A Novel Non-Coherent SCMA with Massive MIMO

Qu Luo, Haifeng Wen, Gaojie Chen, *Senior Member, IEEE,* Zilong Liu, *Senior Member, IEEE,*
Pei Xiao, *Senior Member, IEEE,* Yi Ma, *Senior Member, IEEE,* Amine Maaref, *Senior Member, IEEE.*

***Abstract*—The synergistic amalgamation of sparse code multiple access (SCMA) and multiple-input multiple-output (MIMO) technologies can be exploited for improving spectral efficiency and providing enhanced wireless services to massive users. In this case, however, channel estimation is a burning issue with the increasing number of users and/or antennas. To tackle this problem, we propose a novel non-coherent transmission scheme for SCMA, referred to as NC-SCMA. In the proposed NC-SCMA, each user first maps its binary data to sparse codewords, and then perform differential modulation on the non-zero dimensions. Upon receiving all users' signals, we leverage the channel hardening effect to carry out differential demodulation and multi-user detection without any instantaneous channel state information. In addition, the design of the sparse codebooks in the NC-SCMA system is investigated with the aid of the pair-wise probability. Numerical results demonstrate the superiority of the proposed technique over the benchmark scheme in terms of bit error rate performance.**

***Index Terms*—SCMA, massive MIMO, differential modulation, non-coherent transmission, message passing algorithm.**

I. Introduction

Sparse code multiple access (SCMA) is a promising non-orthogonal multiple access (NOMA) scheme for massive machine-type communications (mMTC) in 5G-and-beyond communication systems [1], [2]. As a generalization of low density signature (LDS) using sparse sequence spreading of constellation symbols [3], every SCMA user's instantaneous input message bits are directly mapped to a multi-dimensional (MD) sparse codeword drawn from a carefully designed sparse codebook [4], [5]. At the same time, massive multiple-input multiple-output (MIMO) technology is another emerging paradigm for significantly boosted spectral efficiency [6]. The combination of MIMO and NOMA, referred to as MIMO-NOMA, is thus a promising wireless technique for tremendous performance improvement [7], [8].

In the existing works on MIMO-NOMA, it is often assumed that the instantaneous channel state information (CSI) for all users is available at the base station (BS) [6]–[8]. As the data packets in mMTC are typically short, however, the traditional "estimation-and-then-detection" method may result in excessive pilot overhead and unaffordable complexity at the receiver. The situation gets severe in high mobility communication environments as the system does not afford tedious channel estimation due to the rapidly varying wireless channels. In order to avoid channel estimation, whilst at the same time, attain enhanced spectral efficiency, non-coherent communication was proposed in MIMO-NOMA systems. In [9], the authors proposed a joint constellation design for two users by enumerating all possible orderings that satisfy design constraints. A new constellation domain-based approach was proposed in [10], [11] by employing energy detector with the unipolar pulse-amplitude modulation constellations. Recently, an LDS based non-coherent transmission scheme was proposed in [12] which can support overloaded users whilst achieving a comparable BER performance to that of point-to-point systems. However, their proposed scheme only performs efficiently with the binary phase-shift keying (BPSK) modulation and is limited to LDS system whose BER performance is generally inferior to SCMA.

In this letter, a novel non-coherent SCMA (NC-SCMA) architecture with optimized sparse codebooks is introduced. For non-coherent transmission, an efficient differential modulation scheme after the SCMA encoding is first proposed. Sparse codebooks are optimized in order to cater for certain unique constraint imposed by the phase rotation in differential modulation. It is noted that traditional SCMA codebooks are inapplicable due to such a new constraint. At the BS, upon receiving the superimposed signals from all users, differential demodulation is first carried out. Multiuser decoding is then performed based on messaging passing algorithm (MPA) with only statistical channel information by utilizing the channel hardening [13], with which the fading channels behave like Gaussian ones. Numerical results show that the proposed NC-SCMA with our newly designed codebooks can achieve about 3 dB and 6 dB gains over the non-coherent scheme in [12] at BER $= 10^{-4}$ for the modulation orders of $M = 2$ and $M = 4$, respectively.

Notations: The n-dimensional complex and binary vector spaces are denoted as $\mathbb{C}^n$ and $\mathbb{B}^n$, respectively. Similarly, $\mathbb{C}^{k\times n}$ and $\mathbb{B}^{k\times n}$ denote the $(k \times n)$-dimensional complex and binary matrix spaces, respectively. $\mathbf{I}_n$ denotes an $n \times n$-dimensional identity matrix. $\mathrm{tr}(\mathbf{X})$ denotes the trace of a square matrix $\mathbf{X}$. $\mathrm{diag}(\mathbf{x})$ gives a diagonal matrix with the diagonal vector of $\mathbf{x}$. $(\cdot)^{\mathcal{T}}$, $(\cdot)^{\dagger}$ and $(\cdot)^{\mathcal{H}}$ denote the transpose, the conjugate and the Hermitian transpose operation, respectively. $\|\mathbf{x}\|_2$ and $|x|$ return the Euclidean norm of vector $\mathbf{x}$ and the absolute value of x, respectively.

