IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 October 2023, accepted 20 November 2023, date of publication 24 November 2023,
date of current version 29 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3336677

RESEARCH ARTICLE

Performance Analysis and Phase Shift Design of IRS-Aided Uplink OTFS-SCMA

JIARUI LI, (Graduate Student Member, IEEE), AND YI HONG, (Senior Member, IEEE)
Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, VIC 3800, Australia

Corresponding author: Jiarui Li (jiarui.li@monash.edu)

This work was supported by the Australian Research Council through Discovery Project under Grant DP200100096.

ABSTRACT In this paper, we consider an intelligent reflecting surface (IRS)-aided orthogonal time frequency space (OTFS)-based uplink sparse code multiple access (SCMA) communications system. We first conduct performance analysis in terms of pairwise error probability (PEP) and derive an upper bound on word error probability (WEP). According to this bound, we establish a system design criterion and propose two IRS phase shifts design algorithms using semidefinite relaxation (SDR) and gradient ascent (GA) methods. The computational complexity of these algorithms is discussed. Next, we derive an upper bound on the average bit error rate (BER) and investigate the system performance in terms of diversity and signal-to-noise (SNR) gains. Further, to recover transmitted information bits, we present a modified joint iterative Gaussian approximated message passing (MP) detection and SCMA decoding algorithm, enabling detection and SCMA demapping to occur in each iteration. Finally, simulation results demonstrate that our proposed IRS design algorithms achieve better error performance compared to the known approaches.

INDEX TERMS Intelligent reflecting surface (IRS), orthogonal time frequency space (OTFS), sparse code multiple access (SCMA), word error probability, phase shifts design.

I. INTRODUCTION

With the popularizing of devices in fifth-generation (5G) and beyond wireless networks, the demand for accommodating an ever-increasing number of users while improving spectral efficiency has rapidly grown. Non-orthogonal multiple access (NOMA) has drawn much attention as a solution, allowing multiple users to simultaneously access limited resource elements (REs), such as frequency bands and time slots [1], [2]. Among various NOMA solutions, sparse code multiple access (SCMA), a type of code-domain NOMA, achieves multiplexing in code domain and is able to effectively distinguish multiple users by assigning unique sparse codebooks [3], [4], [5].

On the other hand, as high-speed trains and vehicles become increasingly prevalent, the need for wireless communications in high-mobility scenarios has led to the development of a novel technique, orthogonal time frequency space (OTFS) (see [6], [7], [8], [9], [10], [11], [12], [13] and references therein). In OTFS, the modulation operates on delay-Doppler (DD) domain instead of the conventional time-frequency (TF) domain, enabling it to combat the dynamics of time-varying multipath channel [8]. To further improve spectrum efficiency and support massive connectivity in high-mobility communications, the integration of SCMA with OTFS has attracted tremendous research attention [14], [15], [16].

Recently, the introduction of intelligent reflecting surface (IRS) has been identified as a promising solution to enhance system performance in 5G and beyond networks (see [17], [18], [19], [20] and references therein). With the ability to reshape electromagnetic wave propagation through passive reconfiguration, IRS presents its potential to create smart radio environment. By controlling all passive elements at IRS, we can change the phases of incident radio frequency signals to support diverse user requirements [17]. Considering the radio environment reconfigurable feature offered by IRS, it is expected that the combination of IRS and OTFS-SCMA

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

can further enhance overloaded multiuser system under high-mobility scenarios [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31].

A. IRS-AIDED SCMA

Several research has been conducted to investigate system model and performance of IRS-aided SCMA system (see [21], [22], [23], [24], [25], [26] and references therein). In [21], the authors demonstrated transmission and detection scheme for IRS-aided uplink SCMA. In [22], the authors considered coherent IRS beamforming design and tackled diversity analysis for uplink IRS-assisted SCMA system, its extension to downlink can be found in [23]. In [24], the authors proposed an alternating optimization (AO)-based algorithm for IRS beamforming design, while in [25], the authors presented an algorithm with lower computational complexity. In [26], the authors considered sum-rate maximization problem for IRS-aided downlink SCMA.

B. IRS-AIDED OTFS

Considering high-mobility communications scenarios, IRS-aided OTFS has been investigated in [27], [28], [29], [30], and [31]. In [27], the authors demonstrated the transmission scheme and input-output relationship of IRS-assisted OTFS in TF domain and DD domain, respectively. In [28], the authors considered a single-input single-output (SISO) system and designed IRS phase shifts by generating a large number of random sampling to maximize the Frobenius norm of effective end-to-end DD domain channel matrix. A follow-up research which considered fractional delay and Doppler values was demonstrated in [29]. In [30], the authors investigated OTFS-based IRS-aided space–air–ground integrated networks (SAGINs) and IRS phase shifts were designed according to line-of-sight (LoS) path gain. In [31], the authors investigated an IRS-aided multiple-input multiple-output (MIMO)-OTFS system under single-user scenario and proposed a matching-based IRS beamforming design algorithm, where IRS elements were configured by matching with the strongest cascaded channel path gain.

C. OUR CONTRIBUTIONS

In this paper, we consider an IRS-aided uplink OTFS-SCMA communications system, where multiple high-mobile users simultaneously communicate to a multi-antenna base station (BS) over a limited number of REs. An IRS is adopted between users and BS to assist communications. We assume all users are in high mobility scenarios where the users-IRS links are time-varying multipath channels. The transmit information symbols are first mapped to SCMA codewords, followed by OTFS modulation. At BS, we perform OTFS demodulation and recover the information bits via joint detection and SCMA decoding algorithm. The detailed contributions are summarized below.

- We present an IRS-aided OTFS-SCMA system and conduct performance analysis in terms of pairwise error probability (PEP) conditioned on the channel realization and union bound on the word error probability (WEP).
- To minimize the union bound on WEP, we design IRS phase shifts using the following two approaches: *semidefinite relaxation (SDR)-based approach and gradient ascent (GA)-based approach*. The computational complexity is provided. By simulation, we show that both algorithms can achieve better error performance compared to the known approaches.
- We investigate the *average* error performance over all channel realizations and derive an upper bound on average bit error rate (BER). We also demonstrate the diversity and signal-to-noise (SNR) gain provided by the proposed system.
- We present a joint modified Gaussian approximated message passing (MP) detection and SCMA decoding to recover transmit bits of all users. Compared to the original message passing algorithm (MPA) in [8], we make several adjustments to the proposed IRS-aided OTFS-SCMA system. The detailed detection and decoding procedures are also illustrated.

The rest of the paper is organized as follows. Section II introduces the system model. Section III presents problem formulation and system design criterion, followed by design of IRS phase shifts in Section IV. Section V provides average BER analysis. Section VI demonstrates joint detection and SCMA decoding procedure. Simulation results and conclusions are given in Section VII and Section VIII, respectively.

Notations: Vectors are boldface letters and matrices are boldface capital letters. The notation $\mathbf{I}_n$ denotes identity matrix with size $n \times n$. Then, $\mathbf{F}_n$ represents a n-point discrete Fourier transform (DFT) matrix and ${\mathbf{F}_n}^H$ represents a n-point inverse DFT matrix. The notation $\mathbb{1}_n$ denotes an all-ones row vector of size n. The notations $\mathrm{tr}(\cdot)$, $\mathbb{E}[\cdot]$, $\mathrm{Var}[\cdot]$, $\mathrm{Cov}(\cdot,\cdot)$, $\|\cdot\|$, and $|\cdot|$ denote trace, expectation, variance, covariance, norm, and cardinality operators. Superscripts $(\cdot)^T$ represents transpose and $(\cdot)^H$ represents conjugate transpose. The notations $\mathbb{Z}^{a\times b}$, $\mathbb{R}^{a\times b}$ and $\mathbb{C}^{a\times b}$ represent space of matrix with size $a \times b$ in integer value, real value and complex value, respectively. The operators $\Re(\cdot)$ and $\Im(\cdot)$ are used to extract real and imaginary parts, and then $\arg(\cdot)$ denotes phase of complex number. The operators $\otimes$ and $\odot$ represent Kronecker product and Hadamard product, respectively. The notation $\mathrm{vec}(\cdot)$ represents vectorization operator. $Q(\cdot)$ denotes Q-function. The operator $\binom{\cdot}{\cdot}$ represents binomial coefficients, and $\mathbf{A}\,[a, c:d]$ represents a-th row, from c-th to d-th columns entries in matrix $\mathbf{A}$. The notation $\mathcal{C}^n$ represents n-ary Cartesian power of a set $\mathcal{C}$.

II. SYSTEM MODEL

We consider an IRS-aided multiuser uplink communications system as shown in Fig. 1. We assume K single-antenna users simultaneously communicating to a BS of N_r antennas via IRS over $J < K$ REs. We assume all users are in a similar high

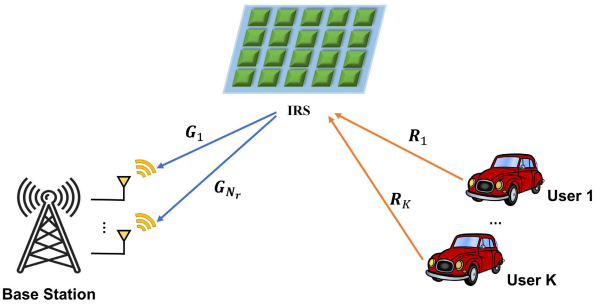

FIGURE 1. An illustration of an IRS-aided multiuser uplink communications system.

mobility environment and the user-IRS link is a time-varying multipath (P paths) channel with the maximum delay $\tau_{\max}$ and Doppler shift $\nu_{\max}$. Further, we assume all the user-IRS links have the same delay and Doppler shifts profiles but independent complex channel gains.

