach similar to the one in [124] is proposed, exploiting the user location geometry by linking the transmission features such as the employed preamble or interleaver to the angle of arrival of the user. The dynamics of the user states and channels are taken into account in a slotted structure in [141], namely, the users can start or end their transmission in any slot, and they can be active in a subset of the slots, and the channel coefficients change from one slot to another. The authors design a two-step decoding algorithm with an inner AMP and an outer variational message passing. Type-based MA, where each device observes the state of a process and transmits it to a common receiver, is investigated in an unsourced setup in [142], where the receiver's goal is to estimate the states and their multiplicity. An achievability bound quantifying the performance of type-based MA is derived in [143]. Moreover, the authors in [144] extend the framework in [142] to the case of fading channels, where a cell-free massive MIMO setup is considered, and a solution exploiting location-based codebook partitioning is proposed. The problem of user identification and authentication is applied to URA in [145] with the aid of a message authentication code. Moreover, the authors in [146] propose a URA scheme that combines physical and medium access layers where the medium access layer slotting carries information. The results show that carrying information by the medium access control layer improves the system performance. A Rician fading channel is considered in [147] through an application of coded compressed sensing. A correlated fading scenario is tackled in [148]. Frequency-selective fading is examined in the context of MIMO URA in [149] and [150]. The authors in [149] propose a tensor-based modulation scheme, while a slotted scheme employing OFDM is studied in [150].

The asynchronous setup is also receiving increasing attention due to its easier implementation in hardware. See [151]–[153] for proposed solutions in the MIMO URA setting. A slotted concatenated coding scheme is proposed in [151] based on SPARCs and tree codes in an OFDM system, where a variation of the ML technique is employed for decoding of the SPARC. The authors in [152] tackle both timing and frequency offsets and propose an asynchronous URA scheme that employs SPARCs, LDPC codes, and multistage orthogonal pilots for transmission. This scheme also incorporates a graph-based channel reconstruction and collision resolution (GB-CR) algorithm for channel estimation and a modified message passing algorithm for decoding. Symbol asynchronicity in URA is investigated in [153].

Furthermore, some works have investigated URA in other communication setups such as cell-free MIMO [154]–[159] and erasure [160], [161] channels. For instance, a cell-free setup is considered in [154] where the coverage area is divided into multiple regions, each served by a receiver unit. Users in each region employ a common codebook, distinct from those in other regions, to transmit their messages. The signals picked up at the receiver units are then jointly processed in a central unit. Joint processing is also considered in [156] and [157], where a combination of channel coding and fixed sparse coding is employed in the former, while an uncoupled CCS scheme utilizing coarse location information for trellis stitching is proposed in the latter. On the other hand, *Level 2* cooperation, where only the symbol estimates are passed from the access points to the central processing unit, is considered in [158] and [159]. Both schemes are pilot-based, and [158] employs the random spreading approach in the data part, while ODMA is utilized in [159].

The common assumption in the URA literature is that all users have the same number of bits to transmit or require the same quality of service. However, this may be impractical in situations where the devices have different payload requirements or power budgets, especially considering their massive number. This problem, known as multi-class URA, is addressed in [162], [163] within the coded compressed sensing framework. More specifically, a scenario is considered in which devices are clustered into two classes with different SNR levels and payload requirements. In the cluster with higher power, devices transmit using a two-layer superposition modulation, while those in the other cluster transmit with the same base constellation as in the first cluster. To recover packets at the receiver, a CCS-based approach is employed, where an outer code is used to stitch message fragments together across time and layers.

We provide a summary of different URA setups and the corresponding solution approaches in Fig. 20.

VI. Conclusions

In this survey, we provided a comprehensive review of the URA literature, with an emphasis on the state-of-the-art. We grouped different URA solutions according to their assumptions on the underlying wireless channels, and explained the representative works in each group in more detail. Moreover, we compared the works in the literature both in terms of setup and performance to gain a clearer understanding of their pros and cons. Across the three major channel models, our survey

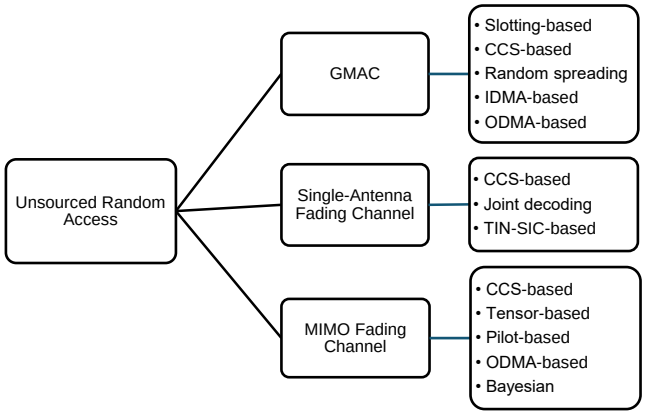


Fig. 20: A summary of different solutions in the URA literature.

demonstrates a consistent evolution of URA research: from foundational achievability bounds on the GMAC, to practical yet computationally constrained designs in single-antenna fading channels, and finally to scalable, high-performing solutions enabled by MIMO. This progression reflects the natural trajectory of URA from theory toward real-world adoption in 6G and IoT systems. The comparative analyses highlight that future URA solutions must jointly optimize energy efficiency, interference management, and computational complexity, while also harnessing spatial diversity and structural coding. Taken together, these insights position URA not only as a theoretical construct but as a practical enabler of massive connectivity, capable of meeting the stringent requirements of next-generation networks.

A. Lessons Learned

The key lessons and insights learned from existing URA solutions can be summarized as follows.

- **ODMA is a promising URA solution.** State-of-the-art URA solutions provide a balance between energy efficiency, scalability, and low-complexity decoding. Among these solutions, ODMA-based schemes excel by offering an effective trade-off between performance, sparsity, and computational complexity. The time-domain sparsity of ODMA-based solutions significantly reduces user interference, making multiuser detection and decoding more manageable even under high user loads, resulting in outstanding performance in both AWGN and fading scenarios, particularly when combined with power diversity. The structured sparsity introduced by ODMA is especially advantageous in massive MIMO systems, where spatial diversity can further enhance separation and decoding accuracy. Additionally, ODMA-based solutions are promising for asynchronous URA, as the ODMA transmission patterns can be utilized instead of commonly employed concatenated preambles to estimate the start times of packets necessary for asynchronous detection, thereby increasing system efficiency. Overall, ODMA-based schemes are promising candidates for scalable, practical, and high-performance solutions for anticipated URA applications, such as massive IoT.
- **Slotted transmission is a simple yet effective interference management technique in URA.** Interference management is one of the major challenges in URA. Among various interference management approaches in the URA literature, ODMA and slotted transmission are promising yet straightforward techniques that introduce structured sparsity by limiting the transmission of user packets to a small portion of the frame, thereby reducing interference. More specifically, by dividing the transmission frame into multiple slots and allowing each user to randomly select one or more slots for transmission, slotting-based solutions decrease the number of simultaneously active users per slot. This localized activity reduces multiuser interference, thereby enabling more accurate signal recovery and effective use of decoding techniques such as SIC. Furthermore, slotted transmission simplifies the decoding process by confining computations to smaller subsets of users, thus decreasing the overall computational complexity compared to full-frame schemes. As a result, slotting serves as a key design element for practical and low-complexity URA implementations.
- **Pilot design and proper channel estimation are critical in URA scenarios over fading channels.** In fading environments, reliable channel estimation is essential for accurate user separation and decoding. Since assigning dedicated pilots to each user is infeasible in URA due to scalability constraints, pilot-based solutions often require users to draw pilot sequences from a common, non-orthogonal codebook. This introduces the risk of pilot collisions, which can lead to severe interference and decoding failures. Therefore, minimizing the probability of pilot collisions and optimizing pilot properties, such as length, power, and structure, are crucial for effective activity detection, channel estimation, and interference management. Note that existing pilot-free URA schemes eliminate pilot overhead, but often result in higher complexity and less reliable performance compared to their pilot-based counterparts.
- **Massive MIMO is a game changer in the realization of URA schemes.** Provided with a large number of antennas, the URA receiver can better manage interference and enhance channel estimation accuracy, even under heavy user loads. In particular, massive MIMO enables advanced decoding strategies, such as spatial separation of users and improved compressed sensing techniques, that would be infeasible for a receiver with a limited number of antennas. These capabilities position massive MIMO as a foundational enabler for highly efficient, scalable communication in a URA setup.
- **URA schemes are sensitive to various system parameters.** Certain system parameters can significantly affect the performance, computational complexity, and scalability in URA solutions. Therefore, key parameters, such as the number of active users, frame length, number of slots, pilot sequence length, and power allocation, must be carefully optimized to balance trade-offs between interference mitigation, decoding accuracy, and computational efficiency. Improper tuning of these parameters