II. Proposed NC-SCMA

We assume an uplink scenario where J users of each with single antenna communicate over K resources to a BS equipped with N receive antennas. The block diagram of proposed NC-SCMA scheme with massive MIMO is shown in

Qu Luo, Gaojie Chen, Pei Xiao and Yi Ma are with 5G & 6G Innovation Centre, University of Surrey, UK, email:{q.u.luo, gaojie.chen, p.xiao, y.ma}@surrey.ac.uk. Haifeng Wen is with the University of Electronic Science and Technology of China, email: wenhaifeng@std.uestc.edu.cn. Zilong Liu is with the School of Computer Science and Electronics Engineering, University of Essex, UK. email: zilong.liu@essex.ac.uk. Amine Maaref with Canada Research Center, Huawei Technologies Company Ltd., Ottawa, Canada, email:amine.maaref@huawei.com, This work was supported in part by the UK Engineering and Physical Sciences Research Council under Grant EP/P03456X/1.

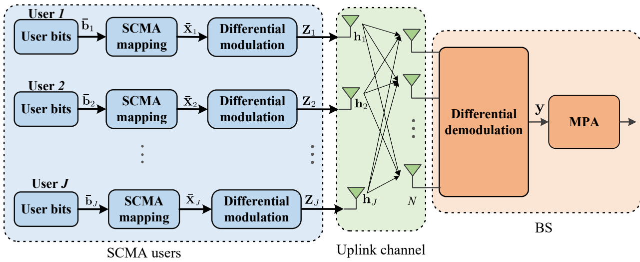


Fig. 1: System model of the proposed NC-SCMA.

Fig. 1. Let $\bar{\mathbf{b}}_j = \left[\mathbf{b}_{j,1}^{\mathcal{T}}, \mathbf{b}_{j,2}^{\mathcal{T}}, \ldots, \mathbf{b}_{j,T}^{\mathcal{T}}\right]^{\mathcal{T}} \in \mathbb{B}^{T\log_2(M)}$ be the message packet to be transmitted, where $\mathbf{b}_{j,t}^{\mathcal{T}} \in \mathbb{B}^{\log_2(M)}, 1 \leq t \leq T$, and M is the modulation order and T denotes the number of symbols to be transmitted for each packet. Then, the message packet $\bar{\mathbf{b}}_j$ will successively go through the operations of SCMA encoding and differential encoding at transmitter side.

A. SCMA encoding

The SCMA encoder successively maps $\log_2 M$ binary bits, i.e., $\mathbf{b}_{j,t}$, into a length-K codeword drawn from pre-designed codebook $\mathcal{X}_j \in \mathbb{C}^K$ with size of M. The mapping process is defined as [14]

$$f_j : \mathbb{B}^{\log_2 M} \rightarrow \boldsymbol{\mathcal{X}}_j \in \mathbb{C}^K, \text{i.e.,} \mathbf{x}_{j,t} = f_j(\mathbf{b}_{j,t}), \tag{1}$$

where $\boldsymbol{\mathcal{X}}_j = \{\mathbf{x}_j^{(1)}, \mathbf{x}_j^{(2)}, \ldots, \mathbf{x}_j^{(M)}\}$ is the codebook set for the jth user with cardinality of M. All the K-dimensional complex codewords of each SCMA codebook are sparse vectors with V non-zero elements and $V < K$. The sparse structure of the J SCMA codebooks can be represented by the indicator (sparse) matrix $\mathbf{F}_{K\times J} = [\mathbf{f}_1, \ldots, \mathbf{f}_J] \in \mathbb{B}^{K\times J}$. The element in $\mathbf{F}_{K\times J}$ is defined as $f_{k,j}$, and the variable node j is connected to resource node k if and only if $f_{k,j} = 1$. Furthermore, let $\boldsymbol{\mathcal{I}}_r(k) = \{j \,|f_{k,j} = 1\}$ and $\boldsymbol{\mathcal{I}}_u(j) = \{k \,|f_{k,j} = 1\}$ denote the set of user indices sharing resource node k and the set of resource indices occupied by user j, respectively, and denote $d_v = |\boldsymbol{\mathcal{I}}_u(j)|$ and $d_f = |\boldsymbol{\mathcal{I}}_r(k)|$. In this paper, the following factor graph with $d_v = 2$ and $d_f = 3$ is employed [14]:

$$\mathbf{F_{4\times 6}} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}. \tag{2}$$

B. Differential modulation

After SCMA encoding, the binary information packet $\bar{\mathbf{b}}_j$ is modulated to a sequence of complex codewords, denoted as $\overline{\mathbf{X}}_j = \left[\mathbf{x}_j^{\mathcal{T}}(1), \mathbf{x}_j^{\mathcal{T}}(2), \ldots, \mathbf{x}_j^{\mathcal{T}}(K)\right]^{\mathcal{T}}$, where $\mathbf{x}_j(k) = [x_{j,k,1}, x_{j,k,2}, x_{j,k,T}] \in \mathbb{C}^T$. Specifically, each user separately performs differential modulation on the non-zero dimensions of $\overline{\mathbf{X}}_j$, i.e., $\mathbf{x}_j(k)$, $k \in \boldsymbol{\mathcal{I}}_u(j)$. Hence, the codewords at the kth entry must have the same energy, i.e., $|x_{j,k}^{(m)}|^2 = E_{j,k}$, $1 \leq m \leq M$, $k \in \boldsymbol{\mathcal{I}}_u(j)$, where $x_{j,k}^{(m)}$ is the jth user's mth codeword at the kth entry and $E_{j,k} > 0$ denotes the corresponding codeword energy. The design of the codebook that satisfies such constraint will be discussed in Section III-B.

