Prototype for 5G New Air Interface Technology SCMA and Performance Evaluation

LU Lei, CHEN Yan, GUO Wenting, YANG Huilian, WU Yiqun, XING Shuangshuang
Huawei Technologies, Co., Ltd., Shanghai, China

Abstract: Sparse code multiple access (SCMA) is a novel non-orthogonal multiple access scheme proposed to meet the challenging demand of the future 5G communications, especially in support of the massive connections. The coded bits from each data stream will be directly mapped as multi-dimensional SCMA codeword in complex domain and then spread onto the physical resource elements in a sparse manner. The number of codewords that can be non-orthogonally multiplexed in one SCMA block can be made much larger than the number of orthogonal resource elements therein, resulting in an overloaded system. The sparsity in the spreading pattern and the design in the multi-dimensional modulator jointly ensure the SCMA codewords can be robustly decoded with low complexity. In this paper, we focus on the low complexity receiver design and verified the superior of an SCMA system via simulations and real-time prototyping. Lab tests and field tests all show that SCMA is a promising candidate for 5G non-orthogonal multiple access which can provide up to 300% overloading that triples the whole system throughput while still enjoying the link performance close to orthogonal transmissions.

Keywords: 5G; non-orthogonal multiple access; sparse code multiple access (SCMA); MAX-Log MPA; massive connectivity; low latency; high reliability; Prototype verification

I. INTRODUCTION

As the fourth generation (4G) commercial networks are continuously deployed worldwide, the research for the fifth generation (5G) is already in full swing. A global network connecting everything – people, machines, cars, pets, and other intelligent devices (to name a few) – is right in the making. In particular, 5G is going to bring the mobile broadband across varied vertical industry segments and enable the concept of internet-of-things (IoT) for real [1,2]. As commonly agreed among major mobile operators and released by NGMN in its white paper [3] that 5G will serve a far more diversified scenario set than 4G, supporting broadband everywhere even in very dense areas, high mobility users in cars or on fast trains, massive connectivity from all types low cost monitors, ultra low latency for Virtual Reality (VR) or Augmented Reality (AR), and ultra reliable transmission for the intelligent control of industry 4.0. Compared with 4G network today, 1ms ultra low delay, 100x higher data rate, 1000x more connections, up to 500km/h speed are extreme ends of the diverse requirement. Though luckily they are not all needed in one application at the same time, any one of the requirement is enough to trigger the rethinking and redesign of the mobile networks, not to mention that some of them are simultaneously demanded.

Air interface design has always been the

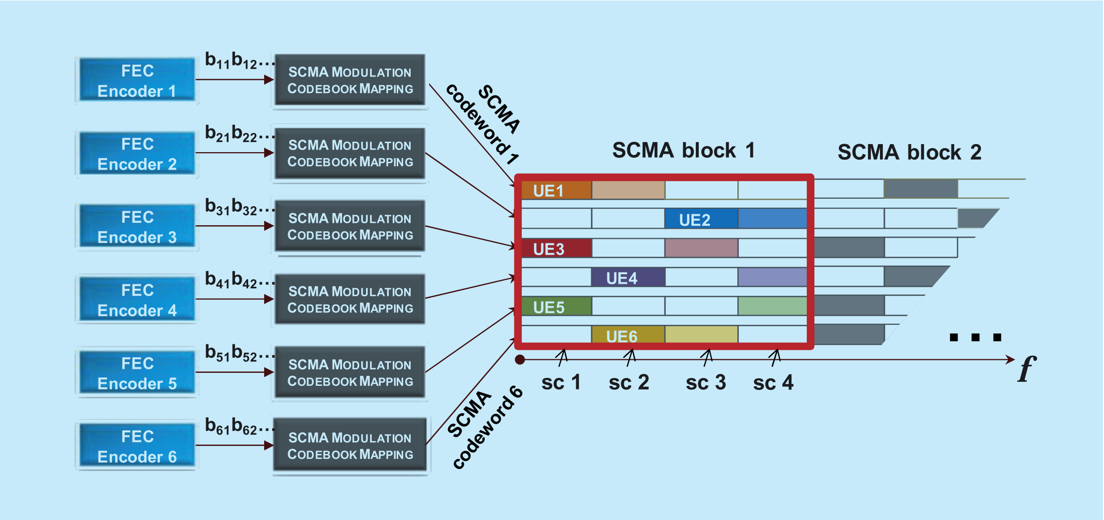

Fig. 1 *Illustration of an uplink multi-user SCMA system*

crown jewel and the foundation of every generation of mobile communications, while the design of waveform and multiple access technology is the very first step on its way of standardization. In current 4G systems, orthogonal frequency division multiple access (OFDMA) techniques are used, which is an orthogonal multiple access scheme. In particular, in OFDMA, radio resource is divided into two-dimensional orthogonal time-frequency grids and each grid can only be used by one user at a time. It is obvious that the number of simultaneously accessible user is strictly proportional to the number of available orthogonal resources, and is thus limited. Facing the 5G requirement of massive connectivity, non-orthogonal multiple access becomes the research focus of 5G multiple access technologies, among which, Sparse code multiple access (SCMA) [4] is a promising candidate. At the transmitter, coded bits are directly mapped to multi-dimensional codewords in complex domain, and codewords from different users are overlapped non-orthogonally in a sparse spreading way; the receiver performs joint multiuser detection followed by channel decoding for data recovering. Thanks to the sparsity, low complexity algorithms could be design to achieve near optimal detection results. The number of non-orthogonally superposed codewords can be much larger than the number of orthogonal resource units. This leads to the advantage of SCMA to serve more users while using the same resources, thus effectively improve the overall system capability of connections.