can lead to severe performance degradation; for instance, too many pilot sequences increase the computational complexity of activity detection, while insufficient pilot diversity raises the risk of collisions. The interdependence of these factors makes robust URA design a delicate task, requiring adaptive parameter configuration based on traffic patterns and deployment scenarios to ensure reliable communication.

B. Future Research Directions

As evident from this survey, numerous works on URA have been conducted in recent years. However, there are still issues, some of which are highly critical, that require further investigation. We will next outline some of these issues and potential methods to address them, and provide several promising future research directions.

1) ***Asynchronous URA*** *:* In URA, the common assumption is that the users transmit simultaneously, or more precisely, their signals are received at the same time at the receiver (e.g., symbol synchronicity). However, due to the nature of users and devices in envisioned URA applications, such as massive sensor networks, this transmission coordination is very challenging. This makes asynchronous URA an important and practically relevant research direction to investigate. Note that even if the transmission synchronization is realized, the synchronous reception cannot be guaranteed. Due to the massive number of potential devices with diverse distances to the receiver, the difference in reception times between many users can be larger than a time slot, which makes the setup asynchronous. Interestingly, it has been recently shown that if asynchronicity is dealt with properly, it can even lead to performance improvements compared to the synchronous case [164].

There have been some recent attempts in the literature to address asynchronous URA [51], [99], [100], [103], [151], [152]. Assuming a bounded time horizon enables these works to convert the problem to one similar to the synchronous scenario, e.g., by employing a limited decoding time frame. For instance, in [100], the authors employ a frequency-domain solution to convert the asynchronous problem to a phase-shift estimation one. Then, conventional synchronous strategies such as TIN-SIC are utilized to decode the active users. On the other hand, [81] and [87] consider a fully asynchronous (i.e., with an unbounded time horizon) setup for URA over GMAC, and a window-based decoding approach.

The primary URA-specific challenge in a fully asynchronous setup for URA over fading channels is the increased computational complexity, which is significantly higher than that of a conventional (non-URA) setup due to the higher number of required preambles. More specifically, in an asynchronous GMAC setup, the preambles only serve to identify the start of the packets at the receiver, so even a very small preamble pool would suffice; while in an asynchronous fading setup, they are also needed to estimate the channel coefficients, which are then used for user separation. In a conventional fading setup, the receiver needs to search only over the assigned preambles, whose number equals the number of active users. However, in a URA setup, due to the lack of coordination between the users and the receiver, the number of available preambles must be significantly higher than in the conventional setup; hence, a much higher computational complexity is required for user activity detection and channel estimation. This computational complexity substantially increases in the asynchronous setup with a sliding window receiver, as the receiver shifts and correlates all possible preambles through the window to find the start of the packets as well as the user channel estimates, repeating this process for each shift of the sliding window. This significantly increases the overall computational complexity, rendering the solution practically infeasible. One way to reduce the complexity is to increase the moving steps of the sliding window, at the cost of an overall performance loss. More sophisticated schemes are needed to alleviate this issue.

2) ***Robust URA****:* Like any other communication system, the performance of a URA system is affected by the presence of imperfections in the assumed models or employed communication blocks. For instance, the presence of hardware impairments, e.g., due to non-linearities of the power amplifiers or phase errors in phase-shifters, is highly probable in many practical URA implementations due to cheap and low-precision devices. So far, this challenge is addressed only in [139], where a robust solution is presented for the case of residual hardware impairments in which the remaining hardware impairments after compensation are statistically modeled as additive and multiplicative noises. The analysis of different hardware impairments on the performance of existing URA solutions and providing robust solutions to mitigate them is an interesting research direction.

Examining the URA literature, we can observe that URA schemes are more sensitive to the choice of design parameters than conventional ones. That is, for the URA schemes to achieve acceptable performance, their design parameters must be optimized whenever a system hyperparameter, such as the active user load or the target PUPE, changes. This makes their standardization and practical employment challenging. Therefore, modifications are needed to make the URA solutions robust against changes in system hyperparameters. One approach is to employ stochastic optimization techniques, such as optimizing the expected value of the cost metric (e.g., PUPE) over the statistics of the intended system hyperparameters, to obtain a set of design parameters that perform well across a wide range of system hyperparameters.

3) ***Machine Learning (ML) for and with URA*** *:* Current solutions look at the URA problem in a single shot. However, the transmission is recurrent and spans several frames in many practical applications. In such cases, ML techniques, such as reinforcement learning, can be employed to adapt to the environment and learn suitable transmission strategies. ML techniques can also be used to simplify certain blocks in the decoding process, such as user activity detection (similarly to the deep learning-based solutions in [165]) and channel estimation. For instance, in [166], the authors demonstrate that employing a dictionary learning-based solution eliminates the need for pilots for channel estimation in URA scenarios over fading channels while maintaining an acceptable performance.

Apart from general challenges in implementing ML solutions in communication systems, such as energy and computational constraints and compatibility with legacy systems, there are some that are specific or more pronounced in a URA setup, which is mostly due to the massive number of potential users (up to the order of millions) and their sporadic activity. Training a neural network with an input of very large size (due to the massive number of potential users) with non-stationary characteristics (due to the sporadic activity) is a very challenging task.

URA can also be employed to enable distributed learning among wireless agents. In federated learning (FL), the goal of the parameter server, which acts as the orchestrator, is to aggregate local model updates from many clients at each iteration, and update the global model by averaging these models [167]. When clients are co-located, and the parameter server is a wireless access point, model updates have to be communicated over a shared wireless channel. In conventional methods, this requires sharing the available channel resources among the active clients; however, in FL, the parameter server is not interested in individual updates. Therefore, in [168], over-the-air computing (OAC) was proposed to exploit the signal superposition property of wireless signals. This would allow the parameter server to directly receive the summation of the model updates.

However, such an analog over-the-air computing scheme would require a certain level of coordination and synchronization among the clients. An alternative digital OAC scheme is proposed in [169] using URA. In OAC with URA, each client quantizes its model update, and sends the corresponding codeword. Both the quantization and channel codebooks are shared among all the clients. Differently from conventional URA, which relies on the fact that active users are likely to transmit independent codewords, and the receiver tries to identify only the transmitted codewords by the active clients, here the users are more likely to transmit the same codeword as their updates tend to converge as the training process progresses. Moreover, the receiver wants to identify not only the active codewords, but also how many clients transmit each codeword. This approach can be extended to other functions, and the identification of the optimal quantization and communication codebooks is an open research direction.