Let $\overline{\mathbf{Z}}_j = \left[\overline{\mathbf{z}}_j^{\mathcal{T}}(1), \overline{\mathbf{z}}_j^{\mathcal{T}}(2), \ldots, \overline{\mathbf{z}}_j^{\mathcal{T}}(K)\right]^{\mathcal{T}} \in \mathbb{C}^{K\times(T+1)}$ denote the differentially modulated symbols for the jth user, where $\overline{\mathbf{z}}_j(k) = [\overline{z}_{j,k,0}, \overline{z}_{j,k,1}, \ldots, \overline{z}_{j,k,T}]$ with $\overline{z}_{j,k,0} = 1$ being the reference symbol, and $\overline{\mathbf{z}}_j(k)$ is obtained as

$$\overline{z}_{j,k,t} = \frac{1}{\sqrt{E_{j,k}}} \overline{z}_{j,k,t-1} x_{j,k,t}, 1 \leq t \leq T, \tag{3}$$

The power scaling factor $1/\sqrt{E_{j,k}}$ in (3) is introduced such that $|z_{j,k,t}|^2 = 1$, which is the central to differential modulation. Finally, the transmitted sequences in frequency domain are obtained as

$$\mathbf{Z}_j = \left[\mathbf{z}_j^{\mathcal{T}}(1), \mathbf{z}_j^{\mathcal{T}}(2), \ldots, \mathbf{z}_j^{\mathcal{T}}(K)\right]^{\mathcal{T}}, \tag{4}$$

where $\mathbf{z}_j(k) = \sqrt{E_{j,k}} \overline{\mathbf{z}}_j(k)$ and the energy factor $\sqrt{E_{j,k}}$ is introduced to ensure the energy remains unchanged after differential modulation.

Denote $\mathbf{r}_{k,t} = [r_{k,t,1}, r_{k,t,2}, \ldots, r_{k,t,N}]^{\mathcal{T}}$ as the collection of all the received signals from the N antennas at the kth subcarrier of the tth transmission. Assume that all the received signal are aligned at the BS [12]. Let $z_{j,k,t}$ be the tth symbol of $\mathbf{z}_j(k)$ and $\mathbf{h}_{j,k,t} = [h_{j,k,1}, h_{j,k,2}, \ldots, h_{j,k,N}]^{\mathcal{T}}$ denote the Rayleigh fading channel coefficient vector at the kth subcarrier from the jth user to the BS. Then, the received signals are given as

$$\begin{aligned} \mathbf{r}_{k,t} &= \sum_{j\in\boldsymbol{\mathcal{I}}_r(k)} \mathbf{h}_{j,k,t} z_{j,k,t} + \mathbf{n}_t, \\ &= \mathbf{H}_{k,t}\mathbf{w}_{k,t} + \mathbf{n}_t, \end{aligned} \tag{5}$$

where $\mathbf{n}_t = [n_{1,t}, n_{2,t}, \ldots, n_{N,t}]^{\mathcal{T}}$ is the Gaussian white noise with zero mean and variance σ_w^2, $\mathbf{w}_{k,t} = \left[z_{j_1}, z_{j_2}, \ldots, z_{j_{d_f}}\right]^{\mathcal{T}}$ and $\mathbf{H}_{k,t} = \left[\mathbf{h}_{j_1,k,t}, \mathbf{h}_{j_2,k,t}, \ldots, \mathbf{h}_{j_{d_f},k,t}\right]$, $j_i \in \boldsymbol{\mathcal{I}}_r(k)$.

C. Differential demodulation

Assuming that the channel coefficient remains unchanged over a duration of at least two adjacent symbols, the BS performs the differential demodulation of two consecutive received symbols at the kth resource as

$$\begin{aligned} y_{k,t} =& \frac{\mathbf{r}_{k,t-1}^{\mathcal{H}}\mathbf{r}_{k,t-1}}{N} = \frac{1}{N}\mathbf{w}_{k,t-1}^{\mathcal{H}}\mathbf{H}_{k,t}^{\mathcal{H}}\mathbf{H}_{k,t}\mathbf{w}_{k,t-1} \\ &+ \underbrace{\frac{1}{N}\left(\mathbf{w}_{k,t-1}^{\mathcal{H}}\mathbf{H}_{k,t}^{\mathcal{H}}\mathbf{n}_t + \mathbf{n}_{t-1}^{\mathcal{H}}\mathbf{H}_{k,t}\mathbf{w}_{k,t} + \mathbf{n}_{t-1}^{\mathcal{H}}\mathbf{n}_t\right)}_{\text{effective noise } W_{k,t}} \\ =& \frac{1}{N}\sum_{j\in\boldsymbol{\mathcal{I}}_r(k)} \|\mathbf{h}_{j,k,t}\|^2 z_{j,k,t-1}^{\dagger} z_{j,k,t} + I_{k,t} + W_{k,t} \\ \overset{\text{(i)}}{=}& \frac{1}{N}\sum_{j\in\boldsymbol{\mathcal{I}}_r(k)} \|\mathbf{h}_{j,k,t}\|^2 x_{j,k,t} + I_{k,t} + W_{k,t}, \end{aligned} \tag{6}$$