In this paper, we will elaborate on how to implement an SCMA system and verify its superior features on providing massive connectivity via both in lab prototype and field test. Specifically, we shall introduce the fundamental design principles in section II, including the basic features of SCMA, the design principles of the SCMA encoder and the decoder. The specification and setup of the SCMA prototype are given in section III, together with the testing performance. Section IV concludes the whole paper.

II. Fundamentals of SCMA

2.1 Design principles and features

Fig. 1 gives an illustration of an uplink multi-user SCMA system. At the transmitter side, for each user equipment (UE), the information bits after channel coding are then fed to the SCMA encoder which maps the coded bits to SCMA codewords. An SCMA codeword is a multi-dimensional constellation vector, which is then spread onto the physical resource elements (REs) of a multi-carrier system. Here we consider OFDM as the basic waveform, so one RE is consisted by one OFDM symbol and one subcarrier. Note that codewords can be assigned to the same UE for large packets transmission or to multiple UEs for small packets massive connections. Such system has the following features.

- *Non-orthogonal* After spreading, at the receiver side, there are multiple modulation symbols from different users superposed together on each RE. For example, in figure 1, on subcarrier 1, symbols from UE1, 3, and 5 are overlapped with each other.
- *Overlaoding* Due to the non-orthogonal nature, each SCMA system can accommodate more data layers than its spreading length. For example, in figure 1, 6 data layers can be supporting by length 4 spreading, resulting in an overloading rate 150%.
- *Sparse Spreading* Different from the traditional CDMA spreading, SCMA is a new frequency domain spreading with low density. The spreading enhances the robustness of link adaptation and is good for coverage improvement, while the sparsity in the spreading help to limit the total interference on each RE and at the same time limit the complexity of the receiver.
- *Multi-dimension* In general the non-zero spread modulation symbols from the same data layer are different but their dependency is optimized to provide a large average distance between any two points on the constellation. A simplified version with the same modulation symbols repeated on the non-zero tones is usually referred as low density spreading (LDS).

2.2 Codebook and encoder

SCMA codebook design is the key to ensure good performance and flexibility of the whole SCMA system. As mentioned above, the SCMA codebook is the joint optimization of multi-dimensional modulation and low density spreading. It maps the coded bits directly to multi-dimensional codewords in complex domain in a sparse manner, with the modulation symbols on the non-zeros REs different but dependent with each other for each codeword. The motivation for such correlated design is to follow the multi-dimensional modulation principles to maximize the average distance between the constellation points. One example of a 16 points mother constellation on two non-zero tones are given in Fig. 2. Note that under the multidimensional modulation design, the distance between constellation point 1010 and 1001 is small on the first tone but enlarged on the second tone, which yields a good average distance. Detailed design method is out of the scope of this paper and can be referred to [5].

Once having the codebook at hand, SCMA encoder simply selects the codeword corresponding to the input coded bit sequence from the SCMA codebook and then map it to the physical resource elements. An example of SCMA encoding process is presented in Fig. 3. For each SCMA block, there are *J*=6 codebooks stored at 6 users, each used for one data layer transmission. Each codebook is of size K-by-M, where each of the M columns stands for a codeword while each of the K row represents one physical RE. In this example, K =4 and M=4, meaning every log2M=2 bits from each data are mapped to a resource block of length K=4. So in total J=6 data layers are spread and mapped to K=4 REs, resulting in J/K = 6/4 = 150% overloading. Note that there is only N=2

Fig. 2 *Illustrative example of SCMA two dimensional mother constellation*

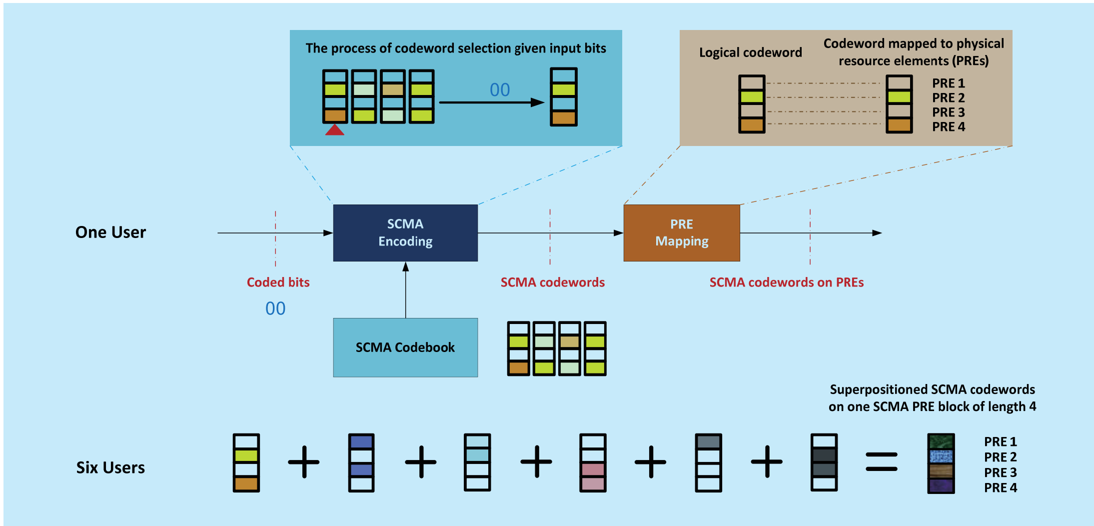

Fig. 3 *Illustrative sketch of the SCMA encoder procedure*

non-zero elements in each codeword, meaning the spreading is done in a low density way.

The set of J SCMA codebooks for one SCMA block can also be represented by Tanner graph [11], taking each RE as a function node (FN) and each data layer as a variable node (VN). The degree of FN, defined as the number of connected (neighboring) VNs for each FN, and the degree of VN, defined as the number of connected (neighboring) FNs for each VN, are two useful parameters to design the structure of the codebook and will greatly impact the complexity in MPA (message passing algorithm) decoding. The important parameters to characterize SCMA systems are summarized in Tab. 1.

