

Unified Framework Towards Flexible Multiple Access Schemes for 5G

SUN Qi, WANG Sen, HAN Shuangfeng, and Chih-Lin I
(China Mobile Research Institute, Beijing 100032, China)

Abstract

Non-orthogonal multiple access (NOMA) schemes have achieved great attention recently and been considered as a crucial component for 5G wireless networks since they can efficiently enhance the spectrum efficiency, support massive connections and potentially reduce access latency via grant free access. In this paper, we introduce the candidate NOMA solutions in 5G networks, comparing the principles, key features, application scenarios, transmitters and receivers, etc. In addition, a unified framework of these multiple access schemes are proposed to improve resource utilization, reduce the cost and support the flexible adaptation of multiple access schemes. Further, flexible multiple access schemes in 5G systems are discussed. They can support diverse deployment scenarios and traffic requirements in 5G. Challenges and future research directions are also highlighted to shed some lights for the standardization in 5G.

Keywords

5G; non-orthogonal multiple access; unified framework; flexible multiple access

1 Introduction

Worldwide initiatives on the 5th generation (5G) wireless communication have been extensively carried out, starting with an investigation on user demands, scenarios, key performance indicators (KPIs) and enabling technologies. A global consensus is first forming that 5G network will be able to sustainably support 1000-fold mobile data traffic growth, improve energy efficiency (EE) and cost efficiency by over 100 times, provide fiber link access data rates and “zero” latency user experience, and be capable of connecting 100 billion devices and capable of delivering a consistent experience across a variety of scenarios including the cases of ultra-high traffic volume density, ultra-high connection density and ultra-high mobility [1]. Three typical usage scenarios of 5G are also identified: enhanced mobile broadband (eMBB), massive machine type communication (mMTC) and ultra-reliable low latency machine type communication (URLLC), targeting different 5G capabilities. Beyond that, the standardization organizations, e.g. 3GPP has started the new research on 5G, studying the new access technology to meet a broad range of use cases.

Multiple access schemes, the most fundamental aspect of the physical layer, to a large extent, are considered as the defining technical feature of each wireless communication generation and have continually evolved in each cellular generation

from frequency division multiple access (FDMA), time division multiple access (TDMA) in 1G and 2G to code division multiple access (CDMA) in 3G and orthogonal frequency-division multiple access/single-carrier FDMA (OFDMA/SC-FDMA) for 4G. Facing the stringent demands of diverse scenarios in 5G, e.g., 1000x higher data rates, massive uplink connectivity and low access latency, the traditional pure orthogonal multiple access is not a good option. Some alternative non-orthogonal multiple access schemes have attracted considerable attention and been identified as a crucial technology component in 5G since they can serve multiple users in the same frequency and time resources via code domain multiplexing and/or power domain multiplexing to enhance system access performance. The non-orthogonal multiple access schemes are potentially able to support massive connections, improve spectrum efficiency and also reduce access latency via the grant free access. Currently, some potential alternative multiple access schemes are being actively studied in 3GPP for 5G, including superposition coding based non-orthogonal multiple access (SPC-NOMA) [2], multi user shared access (MUSA) [3], sparse code multiple access (SCMA) [4], pattern division multiple access (PDMA) [5], resource spread multiple access (RSMA) [6], non-orthogonal coded multiple access (NCMA) [7], and interleave-grid multiple access (IGMA) [8].

In this paper, the principles, advantages and application scenarios of different multiple access techniques are discussed

and compared. In addition, we introduce a unified framework that can merge a wide range of multiple access techniques, which helps to minimize the hardware functional module. Based on the unified framework, some initial work on flexible multiple access schemes is also introduced. Finally, the challenges and future directions are discussed.