where (i) holds since the differentially modulated symbol yields $z_{j,k,t-1}^{\dagger} z_{j,k,t} = E_{j,k}\overline{z}_{j,k,t-1}^{\dagger}\overline{z}_{j,k,t} = x_{j,k,t}$, and

$$I_{k,t} = \frac{1}{N}\sum_{j_1\in\boldsymbol{\mathcal{I}}_r(k)} \sum_{\substack{j_2\in\boldsymbol{\mathcal{I}}_r(k), \\ j_1\neq j_2}} \mathbf{h}_{j_1,k,t}^{\mathcal{H}}\mathbf{h}_{j_2,k,t} z_{j,k,t-1}^{\dagger} z_{j,k,t}, \tag{7}$$

$$\begin{aligned} W_{k,t} =& \frac{1}{N}\Bigg(\Big(\sum_{j\in\boldsymbol{\mathcal{I}}_r(k)} \mathbf{h}_{j_1,k,t} z_{j,k,t-1}\Big)^{\mathcal{H}} \mathbf{n}_{k,t} \\ &+ \mathbf{n}_{k,t-1}^{\mathcal{H}} \sum_{j\in\boldsymbol{\mathcal{I}}_r(k)} \mathbf{h}_{j_1,k,t} z_{j,k,t} + \mathbf{n}_{k,t-1}^{\mathcal{H}}\mathbf{n}_{k,t}\Bigg), \end{aligned} \tag{8}$$

are referred to as the multiple access interference and effective noise, respectively [12]. In (6), due to the channel hardening, the off-diagonal terms of the $\mathbf{H}_{k,t}^{\mathcal{H}}\mathbf{H}_{k,t}$ matrix become increasingly small compared to the diagonal terms, i.e., $\mathbf{H}_{k,t}^{\mathcal{H}}\mathbf{H}_{k,t} \to \mathbf{I}_{d_f}$ for $N \to \infty$ [13]. Hence, the multiple access interference, which involves the off-diagonal terms of the $\mathbf{H}_{k,t}^{\mathcal{H}}\mathbf{H}_{k,t}$ matrix, can be modeled as a Gaussian noise. In addition, the $I_{k,t} + W_{k,t}$ term tends to subject to a Gaussian distribution with with zero mean and σ^2 variance, where $\sigma^2 = \mathbb{E}\{(I_{i,k} + W_{i,k})^\dagger (I_{i,k} + W_{i,k})\} = \frac{1}{N}(d_f(d_f-1) + 2d_f\sigma_w^2 J d_f + 1)$. By central limit theorem, this approximation is accurate for large N. Let $\mathbf{y}_t = [y_{1,t}, y_{2,t}, \ldots, y_{K,t}]^\mathcal{T}$ be the differentially demodulated signals on the K resources, then $\mathbf{y}_t$ can be expressed as

$$\mathbf{y}_t = \sum_{j=1}^{J} \operatorname{diag}\left(\overline{\mathbf{h}}_{j,t}\right)\mathbf{x}_{j,t} + \mathbf{I}_t + \mathbf{W}_t, \tag{9}$$

where $\overline{\mathbf{h}}_j = \frac{1}{N}\left[\|\mathbf{h}_{j,1,t}\|^2, \|\mathbf{h}_{j,2,t}\|^2, \ldots, \|\mathbf{h}_{j,K,t}\|^2\right]^\mathcal{T}$ is the jth user's effective channel vector, $\mathbf{y} = [y_1, y_2, \ldots, y_K]^\mathcal{T}$, $\mathbf{I}_t = [I_{1,t}, I_{2,t}, \ldots, I_{K,t}]^\mathcal{T}$ and $\mathbf{W}_t = [W_{1,t}, W_{2,t}, \cdots, W_{K,t}]^\mathcal{T}$.

III. Implementation issues of NC-SCMA

A. Channel hardening enabled MPA

Thanks to the sparsity of the SCMA codewords, the MPA detector is applied to reduce the decoding complexity. Define $L_{r_k\to u_j}^{(q)}(\mathbf{x}_{j,t})$ as the belief message from resource node r_k to variable node u_j at the qth iteration of codeword $\mathbf{x}_j$. Similar, let $L_{u_j\to r_k}^{(q)}(\mathbf{x}_{j,t})$ be the belief message updated from variable node u_j to resource node r_k. The iterative message exchange between resource nodes and variable nodes is computed as

$$\begin{aligned} &L_{r_k\to u_j}^{(q)}(\mathbf{x}_{j,t}) \\ &= \sum_{\substack{\mathbf{x}_{j,t}=\mathbf{x} \\ i\in\boldsymbol{\mathcal{I}}_r(k)\setminus\{j\} \\ \mathbf{x}_{i,t}\in\mathcal{X}_{i,t}}} \frac{1}{\pi\sigma}\exp\left\{-\frac{\left|y_{k,t} - \sum\limits_{i\in\boldsymbol{\mathcal{I}}_r(k)} \frac{\|\mathbf{h}_{j,k,t}\|^2}{N} x_{i,k,t}\right|^2}{\sigma^2}\right\} \\ &\qquad \times \prod_{i\in\boldsymbol{\mathcal{I}}_r(k)\setminus\{j\}} L_{u_i\to r_k}^{(q-1)}(\mathbf{x}_{i,t}), \end{aligned} \tag{10}$$

and

$$L_{u_j\to r_k}^{(q)}(\mathbf{x}_{j,t}) = \alpha_j \prod_{\ell\in\boldsymbol{\mathcal{I}}_u(j)\setminus\{k\}} L_{r_\ell\to u_j}^{(q-1)}(\mathbf{x}_{j,t}), \tag{11}$$

where α_j is a normalization factor. We assume equal probability for the input message $\mathbf{x}_{j,t}$.