Table I *Key parameters for SCMA codebook characterization*

Parameter	Meaning	Example in Fig. 1
J	Number of data layers supported by each SCMA block, different layers can belong to one user or multiple users	6 users each with 1 data layer
K	Number of resource elements in each SCMA block, also known as the spreading factor (SF) of the codebook	4
N	Number of non-zero elements in each codeword	2
M	Number of codewords in each codebook, equals to the maximum constellation points for each data layer on each RE	4
df	Number of connected (neighboring) VNs for each FN	3
dv	Number of connected (neighboring) FNs for each VN	2

2.3 MAX-Log MPA decoder for prototyping

As mentioned above, for a non-orthogonals system like SCMA, there are more than one OFDM symbol superposed on each RE, so joint multi-user detection algorithms are needed. Message passing algorithm (MPA) with affordable complexity can be adopted to achieve near ML performance [6].

The MPA algorithm is performed on the factor graph [7] constructed by the specific codebook. Every FN (representing RE) and VN (representing data layer) which is connected in the tanner graph makes a FN-VN pair. As illustrated in Fig. 4, each VN node is represented by a red circle while each FN node is represented by a blue square. As know from [6], The exp(.) operations in the MPA algorithm may cause very large dynamic ranges and very high storage burden if using lookup table, which is not good news for hardware implementation. Therefore, we learn from the turdo decoder design and change to the log domain for calculation, which removes exp(.) operations by the application of the following Jacobi's logarithm fomula.

$$\log\left(\sum_{i=1}^{N}\exp\left(f_i\right)\right) \approx \max_{i=1,\dots,N}\{f_1, f_2, \dots, f_N\} \quad (1)$$

The input of the MAX-Log MPA algorithm

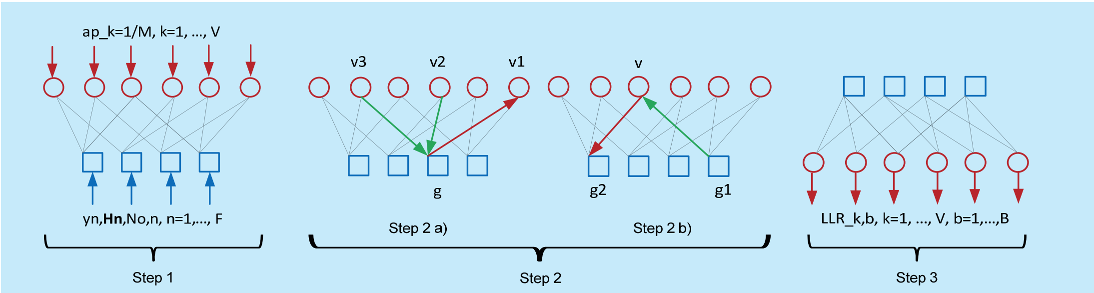

Fig. 4 *Illustrative sketch of the MPA processes*

MAX-Log MPA (see Fig. 4)

Step 1*: Initial calculation of the conditional probability in log domain for each RE n*

$$f(y_n, \{m_i\}, \boldsymbol{H}_n, \Delta_n) = \frac{-1}{\Delta_n}\left\| y_n - \sum_{i=1}^{d_f} h_{n,i} C_{i,g}(m_i) \right\|^2, m_i = 1, \dots, M, n = 1, \dots, K$$

Step 2*: Iterative message passing along edges*

Step 2 a): FN upate for each FN node g, which then passes the extrinsic information to its neighboring VN nodes v_{q} (g to v1, information from v2 and v3 are extrinsic)

$$\boldsymbol{I}_{g\to v_{\mathrm{q}}}^{log}(m_{\mathrm{q}}) = \max_{\{m_i\}_{i\neq \mathrm{q}}} \left\{ f(y_n, \{m_i\}, \boldsymbol{H}_n, \Delta_n) + \sum_{i} \boldsymbol{I}_{v_i\to g}^{log}(m_i) \right\} - \boldsymbol{I}_{v_q\to g}^{log}(m_q)$$

Step 2 b): VN update for each VN node v, which then passes the extrinsic information to its neighboring FN nodes g_{q} (v to g1, information from g2 is extrinsic)

$$\boldsymbol{I}_{v\to g_1}^{log}(m) = \boldsymbol{I}_{g_2\to v}^{log}(m), \qquad \boldsymbol{I}_{v\to g_2}^{log}(m) = \boldsymbol{I}_{g_1\to v}^{log}(m)$$

Step 3*: LLR output at variable node after N_iter iterations*

$$\boldsymbol{I}_{v}^{log}(m) = \sum_{n\in neighbor\ of\ v} \boldsymbol{I}_{g_n\to v}^{log}(m), m = 1, \dots, M$$

$$\mathrm{LLR}(b_i) = \max_{\mathrm{k}\in\{b_i=0\}} \left(\boldsymbol{I}_{v}^{log}(m_k)\right) - \max_{\mathrm{k}\in\{b_i=1\}} \left(\boldsymbol{I}_{v}^{log}(m_k)\right), \mathrm{i} = 1, \dots, \log_2 \mathrm{M}$$

are the received superposed signal y_n, the channel information vector $\mathbf{H}_n$, the noise estimation vector Δ_n, on each RE, and the prior probablility of each codeword, as well as the codebook of each data layer C_i. The output of the algorithm are the log likelihood ratio (LLR) for the coded bits calculated from the probability estimations of each codeword, which are then served as input to the turbo decoder. The MAX-Log MPA algorithm works in the following steps.