*4) **Semantic and Task-Oriented URA**:* Conventional URA frameworks are designed under the assumption that all transmitted bits are equally important, focusing primarily on reliable message recovery. However, in many IoT and machine-type communication scenarios, the goal is not to recover raw data streams but rather to support specific tasks such as anomaly detection, control signaling, or status monitoring. In such cases, transmitting semantic or task-oriented information, i.e., only the most relevant features of the data, can drastically reduce the required payload and improve spectral efficiency. Future research may investigate how semantic encoding techniques can be combined with URA to enable resource-efficient access, for example, by allowing multiple devices reporting correlated measurements to share compressed semantic representations. Such integration would align URA more closely with emerging 6G paradigms, where communication systems are expected to deliver knowledge and actionable insights rather than just data.

*5) **ISAC in URA**:* Existing research on URA has focused primarily on the transmission of information from communication users. However, with the evolution of 6G and beyond, new functionalities such as sensing, computing, and security are becoming critical alongside communication. As the hardware architectures and signal processing methods for these tasks are mostly similar, jointly adopting them significantly reduces hardware costs compared to separate systems while maintaining comparable performance [170]. A promising example of such joint designs is integrated sensing and communication (ISAC) systems.

The problem of ISAC in URA is introduced for the first time in [171], where a massive number of unsourced and sporadically active users are connected to the system to jointly perform sensing and communication tasks. In the resulting system, called unsourced ISAC (UNISAC), users are divided into two main categories: Sensing and communication users. The system aims to decode signals from communication users while detecting signals from active sensing users and estimating their AOAs. The authors derive an achievability bound for UNISAC to demonstrate its effectiveness in enabling both user sensing and user communication within a URA setup. This work is conducted under two key assumptions: the targets are active, and the channels between the users (targets) and the BS are line-of-sight (LoS).

The problem of UNISAC has been further investigated in subsequent works, including: proposing a practical transceiver scheme for UNISAC under LoS channels between the BS and all sensing and communication users [173], [174], as well as under LoS target–BS channels and Rayleigh fading user–BS channels [175]; studying the UNISAC model equipped with fluid antenna systems [176], [177]; and addressing the problem of passive target sensing [180].

There are several promising directions for future research. In existing UNISAC setups, only the problem of user angle estimation has been investigated, whereas the field of sensing is much broader. For example, tracking, surveillance, passive source localization, and target localization are important sensing tasks that warrant further study. Moreover, networks incorporating advanced technologies, such as multi-BS systems, UAV-equipped systems, and RIS-aided systems, can be explored to enhance performance. Another significant research direction is to study UNISAC under realistic channel models, including frequency-selective fading, large-scale fading, Doppler effects, asynchronous transmission errors, and hardware impairments [172].

To explore the challenges that future work on URA-based ISAC may encounter, many aspects still require further investigation. In a URA system, a massive number of users transmit sporadically without being registered or authenticated by the network. Therefore, there is no control over transmitter characteristics such as location, signal structure, and beam direction, which are critical when the transmitters' echoes are used for reliable passive target sensing. Additionally, the random access nature of user transmissions causes heavy interference, a phenomenon uncommon in traditional sensing

techniques. Advanced algorithms are necessary to address these challenges and enhance the system's ability to effectively sense passive targets.

*6) **URA with Feedback**:* In communication systems, feedback has been shown to greatly simplify transmission and reception mechanisms and, in some cases, even improve system performance [178]. The main challenge in employing feedback-based techniques in URA is the lack of coordination between the users and the receiver due to the massive number of potential users. However, even in URA, there can be some level of coordination as long as the signal sent from the receiver is not user-specific, as there is no user identity in URA. For instance, the receiver can broadcast a common pilot or beacon to the users, allowing them to check if they are among the successfully decoded users from the previous round and estimate their channel coefficients or at least their large-scale components, depending on their dedicated time slots. Note that the time dedicated to this operation needs to be considered as extra overhead when comparing with solutions without feedback. This issue is tackled in part in [179]–[182]. In [179], hashes of the messages of successfully decoded users and their estimated channel coefficients are fed back using beamforming techniques. In [180], the authors enhance the approach of [179] by informing both the successfully decoded users and those that were not decoded about their respective transmission outcomes. In [181], with the help of a designed feedback beacon, the users estimate their channels and infer whether they were successfully decoded using a predefined threshold. An application of URA can be in sensor networks where the devices partially observe a common state, which is the main interest of the receiver. For such an application, the authors in [182] devise a mechanism where the receiver broadcasts a reliability indicator of the estimated state after each transmission, using which each device can decide whether to transmit in the next iteration.

Recall that the main novelty of URA is providing a unified framework to decouple user identification and data transmission problems. This makes it a suitable approach in applications where user identification is not feasible due to the massive number of potential users. However, URA remains applicable and is a good candidate in non-massive scenarios with a considerable number of potential users, where user identification is still feasible. In these scenarios, conventional authentication and acknowledgment schemes impose a huge overhead on the system. This issue is tackled in [183].

*7) **RIS-Aided URA**:* In URA, due to the massive number of potential users, it is very probable that the line-of-sight link between some users and the receiver is either not present or heavily attenuated. In this case, employing relaying techniques such as RIS can play a crucial role in enhancing the overall system performance by providing an alternative communication link with potentially many degrees of freedom. However, the huge channel estimation overhead is the main challenge in the use of RIS in URA systems. The channel estimation overhead is already one of the main issues in RIS systems due to the high number of channel coefficients per user as a result of the large number of RIS elements and the potentially massive number of antennas at the receiver (as the benefits of a RIS system are amplified when combined with a massive MIMO receiver). This problem is more pronounced in a RIS-aided URA system as there is a potentially massive number of even users, which renders most existing channel estimation schemes in RIS infeasible. There are a few works that have considered the RIS-aided URA scenario [136]–[138]. Although these works make channel estimation feasible in a RIS-aided URA setup employing an iterative SIC-based approach and time frame slotting strategies, the high computational complexity remains a significant challenge.

Moreover, there are many promising RIS configurations and types that need to be investigated in a URA setup, such as active or hybrid RIS [184] and metasurface-based RIS [185], in order to find suitable candidates. Moreover, the employment of massive arrays such as extra-large massive MIMO and RIS leads to some active users being in the near-field of these arrays, which makes hybrid near-field-far-field analysis necessary. This is in contrast with the existing URA solutions, which are all for the far-field regime.

*8) **Security**:* One of the key concerns in multiple access systems is a robust implementation against security threats, such as jamming and eavesdropping. To address these security threats, some common countermeasures are employed, including *mutual authentication* protocols for access control and authentication, such as the authentication and key agreement (AKA) used in 4G/5G, and data encryption techniques, such as AES-128 in 4G and AEGIS-256 in 5G. These protocols ensure that both users and the receiver (network) verify each other's identities before establishing a secure connection. However, many of these solutions are not feasible in a URA setup, primarily due to the lack of coordination between users and the receiver. For example, key agreement in the AKA protocol relies on some level of coordination between the user and the receiver to establish a secure connection.

To the best of our knowledge, there has not been any work addressing the security aspects of URA. However, some existing security solutions can still be utilized in URA with certain modifications or considerations. For example, light encryption methods can be employed (similarly to those in [186] and [187]) to introduce a certain level of security with a common key among users. In some URA solutions, the key can also be randomly selected from a pool, with each key paired with a transmission signature, such as a pilot sequence in pilot-based schemes and a transmission pattern in ODMA-based ones. Likewise, physical layer security techniques, such as the injection of artificial noise, can also be employed, with the modification that all users utilize the same noise sequence or randomly select it from a pool of available sequences. Existing physical layer authentication techniques based on signal features, such as channel state information or RF fingerprints [188], can also be utilized as long as they do not rely on user-specific signatures.