Clearly, the instantaneous channel power $\|\mathbf{h}_{j,k,t}\|^2$ is required to compute the belief message at the resource node in (10). It is worth mentioning that in massive MIMO systems, the estimation of the channel power is generally less complicated than the full channel information estimation [12], [13]. Moreover, due to the channel hardening in massive MIMO, $\mathbf{H}_{k,t}^{\mathcal{H}}\mathbf{H}_{k,t}$ tends to be an identity matrix, i.e., $\frac{1}{N}\mathbf{H}_{k,t}^{\mathcal{H}}\mathbf{H}_{k,t} \to \mathbf{I}_{d_f}$. Namely, the following asymptotic property holds [13]:

$$\lim_{N\to\infty} \frac{\|\mathbf{h}_{j,k}\|^2}{N} = 1. \tag{12}$$

Hence, by employing the asymptotic property of (12), only statistical channel power is required in MPA.

B. Proposed codebook design for NC-SCMA

In this subsection, we investigate the codebook design for the proposed NC-SCMA system. For simplicity, the subscript t in (9) is omitted whenever no ambiguity arises.

1) Design criteria: Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_J]_{K\times J}$ and $\overline{\mathbf{H}} = \left[\overline{\mathbf{h}}_1, \overline{\mathbf{h}}_2, \ldots, \overline{\mathbf{h}}_J\right]_{K\times J}$ collect J users' transmitted vectors and channel gains, respectively. Due to the multiuser interference, assume that $\mathbf{X}$ is erroneously decoded to another $K \times J$ matrix $\hat{\mathbf{X}}$, $\hat{\mathbf{X}} \neq \mathbf{X}$, i.e., $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_J]_{K\times J}$. Since the noise term in (9) can be modeled as a Gaussian random variable for $N \to \infty$, the conditional pairwise error probability (PEP) between $\hat{\mathbf{X}}$ and $\mathbf{X}$ is given as [1]

$$\begin{aligned} \Pr\{\mathbf{X}\to\hat{\mathbf{X}}|\overline{\mathbf{H}}, N\to\infty\} &= Q\left(\sqrt{\frac{\left\|\sum_{j=1}^{J}\operatorname{diag}(\overline{\mathbf{h}}_j)(\mathbf{x}_j - \hat{\mathbf{x}}_j)\right\|^2}{2\sigma^2}}\right) \\ &= Q\left(\sqrt{\frac{\left\|\sum_{j=1}^{J}(\mathbf{x}_j - \hat{\mathbf{x}}_j)\right\|^2}{2\sigma^2}}\right), \end{aligned} \tag{13}$$

where $Q(\cdot)$ denotes the Q-function, and the last equality holds by applying (12). Let $\mathbf{v} = \sum_{j=1}^{J}\mathbf{x}_j$ denote the superimposed codewords, which constitute a superimposed constellation Φ, i.e., $\Phi = \left\{\sum_{j=1}^{J}\mathbf{x}_j | \forall \mathbf{x}_j \in \boldsymbol{\mathcal{X}}_j, j = 1, 2, \ldots, J\right\}$. We further denote $d_{\min}^2$ as the minimum Euclidean distance (MED) of Φ, which can be obtained as

$$d_{\min}^2 = \min\left\{\| \mathbf{v}_n - \mathbf{v}_m \|^2, \forall \mathbf{v}_n, \mathbf{v}_m \in \Phi, \mathbf{v}_n \neq \mathbf{v}_m\right\}. \tag{14}$$

Then, the error rate of the transmitted vector is expressed as

$$\Pr\{\mathbf{X}\} = \sum_{\hat{\mathbf{X}}, \hat{\mathbf{X}}\neq\mathbf{X}} \Pr\{\mathbf{X}\to\hat{\mathbf{X}}\} \leqslant \left(M^J - 1\right) Q\left(\sqrt{\frac{d_{\min}^2}{2\sigma^2}}\right). \tag{15}$$

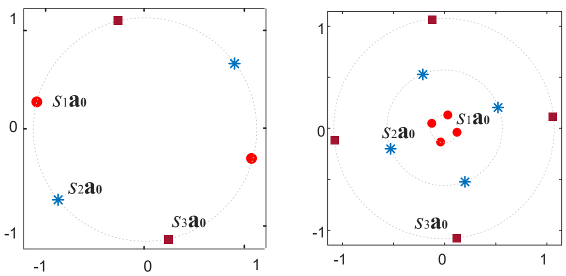
Clearly, by exploiting the channel hardening in the NC-SCMA, the BER performance is mainly dominated by $d_{\min}^2$, which is also a codebook design criteria in the Gaussian channel for the conventional SCMA system [14]. As mentioned in Subsection II-B, another codebook design constraint for NC-SCMA is that the M codewords in each dimension of $\boldsymbol{\mathcal{X}}_j$ should have the same magnitude of $E_{j,k}$. This is in sharp contrast to the conventional codebook design. Hence, the codebook design for NC-SCMA is formulated as

$$\begin{aligned} \max_{\boldsymbol{\mathcal{X}}_j} \quad & d_{\min}^2 && (16) \\ \text{s.t.} \quad & \sum\nolimits_1^J \operatorname{tr}(\boldsymbol{\mathcal{X}}_j^{\mathcal{H}}\boldsymbol{\mathcal{X}}_j) = MJ, && (16a) \\ & |x_{j,k}^{(m)}|^2 = E_{j,k}, E_{j,k} > 0, && (16b) \\ & 1 \le m \le M, 1 \le j \le J. && (16c) \end{aligned}$$