2.4 Projected MAX-Log MPA for low complexity

As we can see that even with the employment of MAX-Log MPA, the complexity of the receiver is still of the order $O(M^{d_f})$. In order to further reduce the complexity of the receiver, we can use the projected constellation methodology to construct the codebook [5]. Fig. 5 gives an example of projecting a 16 point constellation to two-dimensional 9 point constellations. The depnedency between the two dimensions gurantees the feasiblity of decoding each constellation point. For instance, though 0000 collides with other constellation points on the right hand size constellation, it is uniquely decodable on the left hand side constellation. Employing such projected codebook, the complexity reduced from $O(M^{d_f})$ to $O(M_1^{d_f})$, with $M_1<M$. Simple calculations in table 2 give the complexity reduction ratios for different

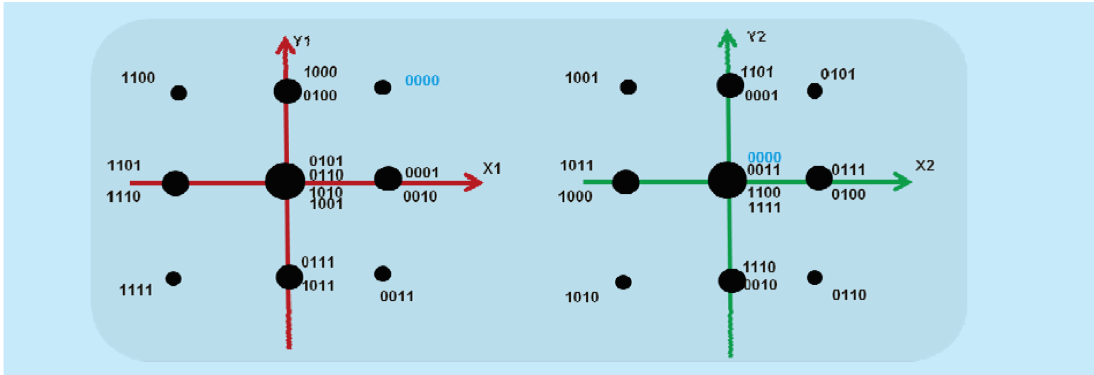

Fig. 5 *Multi-dimensional projected constellation for low complexity receiver*

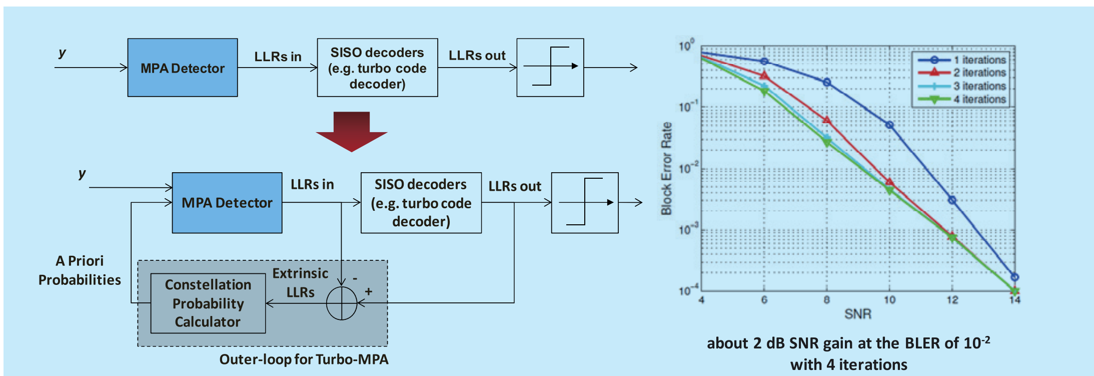

Fig. 6 *Turbo MPA receiver structure and performance evaluation*

configurations. The performance loss due to projection, as shown in [5], is in general less than 0.5 dB in SNR for the above examples.

2.5 Turbo MPA for performance enhancement

To further enhance the robusteness of the multi-user detection, it can further combine the MPA detector with the turbo channel decoder to form a outer-loop turbo MPA receiver structure, as illustrated in Fig. 6. As has been pointed out in [8], such outer-loop decoder can boost the decoding performance especially for large overloaded systems. As shown in Fig. 6, 3 iterations are enough to boost the link level performance for a 300% overloading system by 2 dB.

III. SCMA for Massive Connectivity

As an overloading system, SCMA can support more simultaneous connections than the current non-overloading LTE system, which is suitable for the scenarios with massive connectivity demand, such as hot spot stations, shopping malls, and stadiums, as well as machine types of equipments. In this section, we shall show the link level performance of an SCMA system with different overloading factors, as well as its robustness against codebook collision. While in the next section, we will verify the conclusion here by in lab prototyping and field testing.

3.1 Large and flexible overloading

For a SCMA system, the maximum overloading factor increases as the function of the spreading factor (equal to the number of FN). While keeping the non-zero elements after spreading to be fixed to two (dv=2), for the spreading factor K = 4, 6, 8, and 10, the maximum overloading factor can be (K-1)/2 = 150%, 250%, 350%, and 450%, respectively. Therefore, by tuning the spreading factor K, we can achieve large and

Table II *Complexity reduction due to projection for a 6-by-4 codebook*

Constellation projection	4->3	8->6	16->9
Ratio of complexity reduction	57.8%	57.8%	82.2%