To support such an overloaded IRS-aided uplink system, we adopt an OTFS-based SCMA scheme using $\mathcal{M}$-quadrature amplitude modulation (QAM) signaling, where the detailed system block diagram is given in Fig. 2. For OTFS, we assume that the k-th user, for $k = 1, \ldots, K$, has a bandwidth $B = M\Delta f = \frac{M}{T}$ and a frame duration $T_f = NT$, where $\Delta f = \frac{1}{T}$ denotes subcarrier spacing, $M, N \in \mathbb{Z}$ denote delay and Doppler dimonsions, respectively [7]. The choices of T and Δf should meet the requirements $\nu_{\max} < 1/T$ and $\tau_{\max} < 1/\Delta f$ [8].

A. TRANSMITTER SIDE

As shown in Fig. 2, the k-th user transmits the bit sequence

$$\mathbf{b}_k = \left[b_1^k, \ldots, b_{\frac{MN}{J}\log_2 \mathcal{M}}^k\right],$$

of length $\frac{MN}{J}\log_2 \mathcal{M}$, which is further mapped into $\mathcal{M}$-QAM symbol sequence

$$\mathbf{u}_k = \left[u_1^k, \ldots, u_{\frac{MN}{J}}^k\right],$$

with cardinality $|\mathbf{u}_k| = \frac{MN}{J}$, where the j-th symbol of k-th user, u_j^k, $j \in [1, \frac{MN}{J}]$, $k \in [1, K]$, contains $\log_2 \mathcal{M}$ bits. The symbol sequence $\mathbf{u}_k$ is further processed by an SCMA encoder, mapping each symbol to a unique codeword $\mathbf{c}_j^k$ of length J, i.e., $|\mathbf{c}_j^k| = J$, taken from a codebook $\mathcal{C}_k$ with cardinality $|\mathcal{C}_k| = \mathcal{M}$, leading to a codeword sequence

$$\mathbf{c}_k = \left[\mathbf{c}_1^k, \ldots, \mathbf{c}_{\frac{MN}{J}}^k\right].$$

We then place the sequence $\mathbf{c}_k$ in the DD domain along the delay axis[1] [14], followed by OTFS modulation. For simplicity, we assume that M and N are *integer multiples of SCMA codeword length* J. Hence, we can place a total $\frac{MN}{J}$ codewords, each of length J, in an OTFS DD grid of size $M \times N$, yielding

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{c}_{\frac{M}{J}}^k & \cdots & \mathbf{c}_{\frac{MN}{J}}^k \\ \vdots & \ddots & \vdots \\ \mathbf{c}_1^k & \cdots & \mathbf{c}_{\frac{M(N-1)}{J}+1}^k \end{bmatrix}. \tag{1}$$

We illustrate an example in Fig. 3, with $M = N = 8$, $J = 4$, and 16 codewords along the delay axis.

Considering all K users transmit a total number of $\frac{MNK}{J}$ codewords over MN DD domain REs, then the *overloading factor* of the OTFS-SCMA system is $\Gamma_o = K/J$.

Assuming rectangular pulse shaping, we perform OTFS via inverse discrete Zak transform (IDZT) to convert $\mathbf{X}_k$ from DD domain to time-domain as [31],

$$\mathbf{s}_k = \left(\mathbf{F}_N{}^H \otimes \mathbf{I}_M\right) \mathrm{vec}\left(\mathbf{X}_k\right) = \left(\mathbf{F}_N{}^H \otimes \mathbf{I}_M\right) \mathbf{x}_k, \tag{2}$$

where $\mathbf{x}_k = \mathrm{vec}\left(\mathbf{X}_k\right) \in \mathbb{C}^{MN}$ is the k-th user's codewords vector, taking from set $\mathcal{C}_k^{\frac{MN}{J}}$ containing $\frac{MN}{J}$ codewords in the DD domain.

B. IRS AND CHANNEL MODEL

An IRS with Q passive reflecting elements is adopted to assist communication between users and BS. An embedded controller can control phases and amplitudes of all elements. Let $\boldsymbol{\theta} = \left[\beta_1 e^{j\theta_1}, \ldots, \beta_Q e^{j\theta_Q}\right]^T$ denote IRS reflection vector, where $\theta_q \in [0, 2\pi)$ and $\beta_q \in [0, 1]$, for $q = 1, \ldots, Q$, represent reflection phase and amplitude, respectively. For simplicity, we assume full reflection and $\beta_q = 1, \forall q$ [25].

For the channel setting, we assume all users have the same delay and Doppler shifts profiles but independent complex channel gains. We let $\mathbf{R}_k = \left[\bar{\mathbf{R}}_{k1}^T, \ldots, \bar{\mathbf{R}}_{kQ}^T\right]^T \in \mathbb{C}^{QMN \times MN}$ be the time-domain channel matrix between k-th user and IRS, where

$$\bar{\mathbf{R}}_{kq} = \sum_{p=1}^{P} \gamma_{kq}^p \mathbf{\Pi}^{l_p} \mathbf{\Delta}^{k_p}, \tag{3}$$

denotes the time-varying channel between k-th user and q-th IRS element, for $k \in [1, K]$, $q \in [1, Q]$, $p \in [1, P]$, respectively, γ_{kq}^p, l_p and k_p denote complex channel gain, delay, and Doppler shift of p-th path, $\mathbf{\Pi}$ is cyclic-shift matrix related to delay given in [9] and $\mathbf{\Delta} = \mathrm{diag}\{z^0, z^1, \ldots, z^{MN-1}\}$, where $z = e^{j\frac{2\pi}{MN}}$, is a diagonal matrix related to the Doppler shifts [9].

The IRS-BS channel links can be assumed to be LoS [25], [18], since, in practice, the location of BS is usually fixed and an IRS can be deployed at a desirable location. Specifically, the channel matrix of IRS to i-th receive antenna is denoted as $\mathbf{G}_i = \left[g_1^i \mathbf{I}_{MN}, \ldots, g_Q^i \mathbf{I}_{MN}\right] \in \mathbb{C}^{MN \times QMN}$, where g_q^i denotes channel gain between q-th IRS element and i-th receive antenna. We can write down the cascaded channel matrix between k-th user and i-th receive antenna at BS as

$$\mathbf{H}_k^i = \mathbf{G}_i \mathbf{\Theta} \mathbf{R}_k, \tag{4}$$

[1] The sequence can also be placed along Doppler axis. It was shown in [14] that it would result in similar performance.

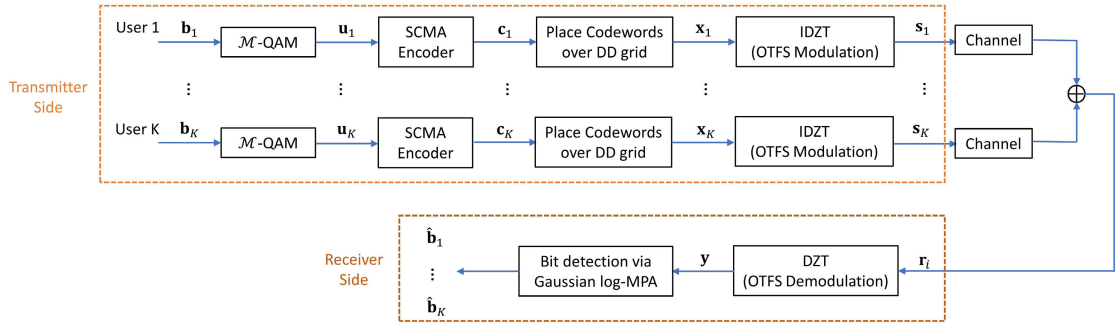

FIGURE 2. The block diagram of IRS-aided OTFS-SCMA multiuser uplink communications.

where $\mathbf{\Theta} = \text{diag}\{\boldsymbol{\theta}\} \otimes \mathbf{I}_{MN} \in \mathbb{C}^{QMN \times QMN}$ is the matrix containing all the IRS reflection phases, for $k \in [1, K]$, $i \in [1, N_r]$.