2) Proposed design approach: It is quite challenging, if not impossible, to directly solve the problem (16) due to the newly introduced constraint (16b) and the prohibitively high computational complexity of transforming the $d_{\min}$ into a convex expression, thus an efficient multi-stage design approach with sub-optimal solutions is developed to address (16). The proposed detailed design is shown in **Algorithm** 1. In Step 1, considering the power constraint introduced by the differential modulation, i.e., (16b). we choose the M-PSK for $\mathbf{a}_0$. In Step 2, the mother constellation $\mathbf{A}_{MC}$ is obtained by the repetition of $\mathbf{a}_0$, i.e., $\mathbf{A}_{MC} = [\underbrace{\mathbf{a}_0^\mathcal{T}, \mathbf{a}_0^\mathcal{T}, \ldots, \mathbf{a}_0^\mathcal{T}}_{V}]^\mathcal{T}$.

Algorithm 1 Codebook construction for NC-SCMA

1: **Step 1** : Choose a one-dimensional basic constellation $\mathbf{a}_0$.
2: **Step 2** : Generate the MC $\mathbf{A}_{MC}$ based on $\mathbf{a}_0$.
3: **Step 3** : Design the jth user's constellation operator $\mathbf{\Lambda}_j \in \mathbb{C}^{N \times V}$.
4: **Step 4** : Search for $\mathbf{\Lambda}_j$ for each user by addressing (16). Generate the jth user's codebook by $\boldsymbol{\mathcal{X}}_j = \mathbf{V}_j \mathbf{\Lambda}_j \mathbf{A}_{MC}$.



(a) $M = 2$ (b) $M = 4$

Fig. 2: Proposed codebooks for NC-SCMA system.

In Step 3, the user operators are usually the combination of phase rotation, power scaling and dimension permutation. Specifically, dimension permutation is first applied to enlarge the MED of $\mathbf{A}_{MC}$. Let $\boldsymbol{\pi}_v$ denote the permutation mapping of the vth dimension, then the V-dimensional constellation can be obtained as

$$\mathbf{A}_{MC}^{*} = \left[\boldsymbol{\pi}_1(\mathbf{a}_0), \boldsymbol{\pi}_2(\mathbf{a}_0), \ldots, \boldsymbol{\pi}_V(\mathbf{a}_0)\right]^{\mathcal{T}}. \tag{17}$$

Next, the phase rotation and power scaling are applied to $\mathbf{A}_{MC}^{*}$ to generate the jth user's codebook. Let $\mathbf{\Lambda}_j = \mathrm{diag}\left(\left[\lambda_1 e^{j\phi_1}, \lambda_2 e^{j\phi_2}, \ldots, \lambda_V e^{j\phi_V}\right]\right)$ denote the combination of phase rotation and power scaling matrix, where $0 < \lambda_v$, $-\pi \leq \phi_v \leq \pi$. Then, the jth user's codebook is generated by $\boldsymbol{\mathcal{X}}_j = \mathbf{V}_j \mathbf{\Lambda}_j \mathbf{A}_{MC}^{*}$, where $\mathbf{V}_j \in \mathbb{B}^{K \times V}$ is the mapping matrix that maps the V-dimensional MC to the K-dimensional sparse codebook [1]. Note that the J codebooks can be represented by the signature matrix $\mathbf{S}_{K \times J} = \left[\mathbf{s}_{K \times J}^{1}, \ldots, \mathbf{s}_{K \times J}^{J}\right]$, where $\mathbf{s}_{K \times J}^{j} = \mathbf{V}_j \mathbf{\Lambda}_j \mathbf{I}_K$. In this paper, the following signature matrix is employed [15]:

$$\mathbf{S}_{4 \times 6} = \begin{bmatrix} 0 & s_1 & s_2 & 0 & s_3 & 0 \\ s_1 & 0 & s_2 & 0 & 0 & s_3 \\ 0 & s_3 & 0 & s_2 & 0 & s_1 \\ s_3 & 0 & 0 & s_2 & s_1 & 0 \end{bmatrix}, \tag{18}$$

where $s_i = \lambda_i e^{j\phi_i}$, $i = 1, 2, \ldots, d_f$. Based on the above discussed design criteria, the parameters λ_i, ϕ_i are optimized to maximize $d_{\min}^2$ under the constraint of (16a). The optimal permutation $\boldsymbol{\pi}_v$ is obtained by exhaustive search with negligible complexity for $V = 2$ and $M \leq 4$. In addition, we employ the genetic algorithm, a classical meta-heuristics algorithm, to address (16).