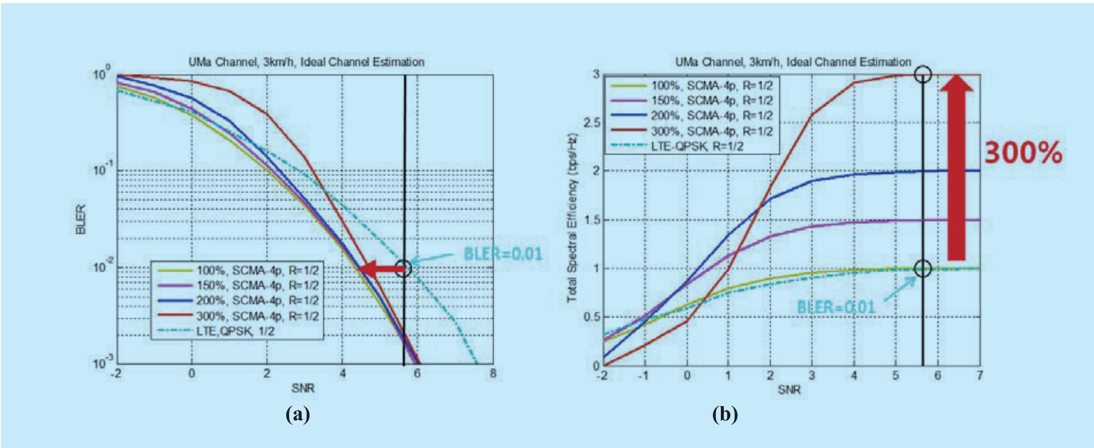

Fig. 7 *The performance of uplink SCMA with high overloading factor*

flexible overloading.

The overloading performance of SCMA is evaluated by link-level simulations. We assume the fading channel is urban-macro, user moving speed is 3 km/h, and channel estimation is perfect. Fig. 7(a) shows the BLER curves of SCMA with the overloading factor 100%, 150%, 200%, and 300%, respectively, and also the 4G LTE OFDMA as baseline. Each UE has the same maximum throughput for all the scenarios. It can be observed that when the overloading factor increases, there is little loss for SCMA and the performance with large overloading can still approach the non-overloading case (100%). This means the inter-user interference has not much influence on the performance. Over the reasonable SNR regime, SCMA performs better than OFDM, and the maximum network throughput can be promoted by 300%, as shown in Fig. 7(b).

3.2 Blind detection and robustness to codebook collision

For sporadic machine type communications (MTC), the packet size is usually very small and the battery life is very limited. The long delay and large signaling overhead in the scheduling based granted access is way too expensive for the low cost MTC equipments to afford. In light of this, contention based grant-free access is proposed to substantially reduce the delay by removing the “asking-for-grant” procedure and

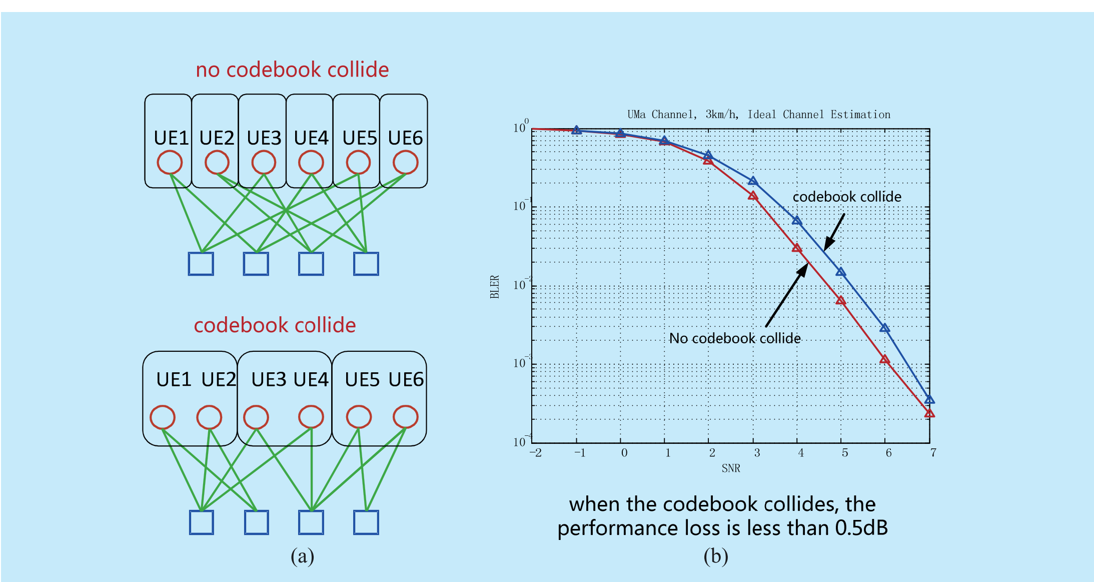

Fig. 8 *codebook collide performance*

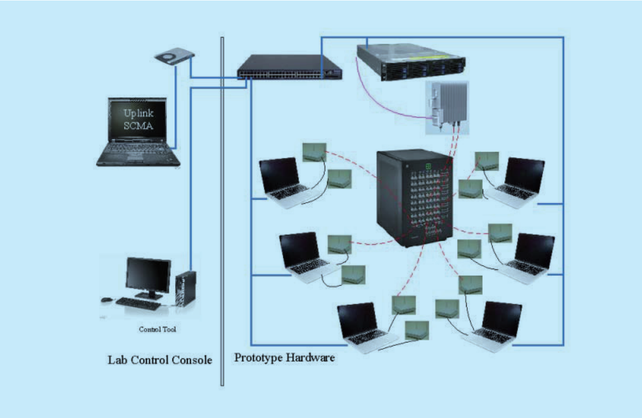

Fig. 9 *The setup of the uplink SCMA demo system*

let users decide the transmission resource to use and enjoy "arrive-and-go" type of service. It is thus very demanding for the multiple access scheme design to support the joint detection of data and user activity states with high reliability, while being tolerant for multi-user collisions. Luckily enough, the blind detection capability of SCMA as well as the collision tolerant nature of the SCMA codebook design, all prove it to be a perfect match for such requirement [9, 10]. In particular, to illustrate the robustness to codebook collision, two different scenarios are depicted in Fig. 8 (a). The upper one has no codebook collision, and the lower one has codebook collision, i.e., UE1 and UE2 use the same codebook. Fig. 8 (b) presents the BLER performance of the two scenarios. The simulation results show that the performance loss due to codebook collision is less than 0.5dB, which is affordable for the real system.