*9) **Standardization**:* There has been no effort in standardizing URA so far. This requires a thorough investigation and analysis of the current competing solutions, and finding a good candidate that can preferably support both the current protocols, perhaps with minor modifications, and the high demands of the future ones. For instance, in [189], with a slight

modification in the 5G new radio (5G-NR) release 16 protocol, the authors achieve an improvement in energy efficiency of about 5 to 10 dB compared to the existing 5G-NR solution for systems with a small number of active users (10 to 50 users), employing an IDMA-based URA solution. However, this solution is not scalable to higher active user loads.

There are two main challenges in incorporating URA in current protocols, such as 5G-NR and beyond, for reasonable active user loads (e.g., in the order of hundreds). The first is the small number of preamble sequences available in 5G-NR as opposed to the massive number of preamble sequences (pilots) needed in the preamble/pilot-based URA solutions. The second is the limited number of access patterns that can be realized, which is vital in transmission pattern-based URA solutions such as ODMA-based ones. This indicates that a crucial step in realizing URA in future protocols is to considerably increase the available pilot or transmission patterns.

This comprehensive survey of URA reveals an active field that has evolved significantly in recent years, with substantial progress in fading and MIMO channels. The research challenges identified above demonstrate both the complexity of remaining questions and the rich opportunities for future investigation. Addressing these challenges will be essential for the practical adoption and deployment of URA solutions in future wireless systems.

REFERENCES

[1] Ericsson Mobility Report: Data and forecasts - IoT connections outlook, 2024. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/iot-connections-outlook

[2] H. Shariatmadari, S. Iraji, and R. Jantti, *From Machine-to-Machine Communications to Internet of Things: Enabling Communication Technologies*. Chapman and Hall/CRC, 2017.

[3] E. Dutkiewicz, X. Costa-Perez, I. Z. Kovacs, and M. Mueck, "Massive machine-type communications," *IEEE Network*, vol. 31, no. 6, pp. 6–7, Dec. 2017.

[4] E. Bigleiri and L. Gyorfi, *Multiple Access Channels: Theory and Practice*. Fairfax, VA: IOS, 2007, published in cooperation with NATO Public Diplomacy Division.

[5] M. Ozates, "Design of low complexity unsourced random access schemes over wireless channels," Ph.D dissertation, EEE Dept., Bilkent Univ., Ankara, 2023. [Online]. Available: https://repository.bilkent.edu.tr/items/dba80064-3b92-4d7b-8fb7-7bfa9f3bc263

[6] M. J. Ahmadi, "Novel unsourced random access algorithms over Gaussian and fading channels," Ph.D dissertation, EEE Dept., Bilkent Univ., Ankara, 2023. [Online]. Available: https://repository.bilkent.edu.tr/items/2b7ecfff-8c92-466e-805f-f7ee89524bbb

[7] N. Abramson, "The ALOHA system-another alternative for computer communications," in *Proc. Fall Joint Comput. Conf.*, Nov. 1970, pp. 281–285.

[8] 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.

[9] M. Mohammadkarimi, M. A. Raza, and O. A. Dobre, "Signature-based nonorthogonal massive multiple access for future wireless networks: Uplink massive connectivity for machine-type communications," *IEEE Veh. Technol. Mag.*, vol. 13, no. 4, pp. 40–50, Dec. 2018.

[10] N. H. Mahmood, R. Abreu, R. Böhnke, M. Schubert, G. Berardinelli, and T. H. Jacobsen, "Uplink grant-free access solutions for URLLC services in 5G new radio," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2019, pp. 607-612.

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

[12] M. Gkagkos, K. R. Narayanan, J. F. Chamberland and C. N. Georghiades, "FASURA: A scheme for quasi-static fading unsourced random access channels," *IEEE Trans. Commun.*, vol. 71, no. 11, pp. 6391-6401, Nov. 2023.

[13] M. J. Ahmadi, M. Kazemi, and T. M. Duman, "Unsourced random access with a massive MIMO receiver using multiple stages of orthogonal pilots: MIMO and single-antenna structures," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1343–1355, Feb. 2024.

[14] 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.*, vol. 23, no. 4, pp. 3437-3449, Apr. 2024.

[15] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy and G. Caire, "Massive access for future wireless communication systems," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 148-156, Aug. 2020.

[16] Y. Li et al., "Unsourced multiple access for 6G massive machine type communications," *China Commun.*, vol. 19, no. 3, pp. 70-87, Mar. 2022.

[17] Z. Gao et al., "Compressive-sensing-based grant-free massive access for 6G massive communication," *IEEE Internet Things J.*, vol. 11, no. 5, pp. 7411-7435, Mar. 2024.

[18] G. Liva and Y. Polyanskiy, "Unsourced multiple access: A coding paradigm for massive random access," *Proc. IEEE*, vol. 112, no. 9, pp. 1214-1229, Sept. 2024.

[19] A. F. M. S. Shah, A. N. Qasim, M. A. Karabulut, H. Ilhan and M. B. Islam, "Survey and performance evaluation of multiple access schemes for next-generation wireless communication systems," *IEEE Access*, vol. 9, pp. 113428-113442, 2021.

[20] S. M. R. Islam, N. Avazov, O. A. Dobre and K. -s. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G Systems: Potentials and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721-742, second quarter 2017.

[21] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2294-2323, third quarter 2018.

[22] R. Hoshyar, F. P. Wathan and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616-1626, Apr. 2008.

[23] R. Hoshyar, R. Razavi and M. Al-Imari, "LDS-OFDM an efficient multiple access technique," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Taipei, Taiwan, 2010, pp. 1-5.

[24] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE 24th Annual Int. Symp. Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, London, UK, 2013, pp. 332-336.

[25] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang and J. Xu, "multiuser shared access for Internet of Things," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Nanjing, China, 2016, pp. 1-5.

[26] L. G. Roberts, "ALOHA packet system with and without slots and capture," *SIGCOMM Comput. Commun. Rev.*, vol. 5, pp. 28–42, Apr. 1975.

[27] E. Casini, R. De Gaudenzi and O. Del Rio Herrero, "Contention resolution diversity slotted ALOHA (CRDSA): An enhanced random access scheme for satellite access packet networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1408-1419, Apr. 2007.

[28] 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.

[29] E. Paolini, G. Liva and M. Chiani, "Coded slotted ALOHA: A graph-based method for uncoordinated multiple access," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6815-6832, Dec. 2015.

[30] G. L. Choudhury and S. S. Rappaport, "Diversity ALOHA- A random access scheme for satellite communications," *IEEE Trans. Commun.*, vol. 31, pp. 450–457, Mar. 1983.

[31] U. Demirhan and T. M. Duman, "Irregular repetition slotted ALOHA with energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 45054517, Sep. 2019.

[32] T. Akyildiz, U. Demirhan, and T. M. Duman, "Energy harvesting irregular repetition ALOHA with replica concatenation," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 955-968, Feb. 2021.

[33] J. Haghighat and T. M. Duman, "Analysis of coded slotted ALOHA with energy harvesting nodes for perfect and imperfect packet recovery scenarios," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 7424-7437, Nov. 2023

[34] J. Haghighat and T. M. Duman, "An energy-efficient feedback-aided irregular repetition slotted ALOHA scheme and its asymptotic performance analysis," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9808-9820, Dec. 2023.

[35] M. Kazemi, T. M. Duman, and M. Médard, "Collision resolution for random access," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3464–3477, May 2022.