2 Candidate Non-Orthogonal Multiple Access Solutions

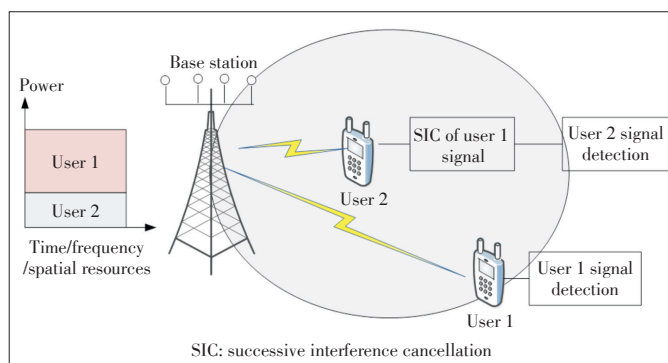
In this section, we introduce the typical candidate NOMA solutions for 5G, which can be basically divided into three categories, i.e., the power domain based, code domain based and interleaver based. Their principles and key features are discussed. At last, we provide their comparison in terms of application scenarios, system performance, receivers, etc.

2.1 Power Domain Based Solutions

2.1.1 SPC-NOMA

NOMA based on superposition coding utilizes power domain for user multiplexing and can be applied for both downlink and uplink. Established by network information theory, non-orthogonal access with successive interference cancellation (SIC)/dirty paper coding (DPC) can achieve the multiuser capacity region both in uplink and downlink. NOMA superposes multiple users in power-domain and exploits channel gain difference between the multiplexed users with the aid of advanced receiver, e.g. the SIC receiver, for user separation. **Fig. 1** shows signal transmission and receiving in downlink NOMA system with two users. Currently the NOMA technique is being discussed in the 3GPP under the study item of “study on downlink multiuser superposition transmission (MUST)” for release 13 [9]. For the study in 3GPP, the study scope of NOMA is very limited, e.g. only about downlink transmission, only for the intra-cell usage and only for data channels.

For 5G system, there are more application scenarios of NOMA technique, such as uplink and control channel, and more advanced NOMA techniques, such as combination with sophisticated multiple-input multiple-output (MIMO) techniques and



▲ Figure 1. Illustration of SPC-NOMA transmission.

inter-cell techniques. In [10]–[12], MIMO NOMA schemes have been studied. Network NOMA which considering the multi-cell scenarios are also studied from EE-SE co-design perspective in [13].

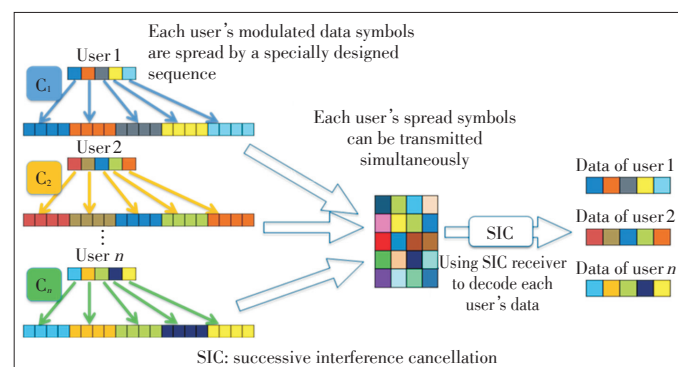
2.2 Code Domain Based Solutions

2.2.1 MUSA

MUSA is a non-orthogonal multiple access scheme operating in code domain. Conceptually, each user's modulated data symbols are spread firstly by a specially designed sequence which facilitates robust SIC implementation compared to the sequences employed by traditional direct-sequence CDMA (DS-SS-CDMA). Then, each user's spread symbols are transmitted concurrently on the same radio resource by means of “Shared Access”, which is essentially a superposition process. Finally, decoding of each user's data from superimposed signal can be performed at the base-station side using SIC technology.

The major processing blocks of MUSA transmitter and receiver are illustrated in **Fig. 2**. Symbols of each user are spread by a spreading sequence. Multiple spreading sequences constitute a pool from which each user can randomly pick one. Note that for the same user, different spreading sequences may also be used to different symbols. This may further improve the performance via interference averaging. Then, all spreading symbols are transmitted over the same time-frequency resources. The spreading sequences should have low cross-correlation and can be non-binary. At the receiver, codeword level SIC is used to separate data from different users. The complexity of codeword level SIC is less of an issue in the uplink as the receiver anyway needs to decode the data for all users. The only noticeable impact on the receiver implementation would be that the pipeline of processing may be changed in order to perform SIC operation.