IV. SCMA Prototype with Up To 300% Overloading

4.1 Prototype Setup and Specifications

In order to verify the SCMA technology and its advantages in real communication systems, we have developed an SCMA-based uplink multi-user system prototype on real-time hardware platforms. Our demo system consists of 1 base station with 2 antennas for diversity combined receiving and in total 12 users each with 1 antenna for uplink access and data transmission. The setup of the prototype is shown in fig. 9.

The basic system configurations of our demo system are set to align with the current LTE TDD system. In particular, we use LTE TDD frame structure configuration 1 and take the LTE physical layer with OFDMA orthogonal multiple access as the baseline for performance comparison. The system diagram for SCMA is shown in Fig.10. It is not hard to find that only the SCMA encoder and decoder are different with the current LTE system. This is positive since all the current transmission modes and optimized link design can be reused. The specification of the prototype system is shown in Tab. 3.

The prototype system is built with soft baseband, that is, all the baseband processing is done by CPU instead of FPGA/DSP. At the base station side, one server (Huawei Tecal RH2288)

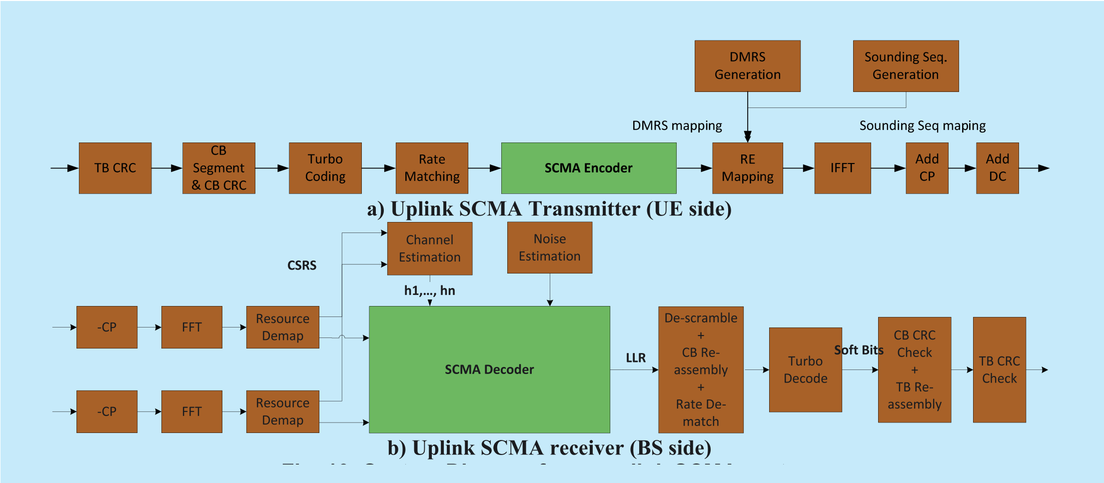

Fig. 10 *System diagram for an uplink SCMA system*

is responsible for all the baseband processing, connecting with the standard commercial radio frequency components (Huawei product RRU3232). At the user side, the CPU of 1 laptop (MacBook Pro ME294CH/A) is used to model the processing of baseband for two users, which is then connecting to two mobile RF modules for testing. A user interface (UI) is developed to show the real-time throughput for each UE and all UEs, supporting also the real-time change of user status and system operation modes.

4.2 In lab and field testing

The prototype can run in either mode, OFDMA or SCMA, separately, and support real-time switching from one to the other. To ensure a fair comparison, we keep the data rate of each user the same to guarantee the same quality of service. We show by the prototype that with the application of SCMA technology over OFDMA, up to 300% overloading gain in the number of connections and network throughput are feasible, compared with the orthogonal multiple access baseline of 4G LTE. For instance, 150% overloading gain can be observed from the fact that given the data volume demand for each user to be 12 physical resource blocks (RBs), a system with a total of 48 RBs can serve at most 4 users using orthogonal LTE OFDMA. However, with SCMA, the codebook design supports 6 users

Table III *Specification for SCMA Prototype*

Mode	Sparse code multiple access
Number of active UEs	12 out of 14
UE Transmit power	23 dBm (max) with open-loop power control
Basic waveform	OFDM / F-OFDM
MIMO mode	1-by-2 SIMO
Center frequency/bandwidth	2.6GHz/20MHz
Scheduled resource	48RBs/4 RBs
Code rate	0.3-0.92
SCMA codebook	24-by-8, 4 points
Frame structure	TDD configuration 1, 4 Subframes for PUSCH

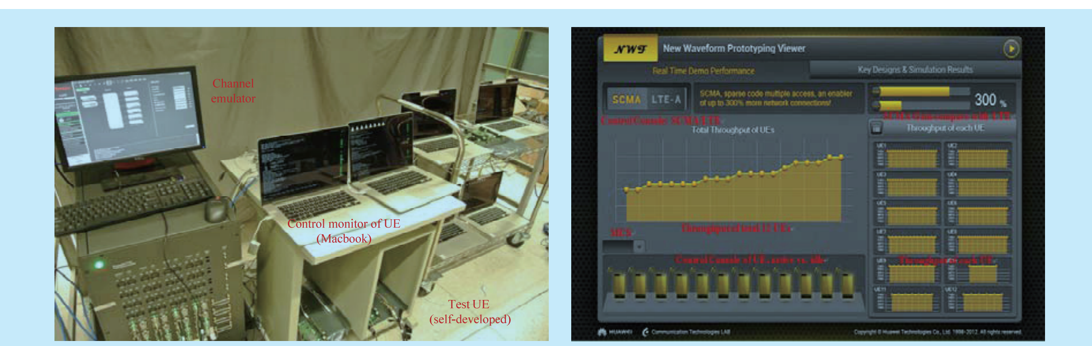

Fig. 11 *Hardware setup and UI of the in lab SCMA demo system group (only 6 out of the 12 are shown)*