[36] C. Pyo et al., "A throughput study of grant-free multiple access for massive wireless communications," in *Proc. Int. Symp. Wireless Personal Multimedia Commun.*, 2017, pp. 529–534.

[37] 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, third quarter 2020.

[38] R. Razavi, M. Dianati, et al., "Non-orthogonal multiple access (NOMA) for future radio access," *5G Mobile Commun.*, 2016, pp. 135–163.

[39] K. Au et al., "Uplink contention based SCMA for 5G radio access," in *Proc. IEEE Global Commun. Workshops (GC Wkshps)*, Austin, TX, USA, 2014, pp. 900-905.

[40] F. Monsees, C. Bockelmann and A. Dekorsy, "Reliable activity detection for massive machine to machine communication via multiple measurement vector compressed sensing," in *Proc. IEEE Global Commun. Workshops (GC Wkshps)*, Austin, TX, USA, 2014, pp. 1057-1062.

[41] A. T. Abebe and C. G. Kang, "Compressive sensing-based random access with multiple-sequence spreading for MTC," in *Proc. IEEE Global Commun. Workshops (GC Wkshps)*, San Diego, CA, USA, 2015, pp. 1-6.

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

[43] L. Liu and W. Yu, "Massive connectivity with massive MIMO - Part I: Device activity detection and channel estimation," *IEEE Trans. Signal. Process.*, vol. 66, no. 11, June 2018.

[44] M. J. Ahmadi and T. M. Duman, "Random spreading for unsourced MAC with power diversity," in *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3995-3999, Dec. 2021.

[45] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, no. 3-4, pp. 154-655, Nov. 2017.

[46] Y. Polyanskiy, H. V. Poor and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307-2359, May 2010.

[47] E. MolavianJazi and J. N. Laneman, "A second-order achievable rate region for Gaussian multi-access channels via a central limit theorem for functions," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6719-6733, Dec. 2015.

[48] K. -H. Ngo, A. Lancho, G. Durisi and A. Graell i Amat, "Unsourced multiple access with random user activity," *IEEE Trans. Inf. Theory.*, vol. 69, no. 7, pp. 4537-4558, July 2023.

[49] A. Glebov, P. Rybin, K. Andreev, and A. Frolov, "Energy efficiency of unsourced random access over the binary-input Gaussian channel," *IEEE Commun. Lett.*, vol. 27, no. 9, pp. 2313-2317, Sep. 2023.

[50] K. -H. Ngo, G. Durisi, A. Graell i Amat, P. Popovski, A. E. Kalør and B. Soret, "Unsourced multiple access with common alarm messages: Network slicing for massive and critical IoT," *IEEE Trans. Commun.*, vol. 72, no. 2, pp. 907-923, Feb. 2024.

[51] J.-S. Wu, P.-H. Lin, M. A. Mross and E. A. Jorswieck, "Worst-case per-user error bound for asynchronous unsourced multiple access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Athens, Greece, 2024, pp. 3207-3212.

[52] J.-S. Wu, P.-H. Lin, M. A. Mross and E. A. Jorswieck, "Wrap-decoding in asynchronous unsourced multiple access with and without delay information," arXiv, 2025. Available: https://arxiv.org/abs/2501.12294.

[53] S. Derya, S. -C. Yu, H. -W. Young, E. A. Jorswieck, P. -H. Lin and S. -C. Lin, "Joint delay and user activity detection in asynchronous massive access," in *Proc. Wireless and Optical Commun. Conf. (WOCC)*, Hsinchu, Taiwan, 2024, pp. 175-179.

[54] F. Ritter, J. Mandelbaum, A. Fengler, H. Jäkel and L. Schmalen, "Removal of small weight stopping sets for asynchronous unsourced multiple access," arXiv, 2025. Available: https://arxiv.org/abs/2501.12186.

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

[56] Z. Zhang, K. Niu and H. Cui, "Improved BCH-polar concatenated scheme for unsourced random Access in Internet of Things," *IEEE Internet Things J.*, vol. 11, no. 19, pp. 32172-32182, Oct. 2024.

[57] 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.

[58] F. Clazzer, F. Amri and M. Grec, "An investigation of the compressed sensing phase in unsourced multiple access," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Dubai, United Arab Emirates, 2024.

[59] E. Marshakov, G. Balitskiy, K. Andreev, and A. Frolov, "A polar code based unsourced random access for the Gaussian mac," in *Proc. IEEE VTC*, pp. 1-5, Honolulu, USA, 2019.

[60] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051-3073, July 2009.

[61] A. K. Tanc and T. M. Duman, "Massive random access with trellis based codes and random signatures," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1496–1499, May 2021.

[62] X. Sui, Z. Si and J. Dai, "Finite-field coding for unsourced multiple access," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, 2023, pp. 691-696.

[63] V. K. Amalladinne, J. 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.

[64] A. Fengler, P. Jung and G. Caire, "SPARCs for unsourced random access," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6894-6915, Oct. 2021.

[65] 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.

[66] H. Cao, J. Xing and S. Liang, "CRC-aided sparse regression codes for unsourced random access," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 1944-1948, Aug. 2023.

[67] P. Agostini, Z. Utkovski, A. Decurninge, M. Guillaud and S. Stańczak, "Constant weight codes with Gabor dictionaries and Bayesian decoding for massive random access," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 2917-2931, May 2023.

[68] D. Jiang and P. Fan, "A raptor code based unsourced random access with coordinated tree-Raptor decoding algorithm," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC Workshops)*, Dalian, China, 2023, pp. 1-6.

[69] E. Nassaji and D. Truhachev, "Dynamic compressed sensing approach for unsourced random access," *IEEE Commun. Lett.*, vol. 28, no. 7, pp. 1644-1648, July 2024.

[70] A. K. Pradhan, V. K. Amalladinne, K. R. Narayanan, and J.-F. Chamberland, "Polar coding and random spreading for unsourced multiple access," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, June 2020, pp. 1-6.

[71] Z. Sun, Y. Xiao, D. Lin and X. Xu, "Performance evaluation of unsourced multiple access with polarization-adjusted convolutional coding," in *Proc. IEEE Veh. Technol. Conf. (VTC2022-Spring)*, Helsinki, Finland, 2022, pp. 1-5.

[72] Z. Han, X. Yuan, C. Xu, S. Jiang and X. Wang, "Sparse Kronecker-product coding for unsourced multiple access," *IEEE Wireless Commun. Lett.*, vol. 10, no. 10, pp. 2274-2278, Oct. 2021.

[73] A. K. Pradhan, V. K. Amalladinne, K. R. Narayanan and J. -F. Chamberland, "LDPC codes with soft interference cancellation for uncoordinated unsourced multiple access," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Montreal, Canada, June 2021, pp. 1-6.

[74] O. del Rio Herrero and R. De Gaudenzi, "A high efficiency scheme for quasi-real-time satellite mobile messaging systems," in *Proc. Int. Workshop Signal Process. Space Commun.*, Rhodes, Greece, 2008, pp. 1-9.

[75] R. Schiavone, G. Liva and R. Garello, "Design and performance of enhanced spread spectrum aloha for unsourced multiple access," *IEEE Commun. Lett.*, vol. 28, no. 8, pp. 1790-1794, Aug. 2024.

[76] J. Dang, Z. Zhang and Y. Wu, "Random spreading with higher order modulation in unsourced random access," *IEEE Commun. Lett.*, vol. 29, no. 4, pp. 829-833, April 2025.

[77] A. K. Pradhan, V. K. Amalladinne, A. Vem, K. R. Narayanan, and J. Chamberland, "Sparse IDMA: A joint graph-based coding scheme for unsourced random access," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7124–7133, Nov. 2022.