MUSA relies on a special family of complex spread sequences that can enjoy relatively low cross-correlation even when they are very short, say, 8 or even 4. The real and imaginary parts of the complex spread sequence can be drawn from an M-ary real value set. For example, for a 3-value set $\{-1, 0, 1\}$, ev-



▲ Figure 2. An example of MUSA with four resources shared by multiple users [1].

Unified Framework Towards Flexible Multiple Access Schemes for 5G

SUN Qi, WANG Sen, HAN Shuangfeng, and Chih-Lin I

ery bit of the complex sequence is drawn from the constellation depicted in Fig. 3 with equal probability.

It should be pointed out that the spread sequences used in MUSA are different from the spreading codes, in the sense that MUSA spreading does not have the low density property. Equipped with the well-optimized spreading sequence and state-of-the-art SIC technology, MUSA is capable of decoupling the multiuser mingled data even if those users are contending to access the system. Potentially a large number of devices are allowed to transmit data at their will, by randomly picking spread sequences, spread the data and send them. In other words, MUSA is suitable for the scenario where the uplink transmissions are not tightly scheduled, and the grants for transmission are not signaled per user basis, and with a high overloading. The relaxed UL synchronization requirement for MUSA allows simple derivation of UL time from a DL synchronization process, which can greatly cut down the battery consumption. Lastly, the code domain superposition nature of MUSA can turn the near-far problem into a near-far advantage. The disparity in the received signal to noise ratio (SNR) across the simultaneously transmitting users can be exploited in MUSA to facilitate SIC. Tight transmit power control is no longer needed, which can further lower the device cost and its power consumption.

2.2.2 SCMA

SCMA is a novel non-orthogonal multiple access scheme with sparse codebooks. The main idea of SCMA is to accommodate more users with limited resources and increase the total network throughput, without sacrificing user experience, which can be overloaded to enable massive connectivity and support grant-free access. There are multiple layers in SCMA, which can be used for user multiplexing. Each layer has a predefined codebook, which consists of multiple codewords. The codewords are composed of multi-dimensional complex symbols, and the codewords in the same codebook have the same sparse pattern. For each layer, the coded bits are directly mapped to codewords, which are selected from layer-specific SCMA codebooks. The codewords of different layers are overlaid in code and power domains and carried over shared time-frequency re-

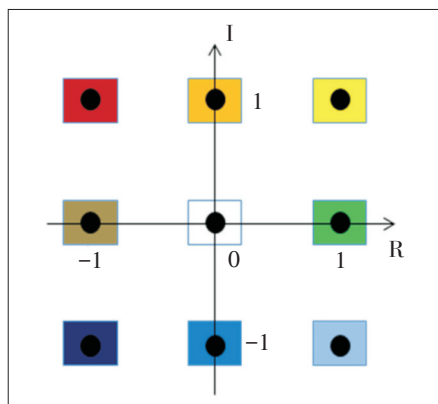


Figure 3. The elements of the complex spreading sequence [3].

sources. Typically, the layer multiplexing may become overloaded if the number of layers is more than the length of the codewords.

Fig. 4 shows an example of bits to codewords mapping in a SCMA system. The codebook design of SCMA has been studied in [14]; it has been shown that with multi-dimensional constellation, shaping and coding gain can be achieved. At the receiver, joint multiuser detection algorithms are needed. Due to the sparsity of the SCMA codeword structure, message passing algorithm (MPA) on factory graph with much lower complexity can be adopted to achieve a suboptimal performance. Some simplified algorithms are proposed [15]–[19] to further reduce the detection complexity.