C. RECEIVER SIDE

At i-th receive antenna of the BS, for $i \in [1, N_r]$, the time-domain received signal, $\mathbf{r}_i \in \mathbb{C}^{MN}$, is given by

$$\mathbf{r}_i = \sum_{k=1}^{K} \mathbf{H}_k^i \mathbf{s}_k + \mathbf{n}, \tag{5}$$

where $\mathbf{n}$ is additive white Gaussian noise (AWGN) vector with each element following $\mathcal{CN}(0, N_0)$, and N_0 is noise power. We then perform OTFS demodulation via DZT to convert $\mathbf{r}_i$ back to DD domain, yielding [31]

$$\mathbf{y}_i = (\mathbf{F}_N \otimes \mathbf{I}_M)\, \mathbf{r}_i \in \mathbb{C}^{MN}. \tag{6}$$

For the ease of notation, we define

$$\mathbf{W}_t = {\mathbf{F}_N}^H \otimes \mathbf{I}_M, \ \mathbf{W}_r = \mathbf{F}_N \otimes \mathbf{I}_M, \tag{7}$$

and

$$\tilde{\mathbf{W}}_r = \mathbf{I}_{N_r} \otimes \mathbf{W}_r, \ \tilde{\mathbf{W}}_t = \mathbf{I}_K \otimes \mathbf{W}_t. \tag{8}$$

Based on (2), (5), (6), the received signal at BS can be expressed as

$$\mathbf{y} = \left[(\mathbf{y}_1)^T, \ldots, \left(\mathbf{y}_{N_r}\right)^T \right]^T = \underbrace{\tilde{\mathbf{W}}_r \mathbf{H}_{\text{eff}} \tilde{\mathbf{W}}_t}_{\mathbf{H}_{\text{all}}} \mathbf{x}_{\text{all}} + \mathbf{n}, \tag{9}$$

where

$$\mathbf{H}_{\text{eff}} = \begin{bmatrix} \mathbf{H}_1^1 & \cdots & \mathbf{H}_K^1 \\ \vdots & \ddots & \vdots \\ \mathbf{H}_1^{N_r} & \cdots & \mathbf{H}_K^{N_r} \end{bmatrix} \in \mathbb{C}^{N_r MN \times KMN}. \tag{10}$$

In (9), $\mathbf{x}_{\text{all}} = \left[\mathbf{x}_1^T, \ldots, \mathbf{x}_K^T\right]^T \in \mathbb{C}^{KMN}$ is a collection of all users' codewords in DD domain and is taken from codewords set $\bar{\mathcal{C}} = \left\{ \mathcal{C}_1^{\frac{MN}{J}} \times \ldots \times \mathcal{C}_K^{\frac{MN}{J}} \right\}$. Assuming channel state information (CSI) is available at the BS [10], [32], [33], we perform log-domain MP detection at BS to obtain the estimated bit sequence $\left[\hat{\mathbf{b}}_1, \ldots, \hat{\mathbf{b}}_K\right]$ of length $\frac{KMN}{J} \log_2 \mathcal{M}$.

III. PROBLEM FORMULATION AND SYSTEM DESIGN CRITERION

Based on channel models in (3) and (4), let $\boldsymbol{\gamma}_k = \{\gamma_{kq}^p\}_{q=1,p=1}^{Q,P} \in \mathbb{C}^{PQ}$ be the channel gain vector of the k-th user-IRS link, and $\mathbf{g}_i = \{g_q^i\}_{q=1}^{Q} \in \mathbb{C}^{Q}$ be the channel gain vector of the IRS-ith antenna link.

Following the similar derivation to [34], we can rewrite the received signal in (9) as

$$\mathbf{y} = \mathbf{\Xi}_{\text{all}}(\mathbf{x}_k)\mathbf{h}_{\text{all}} + \mathbf{n}, \tag{11}$$

where $\mathbf{\Xi}_{\text{all}}(\mathbf{x}_k) \in \mathbb{C}^{N_r MN \times N_r PQK}$ and $\mathbf{h}_{\text{all}} \in \mathbb{C}^{N_r PQK}$ are given by

$$\mathbf{\Xi}_{\text{all}}(\mathbf{x}_k) = \mathbf{I}_{N_r} \otimes \left[\mathbf{\Xi}_1(\mathbf{x}_k), \ldots, \mathbf{\Xi}_K(\mathbf{x}_k)\right],$$
$$\mathbf{h}_{\text{all}} = \left[(\mathbf{h}_1^1)^T, \ldots, (\mathbf{h}_K^1)^T, \ldots, (\mathbf{h}_1^{N_r})^T, \ldots, (\mathbf{h}_K^{N_r})^T \right]^T,$$

with

$$\begin{aligned} \mathbf{\Xi}_k(\mathbf{x}_k) &= \left[\mathbb{1}_Q \otimes (\mathbf{\Phi}_1 \mathbf{x}_k), \ldots, \mathbb{1}_Q \otimes (\mathbf{\Phi}_P \mathbf{x}_k)\right], \\ \mathbf{\Phi}_p &= \mathbf{W}_r \mathbf{\Pi}^{l_p} \mathbf{\Delta}^{k_p} \mathbf{W}_t, \\ \mathbf{h}_k^i &= \left(\left[\mathbb{1}_P \otimes (\mathbf{g}_i \odot \boldsymbol{\theta})^T\right] \odot \boldsymbol{\gamma}_k^T \right)^T, \end{aligned} \tag{12}$$

for $k \in [1, K]$, $i \in [1, N_r]$.

Next, we analyze the error performance of the system in terms of the PEP and union bound. Assuming maximum likelihood (ML) detection is used,[2] we can write the estimated $\mathbf{x}_{\text{all}}$ as

$$\hat{\mathbf{x}}_{\text{all}} = \arg \min_{\mathbf{x}_{\text{all}} \in \bar{\mathcal{C}}} ||\mathbf{y} - \mathbf{\Xi}_{\text{all}}(\mathbf{x}_k)\mathbf{h}_{\text{all}}||^2. \tag{13}$$

For a given channel realization $\mathbf{h}_{\text{all}}$, the WEP is bounded by [34]

$$\begin{aligned} P_e(\mathbf{h}_{\text{all}}) &= \frac{1}{\alpha} \sum_{\mathbf{x}_{\text{all}} \in \bar{\mathcal{C}}} P_e\left(\mathbf{x}_{\text{all}} | \mathbf{h}_{\text{all}}\right) \\ &\leq \frac{1}{\alpha} \sum_{\mathbf{x}_{\text{all}} \in \bar{\mathcal{C}}} \sum_{\substack{\mathbf{x}_{\text{all}} \neq \hat{\mathbf{x}}_{\text{all}} \\ \mathbf{x}_{\text{all}}, \hat{\mathbf{x}}_{\text{all}} \in \bar{\mathcal{C}}}} P_e\left(\mathbf{x}_{\text{all}} \to \hat{\mathbf{x}}_{\text{all}} | \mathbf{h}_{\text{all}}\right), \end{aligned} \tag{14}$$

where $\alpha = \mathcal{M}^{\frac{MNK}{J}}$.

[2] ML detection is considered only for PEP analysis. In practice, receiver can use a standard MPA detection. In this paper, we utilize Gaussian log-MPA detection as demonstrated in the Section VI.

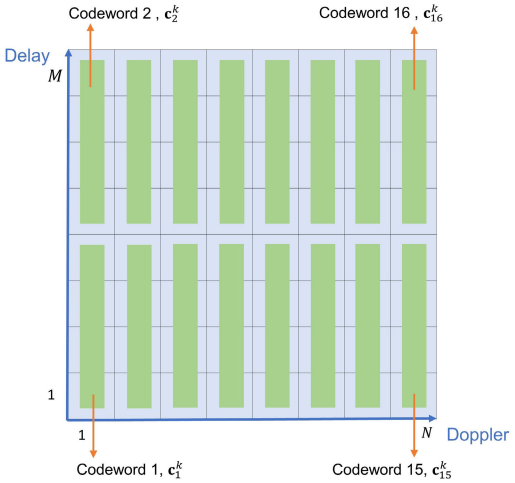

FIGURE 3. An illustration of placement of SCMA codewords of k-th user over DD plane to form DD domain matrix $\mathbf{X}_k \in \mathbb{C}^{M\times N}$, where $J = 4$, $M = 8$ and $N = 8$.

Let us define the set of all codeword difference vectors as

$$\mathcal{L} = \{\boldsymbol{\delta} = \mathbf{x}_{\text{all}} - \hat{\mathbf{x}}_{\text{all}} | \mathbf{x}_{\text{all}} \neq \hat{\mathbf{x}}_{\text{all}},\ \mathbf{x}_{\text{all}}, \hat{\mathbf{x}}_{\text{all}} \in \bar{\mathcal{C}}\}. \tag{15}$$

Let

$$\boldsymbol{\Xi}_{\text{all}}(\boldsymbol{\delta}) = \boldsymbol{\Xi}_{\text{all}}(\mathbf{x}_k) - \boldsymbol{\Xi}_{\text{all}}(\hat{\mathbf{x}}_k).$$

Then the PEP for a given channel realization $\mathbf{h}_{\text{all}}$ is

$$P_e\left(\mathbf{x}_{\text{all}} \to \hat{\mathbf{x}}_{\text{all}} | \mathbf{h}_{\text{all}}\right) = Q\left(\sqrt{\frac{||\boldsymbol{\Xi}_{\text{all}}(\boldsymbol{\delta})\mathbf{h}_{\text{all}}||^2}{2N_0}}\right), \tag{16}$$

which represents the probability of detecting $\hat{\mathbf{x}}_{\text{all}} \neq \mathbf{x}_{\text{all}}$, when $\mathbf{x}_{\text{all}}$ is transmitted, for $\mathbf{x}_{\text{all}}, \hat{\mathbf{x}}_{\text{all}} \in \bar{\mathcal{C}}$.

Observing that

$$||\boldsymbol{\Xi}_{\text{all}}(\boldsymbol{\delta})\mathbf{h}_{\text{all}}||^2 = \mathbf{h}_{\text{all}}^H \mathbf{A}(\boldsymbol{\delta})\mathbf{h}_{\text{all}}, \tag{17}$$

with $\mathbf{A}(\boldsymbol{\delta}) = \boldsymbol{\Xi}_{\text{all}}^H(\boldsymbol{\delta})\boldsymbol{\Xi}_{\text{all}}(\boldsymbol{\delta})$ denoting *delay-Doppler shifted codeword distance matrix*. We have the following Proposition.