[78] Y. Yuan, Y. Huang, C. Yan, S. Wang, S. Ma and X. Shen, "Unsourced sparse multiple access: A promising transmission scheme for 6G massive communication," *IEEE Commun. Mag.*, early access, Mar. 04, 2025.

[79] M. Zheng, Y. Wu, and W. Zhang, "Polar coding and sparse spreading for massive unsourced random access," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Victoria, Canada, Feb. 2020, pp. 1-5.

[80] D. Truhachev, M. Bashir, A. Karami and E. Nassaji, "Low-complexity coding and spreading for unsourced random access," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 774-778, Mar. 2021.

[81] A. Karami, I. Pazouki, R. Soltani and D. Truhachev, "An approach to asynchronous unsourced random access (URA)," *IEEE Wireless Commun. Netw. Conf. (WCNC)*, Dubai, United Arab Emirates, 2024.

[82] 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, July 2023.

[83] G. Song, K. Cai, Y. Chi, J. Guo and J. Cheng, "Super-sparse on-off division multiple access: Replacing repetition with idling," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2251-2263, Apr. 2020.

[84] M. Ozates, M. Kazemi and T. M. Duman, "Unsourced random access using ODMA and polar codes," *IEEE Wireless Commun. Lett.*, vol. 13, no. 4, pp. 1044-1047, Apr. 2024.

[85] J. Yan, Y. Li, G. Song and Z. Zhang, "Enhanced ODMA with pattern collision resolution and parameter design for unsourced multiple access," *IEEE Trans. Commun.*, early access, May 19, 2025.

[86] Z. Zhang, M. J. Ahmadi, J. Dang, K.-K. Wong, Z. Zhang, and C. Masouros, "Sparse code transceiver design for unsourced random access with analytical power division in Gaussian MAC," arXiv, 2025. Available: https://arxiv.org/abs/2505.01988.

[87] M. Ozates, M. Kazemi, G. Liva and D. Gündüz, "A fully asynchronous unsourced random access scheme," in *Proc. IEEE 26th Int. Workshop Signal Process. Artif. Intell. Wireless Commun. (SPAWC)*, Surrey, United Kingdom, 2025, pp. 1-5.

[88] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient coded random access for the wireless uplink," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4694-4708, Aug. 2020.

[89] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static SIMO fading channels at finite blocklength," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, 2013, pp. 1531–1535.

[90] Y. Polyanskiy, *Channel Coding: Non-Asymptotic Fundamental Limits*. Princeton University, 2010.

[91] S. S. Kowshik and Y. Polyanskiy, "Quasi-static fading multiple access with many users and finite payload," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, July 2019, pp. 440-444.

[92] S. S. Kowshik and Y. Polyanskiy, "Fundamental limits of many-user multiple access with finite payloads and fading," *IEEE Trans. Inf. Theory*, vol. 67, no. 9, pp. 5853-5884, Sep. 2021.

[93] K. Andreev, E. Marshakov and A. Frolov, "A polar code based TIN-SIC scheme for the unsourced random access in the quasi-static fading multiple access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Los Angeles, USA, June 2020, pp. 3019-3024.

[94] J. Gao, Y. Wu, T. Li and W. Zhang, "Energy efficiency of MIMO massive unsourced random access with finite blocklength," *IEEE Wireless Commun. Lett.*, vol. 12, no. 4, pp. 743-747, Apr. 2023.

[95] J. Gao, Y. Wu, G. Caire, W. Yang and W. Zhang, "Unsourced random access in MIMO quasi-static Rayleigh fading channels with finite blocklength," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Athens, Greece, 2024, pp. 3213-3218.

[96] K. Andreev, D. Ustinova, and A. Frolov, "Unsourced random access with the MIMO receiver: Projection decoding analysis," *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, pp. 69-73, Jan. 2024.

[97] J. Gao, Y. Wu, G. Caire, W. Yang, H. W. Poor and W. Zhang, "Unsourced random access in MIMO quasi-static Rayleigh fading channels: Finite blocklength and scaling law analyses," *IEEE Trans. Inf. Theory*, vol. 71, no. 6, pp. 4342-4373, June 2025.

[98] K. Andreev, P. Rybin and A. Frolov, "Coded compressed sensing with list recoverable codes for the unsourced random access," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 7886-7898, Dec. 2022.

[99] K. Andreev, S. S. Kowshik, A. Frolov and Y. Polyanskiy, "Low complexity energy efficient random access scheme for the asynchronous fading multiple access," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Honolulu, USA, Sep. 2019, pp. 1-5.

[100] S. S. Kowshik, K. Andreev, A. Frolov and Y. Polyanskiy, "Short-packet low-power coded access for massive multiple access," in *Proc. 53rd Asilomar Conf. Signals, Systems, and Computers*, CA, USA, Sep. 2019, pp. 827-832.

[101] V. K. Amalladinne, K. R. Narayanan, J. -F. Chamberland and D. Guo, "Asynchronous neighbor discovery using coupled compressive sensing," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Brighton, UK, May. 2019, pp. 4569-4573.

[102] E. Nassaji, M. Bashir and D. Truhachev, "Unsourced random access over fading channels via data repetition, permutation, and scrambling," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1029-1042, Feb. 2022.

[103] M. Ozates and T. M. Duman, "Unsourced random access over frequency-selective channels," *IEEE Commun. Lett.*, vol. 27, no. 4, pp. 1230-1234, Apr. 2023.

[104] J. Yun and W. Choi, "Erasure correcting blind detection in unsourced random access for grant-free massive connections," *IEEE Trans. Wireless Commun.*, vol. 23, no. 3, pp. 2428-2439, Mar. 2024.

[105] 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.

[106] 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.

[107] F. Tian, X. Chen, Y. L. Guan and C. Yuen, "Design and analysis of massive uncoupled unsourced random access with Bayesian joint decoding," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 10350-10364, July 2024.

[108] F. Tian, X. Chen, C. Huang and Z. Zhang, "Exploiting matrix information geometry for integrated decoding of massive uncoupled unsourced random access," *IEEE Trans. Commun.*, vol. 72, no. 10, pp. 6178-6191, Oct. 2024.

[109] Z. Zhang, J. Dang, Y. Qi, Z. Zhang, L. Wu and H. Wang, "Blind ODMA for unsourced random access in MIMO and hybrid massive MIMO," *IEEE Internet Things J.*, early access, Sep. 4, 2024.

[110] P. Agostini, Z. Utkovski and S. Stanczak, "BiSPARCs for unsourced random access in massive MIMO," in *Proc. Asilomar Conf. Signals, Systems, and Computers.*, Pacific Grove, CA, USA, 2023, pp. 299-303.

[111] 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.*, vol. 72, no. 6, pp. 3256-3272, June 2024.

[112] 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.

[113] Q. Ruan and J. Zheng, "Unsourced random access for heterogeneous mobility," in *Proc. Int. Conf. Ubiquitous Commun.*, Xi'an, China, 2023, pp. 221-226.

[114] Z. Zhang, J. Dang, Z. Zhang, L. Wu and B. Zhu, "Unsourced random access with uncoupled compressive sensing and forward error correction," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 3555-3560, Feb. 2025.

[115] 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.

[116] A. Decurninge, P. Ferrand and M. Guillaud, "Massive random access with tensor-based modulation in the presence of timing offsets," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Rio de Janeiro, Brazil, 2022, pp. 1061-1066.

[117] Z. Luan, Y. Wu, S. Liang, W. Han, B. Bai and L. Zhang, "Modulation for massive unsourced random access based on tensor block term decomposition," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Rio de Janeiro, Brazil, 2022, pp. 637-643.