[1] $\mathbf{V}_j$ can be constructed according to the position of the '0' elements of $\mathbf{f}_j$ by inserting all-zero row vectors into the identity matrix $\mathbf{I}_V$. For example, $\mathbf{V}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{\mathcal{T}}$ and $\mathbf{V}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}^{\mathcal{T}}$, and $\mathbf{V}_j, j = 3, 4, 5$ and 6 can be generated in a similar way.

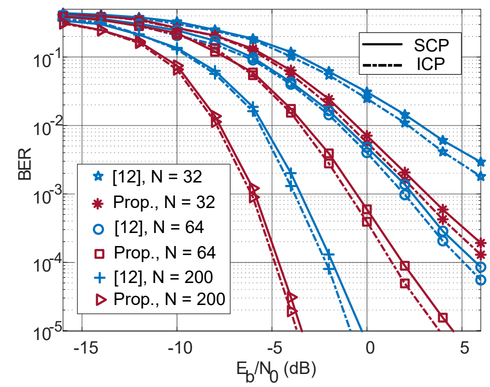


Fig. 3: BER comparison of NC-SCMA with [12] for $M = 2$.

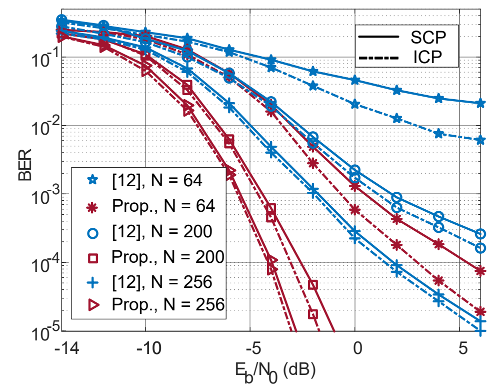


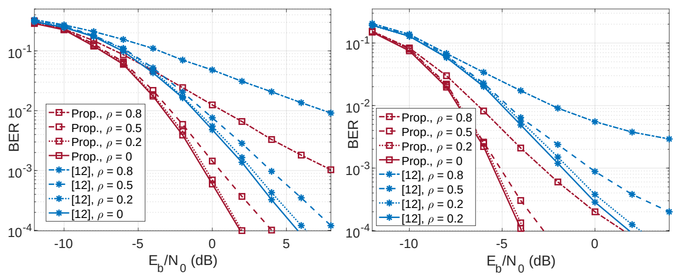
Fig. 4: BER comparison of NC-SCMA with [12] for $M = 4$.

IV. Simulation Results

This Section presents the computational complexity and the BER performance of the proposed NC-SCMA scheme. The number of transmission bits is set to $T = 200$ for each block. The system with $J = 6$, $K = 4$ and various number of receive antennas are considered. Compared to the conventional SCMA, the additional computational complexity introduced by the NC-SCMA is the differential modulation and differential demodulation modules, whose numbers of operations are given by $2JV$ and KN, respectively. Obviously, the additional computational complexity is linear to N and is relatively small compared to that of MPA for medium value of N. For the codebook design, the computational complexity of calculating $d_{\min}^2$ can be approximated as $\mathcal{O}(M^{2J})$. For large size codebook, e.g., $M \geq 16$ and $V \geq 3$, the binary switching algorithm in [16] can be employed to obtain a sub-optimal solution of (17) with the complexity of $\mathcal{O}(M^2)$.

Fig. 2 shows the designed sub-constellations with $M = 2$ and $M = 4$ [2]. As can be seen from this figure, each sub-constellation has distinct phase rotation and power. The constellation points in each sub-constellation has the same energy to enable differential modulation. The MEDs for the 2-ary and 4-ary codebooks are $d_{\min} = 2.0$ and $d_{\min} = 1.14$, respectively, where the codebook energy is normalized to M, i.e., $\mathrm{tr}(\boldsymbol{\mathcal{X}}_j^{\mathcal{H}} \boldsymbol{\mathcal{X}}_j) = M$. We next evaluate the system performance with the above codebooks and compare the proposed system with the non-coherent scheme in [12].

[2] Interested readers can also find our designed codebooks at https://github.com/ethanlq/SCMA-codebook



(a) $M = 2$, $N = 64$ (b) $M = 4$, $N = 256$

Fig. 5: BER performances of the proposed NC-SCMA system in correlated channels.

A. BER performance of NC-SCMA

Fig. 3 presents the BER performance with different number of receive antennas of $M = 2$. The solid lines denote the BER curves of the proposed NS-SCMA by employing the statistical channel power (SCP), i.e., the asymptotic property in (12), whereas the dash lines denote the BER curves under perfect and instantaneous channel power (ICP). Obviously, the proposed NC-SCMA outperforms the scheme in [12]. In particular, the proposed NC-SCMA achieves about 3 dB gain compared to the scheme in [12] at BER = 10^{-4}, where $N = 200$. It can also be seen that the performance gap between the MPA with instantaneous channel information and that with statistical channel information is relatively small, and the gap becomes negligible as N becomes large. This is because the asymptotic approximation in (12) becomes more accurate as N goes larger. Fig. 4 shows the BER performance of the proposed scheme and [12] with $M = 4$. It shows that the proposed NC-SCMA outperforms [12] significantly. For example, our proposed scheme achieves about 6 dB gain compared to [12] at BER = 10^{-4} of $N = 256$.