Proposition 1: For a given channel realization $\mathbf{h}_{\text{all}}$ *and a pre-defined SCMA codebook, we obtain the upper bound on the WEP as*

$$\begin{aligned} P_e(\mathbf{h}_{\text{all}}) &\leq \frac{1}{\alpha}\sum_{\mathbf{x}_{\text{all}}\in\bar{\mathcal{C}}}\sum_{\boldsymbol{\delta}\in\mathcal{L}} Q\left(\sqrt{\frac{\mathbf{h}_{\text{all}}^H\mathbf{A}(\boldsymbol{\delta})\mathbf{h}_{\text{all}}}{2N_0}}\right) \\ &\leq \frac{1}{\alpha}\sum_{\mathbf{x}_{\text{all}}\in\bar{\mathcal{C}}}\sum_{\boldsymbol{\delta}\in\mathcal{L}} \exp\left(-\frac{\mathbf{h}_{\text{all}}^H\mathbf{A}(\boldsymbol{\delta})\mathbf{h}_{\text{all}}}{4N_0}\right) \\ &\leq \frac{-\eta}{4\alpha N_0}\sum_{\mathbf{x}_{\text{all}}\in\bar{\mathcal{C}}}\left(\mathbf{h}_{\text{all}}^H\left[\sum_{\boldsymbol{\delta}\in\mathcal{L}}\mathbf{A}(\boldsymbol{\delta})\right]\mathbf{h}_{\text{all}} + 1\right), \end{aligned} \tag{18}$$

where η *is a small positive real number.*

Proof: See Appendix A. ■

Based on Proposition 1, we have the following system design criterion.

Remark 1 (System Design Criterion): To minimize WEP of proposed IRS-aided communications system for a given channel realization, we need to minimize upper bound in (18) by designing IRS phase shifts vector $\boldsymbol{\theta}$ *associated with cascaded channel vector* $\mathbf{h}_{\text{all}}$.

For notation simplification purpose, we let

$$\boldsymbol{\Psi} = \sum_{\mathbf{x}_{\text{all}}}\sum_{\boldsymbol{\delta}\in\mathcal{L}}\mathbf{A}(\boldsymbol{\delta}).$$

According to the design criterion, given pre-defined SCMA codebook and fixed channel condition $\mathbf{h}_{\text{all}}$, we can formulate the objective problem as follows,

$$\begin{aligned} \text{(P1):}\ &\max_{\boldsymbol{\theta}}\ \mathbf{h}_{\text{all}}^H\boldsymbol{\Psi}\mathbf{h}_{\text{all}} \\ &\text{s.t. } \boldsymbol{\theta} = \left[e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_Q}\right]^T \end{aligned} \tag{19}$$

IV. DESIGN OF IRS BEAMFORMING

In this section, we provide two optimization algorithms aimed at addressing IRS beamforming design problem (P1).

A. SDR-BASED APPROACH

Let $\boldsymbol{\gamma}_{kp} = \left[\gamma_{k1}^p, \ldots, \gamma_{kQ}^p\right]^T \in \mathbb{C}^Q$ denote the complex path gains vector between all IRS elements and k-th user on p-th channel path. Based on (11), the cascaded channel link vector $\mathbf{h}_{\text{all}}$ can be re-expressed as

$$\mathbf{h}_{\text{all}} = \tilde{\mathbf{h}}\boldsymbol{\theta}, \tag{20}$$

where

$$\tilde{\mathbf{h}} = \left[\mathbf{B}_1^{1\,T}, \ldots, \mathbf{B}_K^{1\,T}, \ldots, \mathbf{B}_1^{N_r\,T}, \ldots, \mathbf{B}_K^{N_r\,T}\right]^T, \tag{21}$$

with $\mathbf{B}_k^i = \left[(\mathbf{B}_{k,1}^i)^T, \ldots, (\mathbf{B}_{k,P}^i)^T\right]^T \in \mathbb{C}^{PQ\times Q}$, for $k = 1, \ldots, K$, $i = 1, \ldots, N_r$, where

$$\mathbf{B}_{k,p}^i = \text{diag}\{\mathbf{g}_i \odot \boldsymbol{\gamma}_{kp}\}, \tag{22}$$

for $p = 1, \ldots, P$.

From (19) and (20), we can obtain the objective function as follows:

$$\begin{aligned} \text{(P2):}\ &\max_{\boldsymbol{\theta}}\ \boldsymbol{\theta}^H\underbrace{\tilde{\mathbf{h}}^H\boldsymbol{\Psi}\tilde{\mathbf{h}}}_{\mathbf{W}}\boldsymbol{\theta} \\ &\text{s.t. } \boldsymbol{\theta} = \left[e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_Q}\right]^T \end{aligned} \tag{23}$$

Let $\boldsymbol{\Theta} = \boldsymbol{\theta}\boldsymbol{\theta}^H$, which is a positive semidefinite matrix with rank one. Since $\text{rank}(\boldsymbol{\Theta}) = 1$ is a non-convex constraint, we apply SDR to relax this constraint and problem (P2) reduces to

$$\begin{aligned} \text{(P3):}\ &\max_{\boldsymbol{\Theta}}\ \text{tr}(\mathbf{W}\boldsymbol{\Theta}) \\ &\text{s.t. } \boldsymbol{\Theta}_{q,q} = 1, q = 1, 2, \ldots, Q, \\ &\quad \boldsymbol{\Theta} \succeq 0 \end{aligned} \tag{24}$$

The semidefinite problem (P3) can then be solved by existing convex optimization solver, such as CVX [35]. Generally, the

solution generated by CVX solver does not meet the rank one constraint. Thus a low-rank approximation needs to be conducted. We first apply eigenvalue decomposition on CVX solver solution on problem (P3) as $\hat{\boldsymbol{\Theta}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is unitary matrix and $\boldsymbol{\Lambda}$ is diagonal matrix, both with sizes $Q \times Q$. Then we generate rank-one approximation by using Gaussian randomization technique [36]. We can construct

$$\tilde{\boldsymbol{\Theta}} = (\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{w})^H(\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{w}), \tag{25}$$

where $\mathbf{w} \in \mathbb{C}^Q$ follows $\mathbf{w} \sim \mathcal{CN}(0, \mathbf{I}_Q)$. By generating a large amount of realizations of $\mathbf{w}$, the optimized IRS phase shift vector

$$\boldsymbol{\theta}^*_{\text{SDR}} = \exp\left(j \cdot \arg(\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{w})\right). \tag{26}$$

is chosen to maximize problem (P3) among all $\mathbf{w}$'s.

B. GA-BASED APPROACH

It can be observed that (23) can be regarded as a multi-variable function of IRS phase shifts $\boldsymbol{\theta}$, which can be maximized by updating parameters in the direction of positive gradient. Define $F(\boldsymbol{\theta}) = \boldsymbol{\theta}^H\mathbf{W}\boldsymbol{\theta}$. With an initial set of $\boldsymbol{\theta}^{(1)}$, the parameters of $F(\boldsymbol{\theta})$ are updated according to

$$\boldsymbol{\theta}^{(n+1)} = \boldsymbol{\theta}^{(n)} + \delta\nabla F(\boldsymbol{\theta}), \tag{27}$$

where $\delta \in \mathbb{R}$ represents updating step size and $\nabla F(\boldsymbol{\theta})$ represents gradient of function, which is given by

$$\nabla F(\boldsymbol{\theta}) = -2\Im\left(\boldsymbol{\Omega}\mathbf{t}^T\right), \tag{28}$$

where

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\omega}_1 & & & \\ & \boldsymbol{\omega}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{\omega}_Q \end{bmatrix} \in \mathbb{C}^{Q\times Q(Q-1)} \tag{29}$$

is composed of block submatrix $\boldsymbol{\omega}_q \in \mathbb{C}^{1\times(Q-1)}$, which is transpose of q-th column of $\mathbf{W}$ without diagonal elements for $q \in [1, Q]$. Also, $\mathbf{t}^T = \{e^{j(\theta_a-\theta_b)}\}_{a,b=1,a\neq b}^{Q} \in \mathbb{C}^{1\times Q(Q-1)}$, where $e^{j(\theta_a-\theta_b)}$, $a \neq b$, are placed in vector $\mathbf{t}^T$ according to a and b's ascending order. For example, with $Q = 3$, we have $\mathbf{t}^T = [e^{j(\theta_1-\theta_2)}, e^{j(\theta_1-\theta_3)}, e^{j(\theta_2-\theta_1)}, e^{j(\theta_2-\theta_3)}, e^{j(\theta_3-\theta_1)}, e^{j(\theta_3-\theta_2)}]$. The above parameter updating process is repeated until reaches convergence, i.e., the difference of objective function values in (23) between two consecutive iterations falls below threshold ς. During each iteration, the step size δ is selected to ensure that values in (23) increase as $\boldsymbol{\theta}$ updates. This indicates that (23) converges to at least a local maxima.

C. COMPLEXITY ANALYSIS

In Table 1, we compare computational complexity among our two proposed IRS design algorithms and matching-based algorithm in [31], where ε denotes threshold for SDR algorithm and I_t denotes iteration number of GA-based approach until reach convergence.

TABLE 1. Comparison of computational complexity.

SDR-based approach	$\mathcal{O}\left((Q^{6.5}+Q^{3.5})\log(\frac{1}{\varepsilon})\right)$
GA-based approach	$\mathcal{O}\left(Q^2 I_t\right)$
Matching-based approach	$\mathcal{O}\left(N_r PKQ\right)$

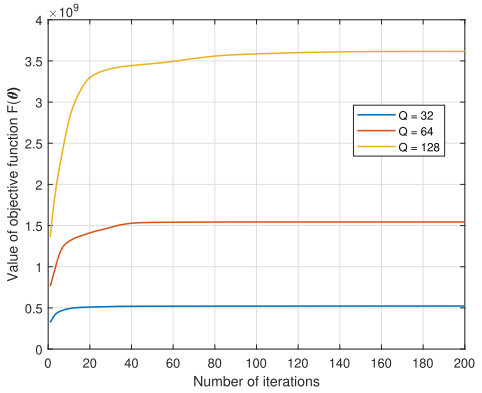

FIGURE 4. Convergence behavior of GA-based approach.