[118] J. Fang et al., "Polar-coded tensor-based unsourced random access with soft decoding," *IEEE Trans. Commun.*, early access, Mar. 25, 2025.

[119] M. Ozates, M. Kazemi and T. M. Duman, "An ODMA-based unsourced random access scheme with a multiple antenna receiver," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Cape Town, South Africa, 2024, pp. 1857-1862.

[120] 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.

[121] Y. Su and J. Zheng, "Index modulation multiple access for unourced random access," *IEEE Wireless Commun. Lett.*, vol. 12, no. 5, pp. 794-798, May 2023.

[122] Y. Su, J. Zheng, Z. Liang, and J. Ni, "Massive MIMO enabled unsourced random access through LDPC coding and random spreading," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Nanjing, China, Nov. 2022, pp. 178-183.

[123] Q. Sun, Z. Zhang, J. Dang and Z. Zhang, "Non-Negative Least Squares Exploiting Multiple Measurements in Unsourced Random Access," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Hefei, China, 2024, pp. 536-542,

[124] 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.

[125] M. Hao, G. Song, Y. Li, and J. Yan, "Superimposed pilot ODMA transmission for massive MIMO unsourced random access," *IEEE Commun. Lett.*, vol. 29, no. 2, pp. 333-337, Feb. 2025.

[126] Z. Zhang, J. Dang, Z. Zhang, L. Wu, B. Zhu and Y. Wu, "Unsourced random access via random scattering with turbo probabilistic data association detector and treating collision as interference," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 17899-17914, Dec. 2024.

[127] Z. Zhang, J. Dang and Z. Zhang, "Probabilistic ODMA receiver with low-complexity algorithm for MIMO unsourced random access," *IEEE Trans. Veh. Technol.*, early access, May 15, 2025.

[128] J. -F. Determe, J. Louveaux, L. Jacques and F. Horlin, "On the exact recovery condition of simultaneous orthogonal matching pursuit," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 164-168, Jan. 2016.

[129] Z. Zhang, J. Dang, Y. Qi, Z. Zhang, L. Wu and H. Wang, "Efficient ODMA for unsourced random access in MIMO and hybrid massive MIMO," *IEEE Internet Things J.*, vol. 11, no. 23, pp. 38846-38860, Dec. 2024.

[130] Z. Zhang, J. Dang, Z. Zhang, L. Wu and B. Zhu, "Unsourced random access via random dictionary learning with pilot-free transceiver design," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 17884-17898, Dec. 2024.

[131] Z. Zhang, J. Dang and Z. Zhang, "LIM-URA: A less-is-more strategy for quasi-static fading unsourced random access," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Milan, Italy, 2025, pp. 1-6.

[132] Z. Han, X. Yuan, C. Xu and X. Wang, "Receiver design for MIMO unsourced random access with sparse Kronecker-product coding," *IEEE Wireless Commun. Lett.*, vol. 12, no. 1, pp. 45-49, Jan. 2023.

[133] J. C. Jiang and H. M. Wang, "A fully Bayesian approach for massive MIMO unsourced random access," *IEEE Trans. Commun.*, vol. 71, no. 8, pp. 4620-4635, Aug. 2023.

[134] W. Yan, X. Yuan, Z.-Q. He and X. Kuai, "Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems," *IEEE J. Select. Areas Commun.*, vol. 38, no. 8, pp. 1793–1808, Aug. 2020.

[135] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[136] M. J. Ahmadi, M. Kazemi, and T. M. Duman, "RIS-aided unsourced random access," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 3270–3275.

[137] X. Shao, L. Cheng, X. Chen, C. Huang and D. W. K. Ng, "Reconfigurable intelligent surface-aided 6G massive access: Coupled tensor modeling and sparse bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10145-10161, Dec. 2022.

[138] M. J. Ahmadi, M. Kazemi, and T. M. Duman, "RIS-aided unsourced multiple access (RISUMA): Coding strategy and performance limits," *IEEE Trans. Wireless Commun.*, vol. 24, no. 7, pp. 6225–6239, Jul. 2025.

[139] M. Ozates, M. Kazemi and T. M. Duman, "Unsourced random access with hardware impairments," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, 2023, pp. 6946-6951.

[140] R. Soltani, D. Truhachev, A. Bayesteh and M. Vameghestahbanati, "Location-based unsourced random access," *IEEE Access*, vol. 12, pp. 51166-51175, 2024.

[141] J. -C. Jiang and H.-M. Wang, "Dynamic unsourced random access with massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9460-9473, Aug. 2024.

[142] K. -H. Ngo, D. P. Krishnan, K. Okumus, G. Durisi and E. G. Ström, "Type-based unsourced multiple access," in *Proc. IEEE 25th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Lucca, Italy, 2024, pp. 911-915.

[143] D. P. Krishnan, K. Okumus, K. -H. Ngo and G. Durisi, "An achievability bound for type-based unsourced multiple access," arXiv, 2025. Available: https://arxiv.org/abs/2504.19916.

[144] D. P. Krishnan, K. Okumus, K. -H. Ngo and G. Durisi, "Type-based unsourced multiple access over fading channels with cell-free massive MIMO," arXiv, 2025. Available: https://arxiv.org/abs/2504.19954.

[145] R. Kotaba et al., "How to identify and authenticate users in massive unsourced random access," *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3795-3799, Dec. 2021.

[146] M. Bayanifar, S. Ghasemi, R. -A. Pitaval, B. M. Popović and O. Tirkkonen, "Information carrying slotting principles for unsourced coded slotted random access," in *Proc. IEEE 35th Int. Symp. Personal, Indoor and Mobile Radio Commun. (PIMRC),* Valencia, Spain, 2024, pp. 1-6.

[147] F. Tian, X. Chen, L. Liu and D. W. Kwan Ng, "Design of massive unsourced random access over Rician channels," in *Proc. IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Lucca, Italy, 2021, pp. 341-345.

[148] X. Xie, Y. Wu, J. Gao and W. Zhang, "Massive unsourced random access for massive MIMO correlated channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Taipei, Taiwan, 2020, pp. 1-6.

[149] S. Kharbech and A. Decurninge, "Massive unsourced random access with tensor-based modulation and multi-antenna receiver over frequency-selective channels," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Lyon, France, 2024, pp. 1077-1081.

[150] Z. Liang, K. Fukawa and Y. Chang, "Slotted unsourced random access using massive MIMO over frequency selective channels," *IEICE Trans. Commun.*, early access, Dec. 16, 2024.

[151] W. Wang, J. You, S. Liang, W. Han and B. Bai, "Slotted concatenated coding scheme for asynchronous uplink unsourced random access with a massive MIMO receiver," in *Proc. IEEE 33rd Annual Int. Symp. Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Kyoto, Japan, 2022, pp. 246-252.

[152] T. Li et al., "Asynchronous MIMO-OFDM massive unsourced random access with codeword collisions," *IEEE Trans. Wireless Commun.*, vol. 24, no. 1, pp. 84-100, Jan. 2025.

[153] O. Musa, P. Jung and G. Caire, "Message and activity detection for an asynchronous random access receiver using AMP," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Rome, Italy, 2023, pp. 1940-1945.

[154] E. Gkiouzepi, B. Çakmak, M. Opper and G. Caire, "Joint message detection and channel estimation for unsourced random access in cell-free user-centric wireless networks," *IEEE Trans. Inf. Theory*, vol. 71, no. 5, pp. 3614-3643, May 2025.

[155] E. Gkiouzepi, B. Çakmak and G. Caire, "2-step RACH and unsourced massive random access in cell-free user-centric networks," in *Proc. 58th Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2024, pp. 1126-1133.