B. Performance evaluation in correlated channels

For massive MIMO systems, antennas may be compactly arranged, thus the correlation issue cannot be ignored [17]. Hence, we further evaluate our proposed NC-SCMA scheme in correlated channels. Specifically, we consider the Kronecker correlation model, where the channel matrix for each user is modeled as [17]

$$\mathbf{H}_{\text{cor},j} = \mathbf{R}^{1/2}\,\mathbf{H}_j, \tag{19}$$

where $\mathbf{H}_j = [\mathbf{h}_{j,1}, \mathbf{h}_{j,2}, \ldots, \mathbf{h}_{j,K}]^{\mathcal{T}} \in \mathbb{C}^{N\times K}$ is the propagation channel over the air of user j, and $\mathbf{R}_{N\times N}$ is the correlation matrix at BS, which takes the following expression

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho^4 & \cdots & \rho^{(N-1)^2} \\ \rho & 1 & \rho & \ddots & \rho^{(N-2)^2} \\ \rho^4 & \rho & 1 & \ddots & \rho^{(N-3)^2} \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ \rho^{(N-1)^2} & \cdots & & \cdots & 1 \end{bmatrix}, \tag{20}$$

where $0 \le \rho \le 1$ denotes the correlation factor.

Fig. 5 shows the BER comparison of the NC-SCMA and the scheme in [12] under the correlated channels with $M = 2, N = 64$ and $M = 4, N = 256$, where the statistic channel information is considered. It is shown that for both $M = 2$ and $M = 4$, the BER performance degrades noticeably for $\rho = 0.8$; however, the performance loss is small when $\rho \le 0.5$. Moreover, the proposed NC-SCMA is also more robust against the channel correlation than that of [12].

V. Conclusion

In this letter, a novel NC-SCMA system with massive MIMO has been introduced to enable multi-user detection with statistical channel information. Differential modulation has been introduced after the SCMA encoding. At the receiver side, the differential demodulation and channel hardening enabled MPA detection have been developed for NC-SCMA system. The detailed design criteria and related implementation issues of the sparse codebooks in the NC-SCMA have also been investigated. Numerical results have shown that our proposed system achieves improved BER performance than the benchmark scheme and is robust to the channel correlation.

References

[1] Z. Liu and L.-L. Yang, “Sparse or dense: A comparative study of code-domain NOMA systems,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4768–4780, Aug. 2021.

[2] L. Yu *et al.*, “Sparse code multiple access for 6G wireless communication networks: Recent advances and future directions,” *IEEE Commun. Stand. Mag.*, vol. 5, no. 2, pp. 92–99, Apr. 2021.

[3] R. Hoshyar, F. P. Wathan, and R. Tafazolli, “Novel low-density signature for synchronous CDMA systems over AWGN channel,” *IEEE J. Sel. Areas Commun.*, vol. 56, no. 4, pp. 1616–1626, Mar. 2008.

[4] M. Taherzadeh *et al.*, “SCMA codebook design,” in *IEEE VTC2014-Fall*, Vancouver, Canada, Dec. 2014, pp. 1–5.

[5] Q. Luo *et al.*, “An error rate comparison of power domain non-orthogonal multiple access and sparse code multiple access,” *IEEE Open J. Commun. Soc.*, vol. 2, no. 4, pp. 500–511, Mar. 2021.

[6] M. Wang, F. Gao, S. Jin, and H. Lin, “An overview of enhanced massive MIMO with array signal processing techniques,” *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 5, pp. 886–901, Sept. 2019.

[7] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, “Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2413–2424, Oct. 2017.

[8] Z. Ding, R. Schober, and H. V. Poor, “A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4438–4454, 2016.

[9] M. Chowdhury, A. Manolakos, and A. Goldsmith, “Scaling laws for noncoherent energy-based communications in the simo mac,” *IEEE Trans Inf. Theory*, vol. 62, no. 4, pp. 1980–1992, Apr. 2016.

[10] Y.-Y. Zhang, J.-K. Zhang, and H.-Y. Yu, “Physically securing energy-based massive MIMO MAC via joint alignment of multi-user constellations and artificial noise,” *IEEE J. Sel. Areas Commun.,*, vol. 36, no. 4, pp. 829–844, Apr. 2018.

[11] H. Xie, W. Xu, H. Q. Ngo, and B. Li, “Non-coherent massive MIMO systems: A constellation design approach,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3812–3825, Mar. 2020.

[12] T. Wang, L. Shi, K. Cai, L. Tian, and S. Zhang, “Non-coherent NOMA with massive MIMO,” *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 134–138, Oct. 2020.

[13] T. L. Narasimhan and A. Chockalingam, “Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems,” *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 847–860, Apr. 2014.

[14] Q. Luo *et al.*, “A novel multi-task learning empowered codebook design for downlink SCMA networks,” *IEEE Wireless Commun. Lett.*, 2022.

[15] X. Li *et al.*, “Design of power-imbalanced SCMA codebook,” *IEEE Tran. on Veh. Techno.*, vol. 71, no. 2, pp. 2140–2145, Feb. 2022.

[16] K. Zeger and A. Gersho, “Pseudo-gray coding,” *IEEE Trans. Commun.*, vol. 38, no. 12, pp. 2147–2158, Dec. 1990.

[17] P. S. Taluja and B. L. Hughes, “Diversity limits of compact broadband multi-antenna systems,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 326–337, Jan. 2013.