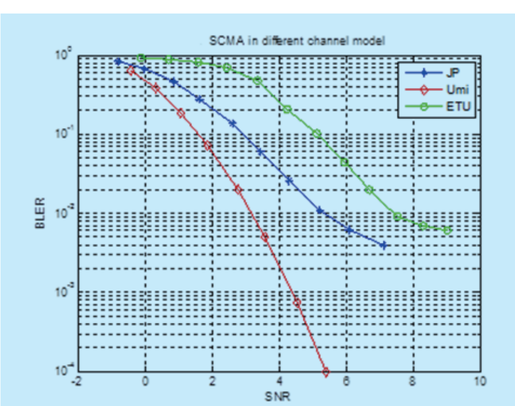

Fig. 12 *SCMA prototype BLER test use different channel model*

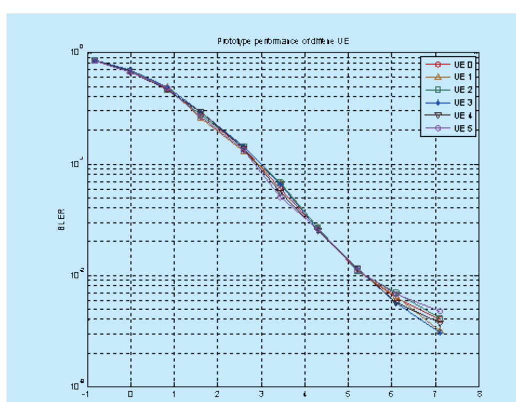

Fig. 13 *BLER performance of different UEs in one SCMA*

with the same amount of data to share the 48 RBs simultaneously, thus the equivalent delivered amount of data is actually 12×6 = 72 RBs other than 12×4 = 48 RBs, resulting in the throughput gain of about 72/48 = 150%. The 300% gain is supported in a similar way but uses a different codebook with larger spreading factor and thus larger number of data layers. In our prototype, we use a 24-by-8 SCMA codebook to allow 12 users each with 2 data streams to access and transmit simultaneously with SCMA, while for LTE OFDMA, only 4 users out of 12 can transmit. A look of the physical hardware of the in lab prototype system and the UI is shown in Fig. 11.

For the SCMA prototyping performance evaluation, we use Anite fading simulator FS32 to test SCMA in different channel, we configure the fading simulator to ETU, Umi and ETU channel model, and test the BLER performance of SCMA prototype, which is show in Fig. 12.

As we can see from Fig. 13 that in the non-orthogonal SCMA mode, the group of UEs which share the same frequency and time resource element has almost the same performance as in the orthogonal LTE mode, which proves the link robustness of SCMA and aligns with the simulation conclusion that the multi-user link performance approaches the single user performance even with large overloading rate.

Moreover, we may further combine SCMA with f-OFDM to improve the spectrum mask and to reduce the out-of-band emission. It has been proved through the prototype that the integration of SCMA and f-OFDM not only provide better localized spectrum than traditional OFDM, but also support asynchronous access between different SCMA groups. The spectrum mask after employing f-OFDM is shown in Fig. 14. Left figure shows the spectrum mask of f-OFDM, while the right figure shows the spectrum mask