[156] Y. Hu et al., "Optimized unsourced random access schemes with sparse-correction-based approximate message passing for massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 74, no. 1, pp. 1104-1120, Jan. 2025.

[157] Z. Zhang, J. Dang, Z. Zhang, L. Wu and B. Zhu, "Uncoupled unsourced random access: Exploiting geographical diversity of access points," *IEEE Trans. Veh. Technol.*, early access, Jan. 20, 2025.

[158] M. Gkagkos, J.-F. Chamberland, C. N. Georghiades and K. R. Narayanan, "Scalable cell-free massive MIMO unsourced random access system," arXiv, 2023. Available: https://arxiv.org/abs/2304.06105.

[159] M. Ozates, M. Kazemi, E. Jorswieck and D. Gunduz, "ODMA-based cell-free unsourced random access with successive interference cancellation," arXiv, 2025. Available: https://arxiv.org/abs/2501.10201.

[160] W. W. Zheng, J. R. Ebert, S. Rini and J.-F. Chamberland, "Coding for the unsourced A-channel with erasures: The linked loop code," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Helsinki, Finland, 2023, pp. 1534-1538.

[161] W. W. Zheng, J. R. Ebert, S. Rini and J. -F. Chamberland, "Coding for the Unsourced B-Channel with erasures: Enhancing the linked loop Code," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Seoul, Republic of Korea, 2024, pp. 9221-9225.

[162] V. K. Amalladinne, A. Hao, S. Rini and J. -F. Chamberland, "Multi-class unsourced random access via coded demixing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2021, Melbourne, Australia, pp. 3080-3085.

[163] S. Rini, V. K. Amalladinne and J. -F. Chamberland, "An exploration of the heterogeneous unsourced multiple access," in *Proc. 55th Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2021, pp. 954-958.

[164] A. Fengler, A. Lancho, K. Narayanan and Y. Polyanskiy, "On the advantages of asynchrony in the unsourced MAC," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Taipei, Taiwan, 2023, pp. 2523-2528.

[165] J. H. I. de Souza and T. Abrão, "Deep learning-based activity detection for grant-free random access," *IEEE Syst. J.*, vol. 17, no. 1, pp. 940-951, Mar. 2023.

[166] Z. Zhang, J. Dang, Z. Zhang, L. Wu, B. Zhu and L. Wang, "Pilot-free unsourced random access via dictionary learning and error-correcting codes," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 7488-7502, July 2024.

[167] M. Chen et al., "Distributed Learning in Wireless Networks: Recent Progress and Future Challenges," in *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 12, pp. 3579-3605, Dec. 2021.

[168] M. M. Amiri and D. Gündüz, "Federated learning over wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3546-3557, May 2020.

[169] L. Qiao, Z. Gao, M. Boloursaz Mashhadi and D. Gunduz, "Massive digital over-the-air computation for communication-efficient federated edge learning," *IEEE J. Select. Areas Commun.*, vol. 42, no. 11, pp. 3078-3094, Nov. 2024.

[170] H. Nikbakht, M. Wigger, S. S. Shitz, and H. V. Poor, "A Memory-Based Reinforcement Learning Approach to Integrated Sensing and Communication," in *Proc. 2024 58th Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2024, pp. 433-437.

[171] M. J. Ahmadi, R. F. Schaefer, H. V. Poor, "Integrated sensing and communications for unsourced random access: Fundamental limits," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Cape Town, South Africa, 2024, pp. 1365-1370.

[172] M. J. Ahmadi, R. F. Schaefer, and H. V. Poor, "Integrated sensing and communications for unsourced random access: Fundamental limits and practical model," arXiv, 2024. Available: https://arxiv.org/abs/2404.19431.

[173] M. J. Ahmadi, R. F. Schaefer and H. V. Poor, "A practical framework for unsourced integrated sensing and communication," in *Proc. IEEE 26th*

Int. Workshop Signal Process. Artif. Intell. Wireless Commun. (SPAWC), Surrey, United Kingdom, 2025, pp. 1-5.

[174] Z. Zhang, J. Dang and Z. Zhang, "Integrated sensing and communications for unsourced random access: A two-phase approach with slotted transmission," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Shanghai, China, 2025, pp. 1-6.

[175] Z. Zhang, J. Dang, K.-K. Wong, Z. Zhang, and C. Masouros, "Integrated sensing and communications for unsourced random access: A spectrum sharing compressive sensing approach," arXiv, 2025. Available: https://arxiv.org/abs/2504.17629.

[176] Z. Zhang, K.-K. Wong, and J. Dang, "On fundamental limits for fluid antenna-assisted integrated sensing and communications for unsourced random access ," *IEEE J. Sel. Areas Commun.*, early access, 2025, doi: 10.1109/JSAC.2025.3608113.

[177] Z. Zhang, K.-K. Wong, J. Dang, Z. Zhang, C. Masouros, and C.-B. Chae, "On fundamental limits of slow-fluid antenna multiple access for unsourced random access," *IEEE Wireless Commun. Lett.*, 2025, doi: 10.1109/LWC.2025.3594112.

[178] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341-1365, Oct. 2008.

[179] J. R. Ebert, K. R. Narayanan and J. -F. Chamberland, "HashBeam: Enabling feedback through downlink beamforming in unsourced random access," in *Proc. Asilomar Conf. Signals, Systems, and Computers.*, Pacific Grove, CA, USA, 2022, pp. 692-697

[180] M. J. Ahmadi, M. Kazemi and R. F. Schaefer, "Efficient feedback design for unsourced random access with integrated sensing and communication," arXiv, 2025. Available: https://arxiv.org/abs/2506.20262.

[181] M. Bashir, E. Nassaji, D. Truhachev, A. Bayesteh and M. Vameghestahbanati, "Unsourced random access with threshold-based feedback," *IEEE Trans. Commun.*, vol. 71, no. 12, pp. 7072-7086, Dec. 2023.

[182] J. Che et al., "Unsourced multiple access for mission-critical control systems in industrial internet of things," *IEEE Internet Things J.*, vol. 11, no. 17, pp. 27955-27968, Sep. 2024.

[183] R. Kotaba et al., "Unsourced random access with authentication and joint downlink acknowledgements," in *Proc. Asilomar Conf. Signals, Systems, and Computers.*, Pacific Grove, CA, USA, 2021, pp. 1496-1501.

[184] I. Yildirim, F. Kilinc, E. Basar and G. C. Alexandropoulos, "Hybrid RIS-empowered reflection and decode-and-forward relaying for coverage extension," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1692-1696, May 2021.

[185] J. -B. Gros, V. Popov, M. A. Odit, V. Lenets and G. Lerosey, "A reconfigurable intelligent surface at mmWave based on a binary phase tunable metasurface," *IEEE open J. Commun. Soc.*, vol. 2, pp. 1055-1064, 2021.

[186] D. Ma and Y. Shi, "A lightweight encryption algorithm for edge networks in software-defined industrial Internet of things," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC Workshops)*, Chengdu, China, 2019, pp. 1489-1493.

[187] A. Esfahani et al., "A lightweight authentication mechanism for M2M communications in industrial IoT environment," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 288-296, Feb. 2019.

[188] O. Ureten and N. Serinken, "Wireless security through RF fingerprinting," *Can. J. Electr. Comput. Eng.*, vol. 32, no. 1, pp. 27-33, Winter 2007.

[189] P. Agostini et al., "Enhancements to the 5G-NR 2-step RACH: An unsourced multiple access perspective," in *Proc. IEEE Conf. Standards Commun. Netw. (CSCN)*, Belgrade, Serbia, 2024, pp. 32-35.