Specifically, the semidefinite problem (P3) can be solved by CVX solver based on interior point method, which has worst-case complexity $\mathcal{O}(Q^6 + Q^3)$ during each iteration and requires a total number of $\mathcal{O}\left(\sqrt{Q}\log(\frac{1}{\varepsilon})\right)$ iterations to generate ε-optimal solution [37]. The complexity of GA-based approach mainly comes from calculating $\nabla F(\boldsymbol{\theta})$ in (28) when updating IRS phase shifts, which requires $\mathcal{O}\left(Q^2\right)$ due to sparsity of matrix $\boldsymbol{\Omega}$. For matching-based approach in [31], each IRS phase shift, θ_q, for $q \in [1, Q]$, is determined by exhaustively searching all $N_r PK$ pairs of g_q^i and γ_{kq}^p, for $i \in [1, N_r]$, $k \in [1, K]$, $p \in [1, P]$, to identify the cascaded channel pairs with largest gain.

In Fig. 4, we demonstrate the convergence behavior of GA-based algorithm. It can be seen that the GA approach reaches convergence within about $I_t = 100$ iterations among different values of IRS elements number. From Table. 1, we observe that matching-based approach has the lowest complexity. However, it also exhibits worse error performance compared to our proposed IRS beamforming design, as demonstrated in Section VII. In addition, although SDR-based approach has the highest complexity, the semidefinite problem (P3) can be efficiently solved by well-developed CVX solver without the need to determine step size δ [35].

V. AVERAGE BER ANALYSIS

In this section, we analyze the average error probability of the proposed system with optimized IRS beamforming. We also derive the diversity of the system and SNR gain.

Since the link between k-th user and each IRS element in (11) is associated with the same $\boldsymbol{\Phi}_p\mathbf{x}_k$, for $p \in [1, P]$, $k \in [1, K]$, we can simply consider this link as the one between user and the entire IRS and rewrite (11) as

$$\mathbf{y} = \check{\boldsymbol{\Xi}}_{\text{all}}(\mathbf{x}_k)\check{\mathbf{h}}_{\text{all}} + \mathbf{n}, \tag{30}$$

where $\check{\boldsymbol{\Xi}}_{\text{all}}(\mathbf{x}_k) \in \mathbb{C}^{N_r MN \times N_r PK}$ and $\check{\mathbf{h}}_{\text{all}} \in \mathbb{C}^{N_r PK}$ are given by

$$\check{\boldsymbol{\Xi}}_{\text{all}}(\mathbf{x}_k) = \mathbf{I}_{N_r} \otimes \left[\check{\boldsymbol{\Xi}}_1(\mathbf{x}_k), \ldots, \check{\boldsymbol{\Xi}}_K(\mathbf{x}_k)\right],$$
$$\check{\mathbf{h}}_{\text{all}} = \left[(\check{\mathbf{h}}_1^1)^T, \ldots, (\check{\mathbf{h}}_K^1)^T, \ldots, (\check{\mathbf{h}}_1^{N_r})^T, \ldots, (\check{\mathbf{h}}_K^{N_r})^T\right]^T,$$

with

$$\check{\boldsymbol{\Xi}}_k(\mathbf{x}_k) = [\boldsymbol{\Phi}_1 \mathbf{x}_k, \ldots, \boldsymbol{\Phi}_P \mathbf{x}_k],$$
$$\check{\mathbf{h}}_k^i = \left[\sum_{q=1}^{Q} \mathbf{g}_i \odot \boldsymbol{\theta} \odot \boldsymbol{\gamma}_{k1}, \ldots, \sum_{q=1}^{Q} \mathbf{g}_i \odot \boldsymbol{\theta} \odot \boldsymbol{\gamma}_{kP}\right]^T,$$

for $k \in [1, K]$, $i \in [1, N_r]$.

To simplify notation, we let

$$\sum_{q=1}^{Q} \mathbf{g}_i \odot \boldsymbol{\theta} \odot \boldsymbol{\gamma}_{kp} = \sum_{q=1}^{Q} S_q, \tag{31}$$

where random variable $S_q = g_q^i \gamma_{kq}^p e^{j\theta_q}$, for $q \in [1, Q]$. Given the large number of IRS reflecting elements and based on central limit theorem (CLT) [38], we can approximate $\sum_{q=1}^{Q} S_q$ follows complex Gaussian distribution as $\sum_{q=1}^{Q} S_q \sim \mathcal{CN}(0, N_s)$.

Applying Chernoff bound on Q-function and using (18), we can express the average PEP as

$$\begin{aligned}\bar{P}_e\left(\mathbf{x}_{\text{all}} \to \hat{\mathbf{x}}_{\text{all}}\right) &= \mathbb{E}_{\check{\mathbf{h}}_{\text{all}}}\left[Q\left(\sqrt{\frac{\check{\mathbf{h}}_{\text{all}}^H \check{\mathbf{A}}(\boldsymbol{\delta}) \check{\mathbf{h}}_{\text{all}}}{2N_0}}\right)\right] \\ &\le \mathbb{E}_{\check{\mathbf{h}}_{\text{all}}}\left[\exp\left(-\frac{\check{\mathbf{h}}_{\text{all}}^H \check{\mathbf{A}}(\boldsymbol{\delta}) \check{\mathbf{h}}_{\text{all}}}{4N_0}\right)\right],\end{aligned} \tag{32}$$

where $\check{\mathbf{A}}(\boldsymbol{\delta}) = \check{\boldsymbol{\Xi}}_{\text{all}}^H(\boldsymbol{\delta}) \check{\boldsymbol{\Xi}}_{\text{all}}(\boldsymbol{\delta})$. We define SNR $= 1/N_0$. Following steps in [34], we have the following Proposition.

Proposition 2: For the proposed IRS-aided communications system and given SCMA codebook, the theoretical average PEP and average BER are respectively given by

$$\bar{P}_e\left(\mathbf{x}_{\text{all}} \to \hat{\mathbf{x}}_{\text{all}}\right) \le \left(\prod_{r=1}^{R} \lambda_r\right)^{-1} \left(\frac{N_s}{4N_0}\right)^{-R}, \tag{33}$$

and

$$\begin{aligned}\bar{P}_b &\le \sum_{\mathbf{x}_{\text{all}} \in \bar{\mathcal{C}}} \sum_{\boldsymbol{\delta} \in \mathcal{L}} \frac{d(\boldsymbol{\delta}) J}{\alpha MNK \log_2 \mathcal{M}} \bar{P}_e\left(\mathbf{x}_{\text{all}} \to \hat{\mathbf{x}}_{\text{all}}\right) \\ &\le \frac{J}{\alpha MNK \log_2 \mathcal{M}} \sum_{\mathbf{x}_{\text{all}} \in \bar{\mathcal{C}}} \sum_{\boldsymbol{\delta} \in \mathcal{L}} d(\boldsymbol{\delta}) \left(\prod_{r=1}^{R} \lambda_r\right)^{-1} \left(\frac{N_s}{4N_0}\right)^{-R} \\ &\propto \left(\frac{Q^2}{4N_0}\right)^{-R_{\min}},\end{aligned} \tag{34}$$

where $d(\boldsymbol{\delta})$ represents the number of non-zero elements in $\boldsymbol{\delta}$, λ_r's are non-zero eigenvalues of $\check{\mathbf{A}}(\boldsymbol{\delta})$, R is the rank of $\check{\mathbf{A}}(\boldsymbol{\delta})$, and $R_{\min}$ is the lowest rank representing the diversity of the system. The second step in (34) holds for very high SNR values.

Proof: See Appendix B. ■

Remark 2: Since $\check{\mathbf{A}}(\boldsymbol{\delta})$ is a Gram matrix [39], which can have full rank $N_r P$ only when $\boldsymbol{\Phi}_p(\boldsymbol{\delta})$, for $p \in [1, P]$, are linearly independent. This implies that our considered system can have diversity $R_{\min}$ ranging from N_r to $N_r P$ subject to channel conditions and delay-Doppler grid resolutions.

Remark 3: In logarithm scale, the last step in (34) becomes $-20R_{min}\log_{10}(Q) + 10R_{min}\log_{10}(4N_0)$. To achieve the same average BER performance (i.e., same $\bar{P}_b$), increasing IRS element from Q to τQ, $\tau \in \mathbb{Z}$ leads to an SNR gain of $20\log_{10}(\tau)$.

VI. JOINT DETECTION AND SCMA DECODING

In this section, we present a modified iterative Gaussian approximation-based MP detection and SCMA decoding algorithm to directly estimate the bit log-likelihood ratios (LLRs). We start from the Gaussian approximation-based MPA in [8], where the interference terms (see (35)) are approximated as complex Gaussian random variable to avoid the exhaustive search used in the exact MPA in [40]. Other variants of the Gaussian approximation based MPA for SCMA systems can be found in [15], [41], and [42] and references therein. Our modifications on [8] are given below.

- We compute the likelihood functions of SCMA codewords instead of those of QAM symbols when passing messages from variable nodes (VN) to function nodes (FN). In particular, given that SCMA codewords have different real and imaginary values, the mean and variance of the interference term are computed separately for the real and imaginary parts.
- In the termination stage, we compute bit LLRs based on likelihoods of SCMA codewords.