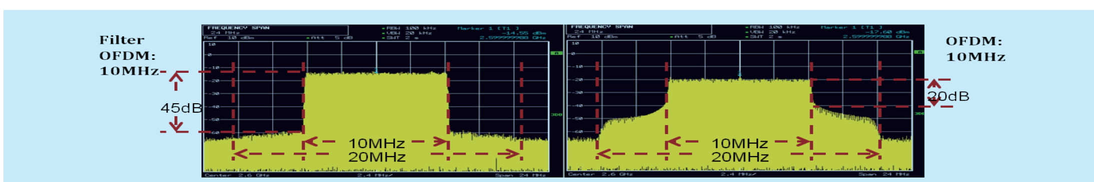

Fig. 14 *Spectrum mask of f-OFDM v.s OFDM*

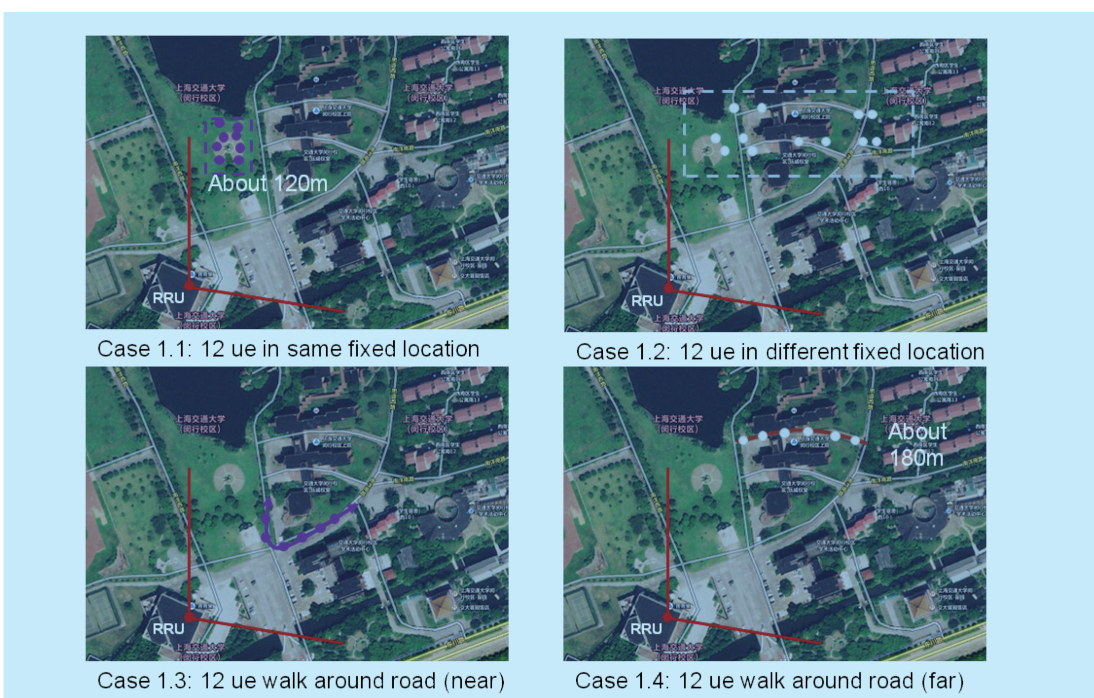

Fig. 15 *Four cases of SCMA field test*

of OFDM as the baseline. Both f-OFDM and OFDM occupy 10 MHz spectrum. From the spectrum analyzer test result, we can see that f-OFDM based system provides 25dB better suppression of channel leakage, thus having much better spectrum mask than OFDM based system.

Besides fading simulator test in Lab, SCMA prototype is deployed to field trial and the performance is tested. We design different test case, UEs are deployed at the different locations to evaluate the performance of SCMA in different conditions. Specifically, we have the following four test cases designed, as shown in Fig. 15.

Case 1.1: 12 UEs closely located in an area without mobility

Case 1.2: 12 UEs locates in an area with distant separation but no mobility

Case 1.3: 12 UE moves along a road about 120 meters from the BS (open-loop power control gives comparatively medium transmit power at UE)

Case 1.4: 12 UE moves along a road about 180 meters from the BS (open-loop power control gives comparatively high power at UE)

In all field trial cases test, we use 20 bytes typical small packets (METIS definition) as payload for both LTE and SCMA, and we limit the scheduling resource in whole system to 4 RBSs in each subframe. The comparative testing results of OFDM and SCMA are shown in Fig. 16. As shown in the figure, SCMA achieves nearly 300% connection gain compared with OFDM.

V. CONCLUSIONS

In this paper, a complete uplink non-orthogonal multiple access system with SCMA technology is presented and implemented. We introduced how SCMA encoder and decoder work and focused on the low complexity receiver design. An SCMA prototype with up to 300% overloading is built to verify the link level simulations results. Results from both in lab testing and field testing support the conclusion that SCMA is a promising candidate for 5G multiple access, which can triple the overall system throughput while still having the link performance close to orthogonal transmissions.

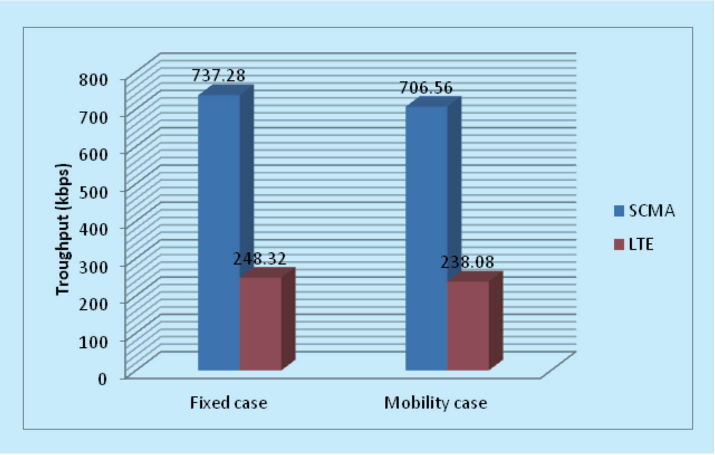

Fig. 16 *SCMA throughput gain over OFDM in field testing*

References

[1] Huawei 5G white paper, http://www.huawei.com/5gwhitepaper/.

[2] IMT-2020 5G white paper, http://www.imt-2020.cn/zh/documents/listByQuery?currentPage=1&content.

[3] NGMN 5G white paper, https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf.

[4] H. Nikopour and H. Baligh, "Sparse Code Multiple Access," IEEE PIMRC, 2013.

[5] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA Codebook Design," IEEE VTC-fall, 2014.

[6] R. Hoshyar, F. Wathan, R. Tafazolli, "Novel low-density signature for synchronous cdma systems over AWGN channel," IEEE Trans Signal Processing, vol. 56, no. 4, Apr 2008.

[7] F. Kshischang, B. Frey, and H. Loeliger, "Factor graphs and the sum-product algorithm," IEEE Trans. Inf. Theory, vol. 47, no. 2, pp. 498-519, Feb. 2001.

[8] Y. Wu, S. Zhang, and Y. Chen, "Iterative multiuser receiver in sparse code multiple access systems," IEEE ICC 2015.

[9] A. Bayesteh, E. Yi, E. , H. Nikopour, H. Baligh, "Blind Detection of SCMA for Uplink Grant-Free Multiple-Access", IEEE ISWCS 2014.

[10] Kelvin Au, Liqing Zhang, Hosein Nikopour, Eric Yi, Alireza Bayesteh, Usa Vilaipornsawai, Jianglei Ma, Peiying Zhu, "Uplink Contention Based SCMA for 5G Radio Access," IEEE Globecom 5G workshop 2014.

[11] H. Nikopour, E. Yi, A. Bayesteh, K. Au, M. Hawryluck, H. Baligh, and Jianglei Ma, "SCMA for Downlink Multiple Access of 5G Wireless Networks," IEEE Globecom 2014.