In addition, different from [15] and [42], we compute the likelihoods from VN to FN rather than approximating them as Gaussian, without incurring significant computational complexity. Further, our method directly calculates bit LLRs estimate at each VN, without extra SCMA demapping used in [15].

Next, we present our modified joint Gaussian approximation based MP detection and SCMA decoding method.

Since $\mathbf{H}_{\text{all}}$ in (9) is a sparse matrix, $\mathbf{y}$ has length $N_r MN$ and $\mathbf{x}_{\text{all}}$ is a vector consisting of a total number of $\frac{MNK}{J}$ transmit SCMA codewords of all users. We consider a factor graph that has $N_r MN$ FNs and $\frac{MNK}{J}$ VNs. Let $\mathcal{I}_d$ and $\mathcal{J}_a$ denote the set of indexes of non-zero elements in the d-th row and a-th column of the factor graph, respectively, for $d = 1, \ldots, N_r MN$, $a = 1, \ldots, \frac{MNK}{J}$. In this factor graph, each FN y_d, is connected to the set of VNs $\{x_a, a \in \mathcal{I}_d\}$. Each VN x_a, is connected the set of FNs $\{y_d, d \in \mathcal{J}_a\}$. The modified message-passing process is given below.

Step 1: **Initialization**. For a-th VN with associated SCMA codebook, $\mathcal{C}_a$, we assume each codeword has equal prior probability, i.e., $P_{x_a}(\mathbf{c}_{am}) = \frac{1}{\mathcal{M}}$, where $\mathbf{c}_{am} \in \mathcal{C}_a$ represents m-th codeword in codebook $\mathcal{C}_a$ for $m = 1, \ldots, \mathcal{M}$. Since we

consider log-MPA detection, all messages are computed in logarithm scale. The initial message passed from VN to FN, $\eta^{\text{ini}}_{x_a\to y_d}(\mathbf{c}_{am})$ are equally likely.

Step 2: **Message passing from FNs to VNs**. Similar to [8], for each FN y_d, a VN x_a is isolated from the interference terms, which can be approximated as Gaussian noise. Consider

$$y_d = \mathbf{h}_{da}x_a(\mathbf{c}_{am}) + \underbrace{\sum_{e\in\mathcal{I}_d, e\neq a}\mathbf{h}_{de}x_e(\mathbf{c}_{em}) + n_d}_{\xi_{de}},$$
$$\text{for } \mathbf{c}_{am}\in\mathcal{C}_a,\ \text{for } \mathbf{c}_{em}\in\mathcal{C}_e, \tag{35}$$

where $\mathbf{h}_{da} = \mathbf{H}_{\text{all}}[d, (a-1)\mathcal{M}+1 : a\mathcal{M}] \in \mathbb{C}^{1\times\mathcal{M}}$ with $\mathbf{H}_{\text{all}}$ given in (9), and $x_a(\mathbf{c}_{am})$ represent complex channel gain and codeword associated with VN x_a, respectively. Similarly, $\mathbf{h}_{de}$ and $x_e(\mathbf{c}_{em})$ are channel gain and codeword associated with VN x_e, respectively. Also, $n_d \sim \mathcal{CN}(0, N_0)$ is AWGN noise.

Then we approximate both $\Re(\xi_{de})$ and $\Im(\xi_{de})$ follow Gaussian distribution with different mean and variance as

$$\mathbb{E}[\xi_{de}] = \sum_{e\in\mathcal{I}_d, e\neq c}\sum_{m=1}^{\mathcal{M}}\mathbf{h}_{de}x_e(\mathbf{c}_{em})\exp\left(\eta_{x_e\to y_d}(\mathbf{c}_{em})\right),$$
$$\mathbb{E}[\Re(\xi_{de})] = \Re\left(\mathbb{E}[\xi_{de}]\right),\ \mathbb{E}[\Im(\xi_{de})] = \Im\left(\mathbb{E}[\xi_{de}]\right) \tag{36}$$

and

$$\begin{aligned}&\text{Var}[\Re(\xi_{de})]\\ &= \frac{N_0}{2} + \sum_{e\in\mathcal{I}_d, e\neq c}\\ &\left[\left(\sum_{m=1}^{\mathcal{M}}\Re\left(\mathbf{h}_{de}x_e(\mathbf{c}_{em})\right)^2\exp\left(\eta_{x_e\to y_d}(\mathbf{c}_{em})\right)\right)\right.\\ &\left.-\left(\sum_{m=1}^{\mathcal{M}}\Re\left(\mathbf{h}_{de}x_e(\mathbf{c}_{em})\right)\exp\left(\eta_{x_e\to y_d}(\mathbf{c}_{em})\right)\right)^2\right],\\ &\text{Var}[\Im(\xi_{de})] = \frac{N_0}{2} + \sum_{e\in\mathcal{I}_d, e\neq c}\\ &\left[\left(\sum_{m=1}^{\mathcal{M}}\Im\left(\mathbf{h}_{de}x_e(\mathbf{c}_{em})\right)^2\exp\left(\eta_{x_e\to y_d}(\mathbf{c}_{em})\right)\right)\right.\\ &\left.-\left(\sum_{m=1}^{\mathcal{M}}\Im\left(\mathbf{h}_{de}x_e(\mathbf{c}_{em})\right)\exp\left(\eta_{x_e\to y_d}(\mathbf{c}_{em})\right)\right)^2\right],\\ &\text{for } \mathbf{c}_{em}\in\mathcal{C}_e.\end{aligned} \tag{37}$$

Then the message passed from FN y_d to VN x_a is given by

$$\eta_{y_d\to x_a}(\mathbf{c}_{am}) = -\frac{\left(\Re\left(y_d - \mathbf{h}_{da}x_a(\mathbf{c}_{am})\right) - \mathbb{E}[\Re(\xi_{de})]\right)^2}{2\text{Var}[\Re(\xi_{de})]} - \frac{\left(\Im\left(y_d - \mathbf{h}_{da}x_a(\mathbf{c}_{am})\right) - \mathbb{E}[\Im(\xi_{de})]\right)^2}{2\text{Var}[\Im(\xi_{de})]},$$
$$\text{for } \mathbf{c}_{am}\in\mathcal{C}_a. \tag{38}$$

Step 3: **Message passing from VNs to FNs**. With normalization, the message passing from each VN is given by

$$\eta_{x_a\to y_d}(\mathbf{c}_{am}) = \log\left(\frac{1}{\mathcal{M}}\right) + \sum_{f\in\mathcal{J}_a, f\neq d}\eta_{y_f\to x_a}(\mathbf{c}_{am}) - \log\left[\sum_{m=1}^{\mathcal{M}}\exp\left(\sum_{f\in\mathcal{J}_a, f\neq d}\eta_{y_f\to x_a}(\mathbf{c}_{am})\right)\right],$$
$$\text{for } \mathbf{c}_{am}\in\mathcal{C}_a. \tag{39}$$

In addition, we apply damping factor on message updating to control convergence speed [8]. The message passing from VNs to FNs at n-th iteration with damping factor, ζ, can be expressed as

$$\eta_{x_a\to y_d}(\mathbf{c}_{am})^{(n)} = \log\Big[\zeta\exp\left(\eta_{x_a\to y_d}(\mathbf{c}_{am})^{(n)}\right) + (1-\zeta)\exp\left(\eta_{x_a\to y_d}(\mathbf{c}_{am})^{(n-1)}\right)\Big],$$
$$\text{for } \mathbf{c}_{am}\in\mathcal{C}_a. \tag{40}$$

Step 4: **Termination and bit decision**. After repeating step 2 and step 3 for several times, the final belief at each VN, x_a, is given by

$$P_{x_a}(\mathbf{c}_{am}) = \log\left(\frac{1}{\mathcal{M}}\right) + \sum_{d\in\mathcal{J}_a}\eta_{y_d\to x_a}(\mathbf{c}_{am}),$$
$$\text{for } \mathbf{c}_{am}\in\mathcal{C}_a. \tag{41}$$

Then we can calculate bit LLR for each VN, x_a, as

$$\text{LLR}^t_a = \log\left(\frac{P(b_t=0)}{P(b_t=1)}\right) \tag{42}$$

for $t = 1,\ldots,\log_2(\mathcal{M})$, where $P(b_t = 0)$ and $P(b_t = 1)$ represents the probability of t-th bit of transmit symbol associated with VN x_a is equal to 0 and 1, respectively.

Remark 4: (Complexity analysis of our modified MP detection and decoding algorithm) During each iteration, we compute mean and variance at each FN in step 2 (i.e., message passing from FNs to VNs), which has computational complexity of $\mathcal{O}(N_rMN|\mathcal{I}_d|\mathcal{M})$. The complexity in step 3 (i.e., message passing from VNs to FNs) is $\mathcal{O}\left(\frac{MNK}{J}|\mathcal{J}_a|\mathcal{M}\right)$. Thus the overall complexity can be described as $\mathcal{O}\left(I_gMN\mathcal{M}\left(N_r|\mathcal{I}_d| + \frac{K}{J}|\mathcal{J}_a|\right)\right)$, where I_g represents the maximum iteration number of Gaussian log-MPA.

VII. SIMULATION RESULTS

In this section, we simulate the error performance of IRS-aided uplink OTFS-SCMA system. We consider Extended Vehicular A (EVA) channel model for users-IRS channel links with simulation parameters given in Table 2 [43]. We consider $K = 6$ users simultaneously communicate to BS via $J = 4$ REs. The SCMA codebook is generated according to [44]. We also set $N_r = 4$ and $\mathcal{M} = 4$.

In Table. 3, we first investigate the impact of damping factor, ζ, on the performance of the detection and decoding algorithm in Section VII, where I_{mp} is average number of

TABLE 2. Simulation parameters.

Parameter	Value
Carrier frequency (f_c)	4 GHz
Number of subcarriers (M)	64
Number of time slots (N)	16
Subcarrier spacing (Δf)	15 kHz
Modulation order ($\mathcal{M}$)	4
User speed	500 Kmph

TABLE 3. BER performance comparison among different values of damping factor ζ in modified MP detection and decoding algorithm, where SNR = −40dB, $J = 4$, $K = 6$, $Q = 32$ and IRS phase shifts are designed based on GA-algorithm.

Damping factor (ζ)	BER	I_{mp}
0.2	0.0494	5
0.4	0.0494	5
0.6	0.0517	5
0.8	0.0821	5

iterations required to reach convergence. We set SNR $=$ -40dB and IRS element number $Q = 32$ with GA-algorithm IRS phase shifts design. It can be seen that BER performance remains unchanged for small values of ζ and deteriorates for large values. Also, given the detection and decoding algorithm reaches convergence within an average of 5 iterations, damping factor will not have a significant impact on detection convergence speed for the proposed IRS-aided OTFS-SCMA system. Thus, we choose $\zeta = 0.2$ in the following simulations.

In Fig. 5, we demonstrate BER versus SNR among different MP detection algorithms. It can be seen that our modified iterative MP detection and decoding algorithm achieves error performance similar to that of MPA in [8] at low SNR values, which demonstrates the effectiveness of our modifications to the proposed IRS-aided OTFS-SCMA system. At high SNRs, our modified MPA achieves better error performance. This can be attributed to the increasing discrepancy in variances between real and imaginary parts of interference term in (35) as SNR values increase. The equal variance assumption in [8] will lead to a degradation of error performance.

In Fig. 6, we demonstrate BER versus SNR for different IRS beamforming design algorithms. It can be seen that both SDR-based approach and GA-based approach can achieve similar error performance, whereas GA-based approach has lower computational complexity compared to SDR-based approach as discussed in Section IV. Although matching-based IRS phase shifts design in [31] has lower complexity compared to GA-based algorithm, it has much higher BER values. For example, with $Q = 64$ and at BER of 10^{-3}, GA-based approach outperforms matching-based approach by about 2.5dB. We can also observe that higher values of IRS element number, Q, will lead to lower BER values. At BER of 10^{-3}, increasing Q from 32 to 64 implies an SNR gain of around 5.5dB, which closely matches with

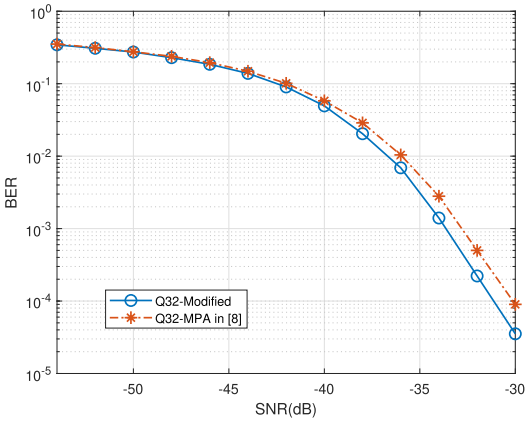

FIGURE 5. BER versus SNR for the proposed IRS-aided uplink OTFS-SCMA system among different MP detection algorithms, where $M = 64$, $N = 16$, $\mathcal{M} = 4$, $J = 4$, $K = 6$ and IRS phase shifts are designed based on GA-algorithm.

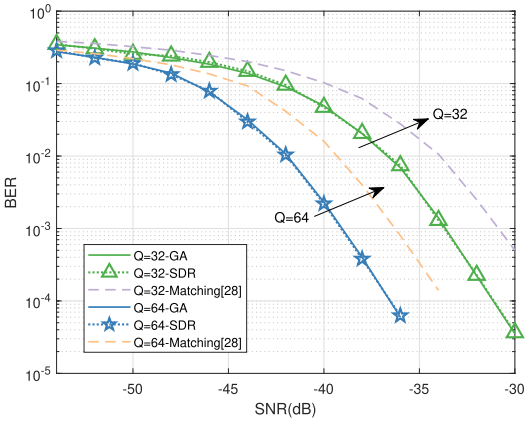

FIGURE 6. BER versus SNR among different IRS design algorithms, where $M = 64$, $N = 16$, $\mathcal{M} = 4$, $J = 4$, $K = 6$.

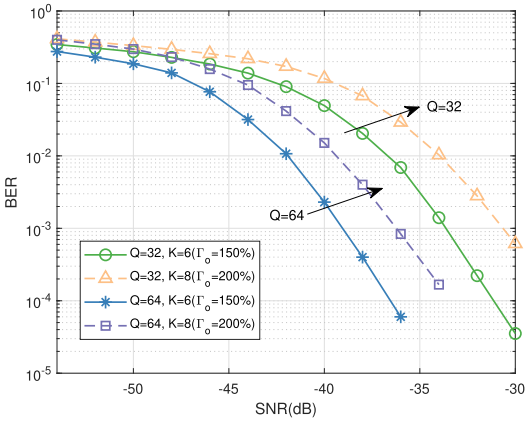

FIGURE 7. BER versus SNR among different overloading factors with GA-algorithm IRS design, where $M = 64$, $N = 16$, $\mathcal{M} = 4$, $J = 4$.

theoretical result $20\log_{10}(\tau) = 6.02$dB, as explained in Section V.

In Fig. 7, we demonstrate BER performance of the proposed IRS-aided uplink OTFS-SCMA communications system with different overloading factors, Γ_o, which is equal to $\frac{K}{J}$ as discussed in subsection II-A. We consider $J = 4$ and compare BER performance among two scenarios,

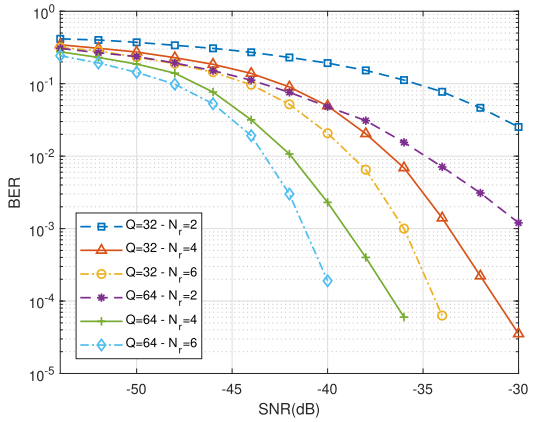

FIGURE 8. BER versus SNR among different number of receive antennas at BS with GA-algorithm IRS design, where $M = 64$, $N = 16$, $\mathcal{M} = 4$, $J = 4$, $K = 6$.

$\Gamma_o = 150\%$ (i.e., $K = 6$) and $\Gamma_o = 200\%$ (i.e. $K = 8$). The IRS phase shifts are designed based on the GA algorithm. From Fig. 7, it can be seen that higher overloading factors will lead to BER performance deterioration due to the increased number of users superimposed on each RE.

In Fig. 8, we demonstrate BER versus SNR for different values of receive antenna at BS, i.e., different N_r. We adopt the GA-algorithm for IRS phase shift design. It can be seen that higher values of N_r lead to better BER performance and higher diversity gain, which validates the analysis in (34).

VIII. CONCLUSION

In this paper, we considered an IRS-assisted uplink OTFS-SCMA communications system. We derived an upper bound on WER, based on which we designed SDR-based and GA-based algorithms to minimize WER. We investigated average BER performance, demonstrating that our considered IRS-aided OTFS-SCMA system can achieve diversity gains ranging from N_r to N_rP. We also showed an SNR gain achieved by increasing the number of IRS elements. To recover information bits of all users, we presented a modified joint iterative MP detection and decoding algorithm. Numerical results demonstrated that our proposed GA-based optimization algorithm can achieve significantly better error performance compared to the known approach without incurring high computational complexity. Our modified detection and decoding algorithm demonstrated its effectiveness for the proposed IRS-aided OTFS-SCMA system. An interesting direction for future research involves exploring the design and performance analysis of IRS-aided communications networks under practical constraints, such as imperfect CSI acquisition and/or discrete phase-shift control of IRS reflections.

APPENDIX A
PROOF OF PROPOSITION 1

Substituting (16) and (17) into (14) yields the first upper bound, and using $Q(x) \leq \exp\left(-\frac{x^2}{2}\right)$ yields the second upper bound.

Assuming inequality $0 < x < \frac{1}{\eta} - \epsilon$ such that $\exp(-x) < -\eta x + 1$ holds, where η and ϵ are small positive real number [25], we have

$$P_e(\mathbf{h}_{\text{all}}) \leq \frac{1}{\alpha} \sum_{\mathbf{x}_{\text{all}}} \sum_{\delta \in \mathcal{L}} \left(\frac{-\eta \mathbf{h}_{\text{all}}^H \mathbf{A}(\delta) \mathbf{h}_{\text{all}}}{4N_0} + 1 \right). \tag{43}$$

The final step in (18) can then be obtained.

APPENDIX B
PROOF OF PROPOSITION 2

Since we approximate $\check{\mathbf{h}}_{\text{all}}$ follows complex Gaussian distribution, (33) can be easily obtained based on the theory of probability distributions [34].

From (31), we can express N_s as

$$\begin{aligned} N_s &= \sum_{q=1}^{Q} \text{Var}\left[S_q\right] + \sum_{q,q'=1, q \neq q'}^{Q} \text{Cov}\left[S_q, S_{q'}\right] \\ &= \sum_{q=1}^{Q} \text{Var}\left[S_q\right] + \sum_{q,q'=1, q \neq q'}^{Q} \rho \sqrt{\text{Var}\left[S_q\right] \text{Var}\left[S_{q'}\right]}, \end{aligned} \tag{44}$$

where ρ represents correlation coefficient between S_q and $S_{q'}$, for $q, q' = 1, \ldots, Q$. We assume that S_q and $S_{q'}$ have similar distribution and we can obtain $\text{Cov}\left[S_q, S_{q'}\right] \approx \rho \text{Var}\left[S_q\right]$ and approximate

$$N_s \propto Q + Q(Q-1) = Q^2. \tag{45}$$

The first step in (34) is obtained based on the definitions of PEP and BER [45]. At very high SNRs, it can be observed that $\check{\mathbf{A}}(\delta)$ with the lowest rank will lead to dominant values in (34). Then the results in (34) can be obtained based on (33) and (45).

REFERENCES

[1] 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, 2nd Quart., 2017.
[2] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.
[3] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2013, pp. 332–336.
[4] M. Rebhi, K. Hassan, K. Raoof, and P. Chargé, "Sparse code multiple access: Potentials and challenges," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1205–1238, 2021.
[5] S. Zhang, X. Xu, L. Lu, Y. Wu, G. He, and Y. Chen, "Sparse code multiple access: An energy efficient uplink approach for 5G wireless systems," in *Proc. IEEE Global Commun. Conf.*, Austin, TX, USA, Dec. 2014, pp. 4782–4787.
[6] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.
[7] Y. Hong, T. Thaj, and E. Viterbo, *Delay-Doppler Communications: Principles and Applications*. Amsterdam, The Netherlands: Elsevier, 2022.
[8] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6501–6515, Oct. 2018.
[9] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix OTFS," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 957–961, Jan. 2019.

[10] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in Delay–Doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.
[11] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Orthogonal time frequency space (OTFS) modulation based radar system," in *Proc. IEEE Radar Conf. (RadarConf)*, Boston, MA, USA, Apr. 2019, pp. 1–6.
[12] B. C. Pandey, S. K. Mohammed, P. Raviteja, Y. Hong, and E. Viterbo, "Low complexity precoding and detection in multi-user massive MIMO OTFS downlink," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4389–4405, May 2021.
[13] K. R. Murali and A. Chockalingam, "On OTFS modulation for high-Doppler fading channels," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2018, pp. 1–10.
[14] K. Deka, A. Thomas, and S. Sharma, "OTFS-SCMA: A code-domain NOMA approach for orthogonal time frequency space modulation," *IEEE Trans. Commun.*, vol. 69, no. 8, pp. 5043–5058, Aug. 2021.
[15] Y. Ge, Q. Deng, D. González G., Y. L. Guan, and Z. Ding, "OTFS signaling for SCMA with coordinated multi-point vehicle communications," *IEEE Trans. Veh. Technol.*, vol. 72, no. 7, pp. 9044–9057, Jul. 2023.
[16] H. Wen, W. Yuan, Z. Liu, and S. Li, "OTFS-SCMA: A downlink NOMA scheme for massive connectivity in high mobility channels," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 5770–5784, Sep. 2023.
[17] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2283–2314, 4th Quart., 2020.
[18] 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.
[19] J. Li and Y. Hong, "Design of an intelligent reflecting surface aided mmWave massive MIMO using X-precoding," *IEEE Access*, vol. 10, pp. 69428–69440, 2022.
[20] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, "Intelligent reflecting surface enhanced wireless networks: Two-timescale beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 2–17, Jan. 2021.
[21] I. Al-Nahhal, O. A. Dobre, and E. Basar, "Reconfigurable intelligent surface-assisted uplink sparse code multiple access," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 2058–2062, Jun. 2021.
[22] S. Sharma, K. Deka, Y. Hong, and D. Dixit, "Intelligent reflecting surface-assisted uplink SCMA system," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2728–2732, Aug. 2021.
[23] S. Sharma, K. Deka, and V. Bhatia, "Intelligent reflecting surface-aided downlink SCMA," *IEEE Syst. J.*, vol. 17, no. 2, pp. 3204–3211, Jun. 2023.
[24] I. Al-Nahhal, O. A. Dobre, E. Basar, T. M. N. Ngatched, and S. Ikki, "Reconfigurable intelligent surface optimization for uplink sparse code multiple access," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 133–137, Jan. 2022.
[25] J. Li and Y. Hong, "Design of intelligent reflecting surface phase shifts for uplink SCMA," *IEEE Commun. Lett.*, vol. 27, no. 5, pp. 1437–1441, May 2023.
[26] S. Chaturvedi, V. A. Bohara, Z. Liu, and A. Srivastava, "Sum-rate maximization of IRS-aided SCMA system," *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 10462–10472, Aug. 2023.
[27] A. Thomas, K. Deka, S. Sharma, and N. Rajamohan, "IRS-assisted OTFS system: Design and analysis," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3345–3358, Mar. 2023.
[28] G. Harshavardhan, V. S. Bhat, and A. Chockalingam, "RIS-aided OTFS modulation in high-Doppler channels," in *Proc. IEEE 33rd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Kyoto, Japan, Sep. 2022, pp. 409–415.
[29] V. S. Bhat, G. Harshavardhan, and A. Chockalingam, "Input-output relation and performance of RIS-aided OTFS with fractional delay-Doppler," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 337–341, Jan. 2023.
[30] C. Xu, L. Xiang, J. An, C. Dong, S. Sugiura, R. G. Maunder, L.-L. Yang, and L. Hanzo, "OTFS-aided RIS-assisted SAGIN systems outperform their OFDM counterparts in doubly selective high-Doppler scenarios," *IEEE Internet Things J.*, vol. 10, no. 1, pp. 682–703, Jan. 2023.
[31] A. S. Bora, K. T. Phan, and Y. Hong, "IRS-assisted high mobility communications using OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 12, no. 2, pp. 376–380, Feb. 2023.
[32] Z. Li, W. Yuan, B. Li, J. Wu, C. You, and F. Meng, "Reconfigurable intelligent surface aided OTFS: Transmission scheme and channel estimation," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 19518–19532, Nov. 2023.
[33] A. Thomas, K. Deka, P. Raviteja, and S. Sharma, "Convolutional sparse coding based channel estimation for OTFS-SCMA in uplink," *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5241–5257, Aug. 2022.
[34] E. Biglieri, P. Raviteja, and Y. Hong, "Error performance of orthogonal time frequency space (OTFS) modulation," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Shanghai, China, May 2019, pp. 1–6.
[35] M. Grant and S. Boyd. (Aug. 2012). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.0*. [Online]. Available: http://cvxr.com/cvx
[36] Z.-Q. Luo, W.-K. Ma, A. M. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
[37] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1268–1279, Mar. 2008.
[38] D. S. Lemons, *An Introduction to Stochastic Processes in Physics*, 1st ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2002.
[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[40] S. Chaturvedi, Z. Liu, V. A. Bohara, A. Srivastava, and P. Xiao, "A tutorial on decoding techniques of sparse code multiple access," *IEEE Access*, vol. 10, pp. 58503–58524, 2022.
[41] J. Dai, K. Niu, and J. Lin, "Iterative Gaussian-approximated message passing receiver for MIMO-SCMA system," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 753–765, Jun. 2019.
[42] Q. Zou and H. Yang, "A concise tutorial on approximate message passing," Mar. 2022, *arXiv:2201.07487*.
[43] *Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) Radio Transmission and Reception, Version 8.6.0*, document TS 36.104, 3GPP, Jul. 2009.
[44] S. Chaturvedi, Z. Liu, V. A. Bohara, A. Srivastava, and P. Xiao, "A tutorial to sparse code multiple access," May 2021, *arXiv:2105.06860*.
[45] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Effective diversity of OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 249–253, Feb. 2020.

JIARUI LI (Graduate Student Member, IEEE) received the bachelor's degree (Hons.) in electronic and communication systems engineering from The Australian National University (ANU). She is currently pursuing the Ph.D. degree with Monash University, Melbourne, Australia. Her research interests include wireless communications, coding theory, and intelligent reflecting surface.

YI HONG (Senior Member, IEEE) received the Ph.D. degree in electrical engineering and telecommunications from the University of New South Wales (UNSW), Sydney.

She is currently an Associate Professor with the Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Australia. Her research interests include communication theory, and coding and information theory with applications to telecommunication engineering. She was a technical program committee member of many IEEE leading conferences. She received the NICTA-ACoRN Earlier Career Researcher Award from the Australian Communication Theory Workshop, Adelaide, Australia, in 2007. She served on the Australian Research Council College of Experts from 2018 to 2020. She was the General Co-Chair of the IEEE Information Theory Workshop 2014, Hobart; the Technical Program Committee Chair of the Australian Communications Theory Workshop 2011, Melbourne; and the Publicity Chair of the IEEE Information Theory Workshop 2009, Sicily. She was an Associate Editor of IEEE Wireless Communication Letters and *Transactions on Emerging Telecommunications Technologies* (ETT).