Besides the codebook and the receiver design, some other challenging issues of SCMA, e.g., the energy efficiency optimization, uplink grant free access, downlink multiuser transmission, and the multi-cell transmission based on SCMA have also been studied. The energy efficiency performance and optimization of SCMA are investigated in [20] and [21]. In [22] and [23], uplink contention based grant-free access based on SCMA has been proposed for 5G radio access. [24] and [25] focus on the downlink multiuser SCMA (MU-SCMA) network. [24] theoretically derives the capacity for downlink Massive MIMO MU-SCMA systems. In [25], a weighted sum rate based user pairing and power sharing algorithm are introduced to the MU-SCMA network. It shows that SCMA can significantly increase the downlink spectral efficiency of 5G wireless cellular networks. Further, SCMA has also been introduced into multi-cell transmission. SCMA based uplink inter-cell interference cancellation technique and open loop joint coordinated multiple point transmission are studied in [26] and [27], respectively.

There are still many challenging issues for SCMA, which need to be solved in the future work. For example, the layer multiplexing in SCMA provides new degree-of-freedom for user scheduling. The algorithms for user grouping and power allocation need to be optimized. In addition, the combination of SCMA and MIMO can be further enhanced.

2.2.3 PDMA

PDMA introduces reasonable diversity between multiple users to promote the capacity, which can obtain higher multiuser

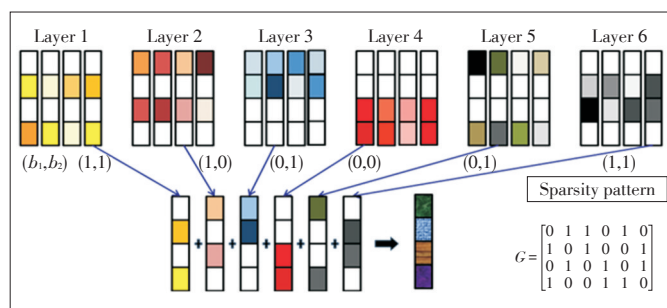


Figure 4. Illustration of SCMA codebooks and the process of bit mapping [1].

multiplexing and diversity gain. It considers the joint design of the transmitter and the receiver based on the optimization point of view for multiuser communication system. At the transmitter side, the non-orthogonal characteristic pattern is used to distinguish users based on the multiple signals domain (including time, frequency and the space domain). At the receiver side, sub-optimal multiuser detection by General SIC based on the features of the user pattern is utilized.

To alleviate the error propagation problem of the SIC receiver, the pattern used in PDMA is generally designed to ensure unequal transmission diversity for each user. In this way, the identical diversity order can be achieved after detection. Inspired by the idea of unequal transmission diversity and sparse coding, an example of pattern and the related resource mapping has been proposed (Fig. 5).

In the example, a code can also be seen as a pattern, which is used to define sparse mapping from data to a group of resources. The code could be represented by a binary vector. The dimension of the vector equals to the number of resources in a group. Each element in the vector corresponds to a resource in a resource group. A “1” means that data shall be mapped to the corresponding resource. Actually, the number of “1” in the code is defined as its transmission diversity order. A code matrix is constructed by all codes sharing on the same resource group. Assuming six users multiplexing on four resource elements (REs). The data for User 1 are mapped to all the four resources in the group, and the data for User 2 are mapped to the first three resources, etc. The order of transmission diversity of the six users is 4, 3, 2, 2, 1, and 1, which is obviously quite different from the SCMA scheme where all the users bear the same transmission diversity.

Generally, if N is the size of resource group (the row number of code matrix), there are $2^N - 1$ possible binary vectors for a code matrix. Assuming K is the column number determined based on overload factor, we can thus choose K patterns out from $2^N - 1$ candidates to construct code matrix. Selection of codes also gives impacts on performance and com-

plexity.

2.2.4 RSMA

RSMA combines the low rate channel code and the scrambling code (and optionally different interleavers) with good correlation properties to separate different transmitters. In RSMA system, all users use the same frequency and time resources to transmit messages to the base station, regardless of the number of concurrent users. In other words, each user’s transmission power can be spread over all the available time and frequency resources.

RSMA can be coupled with various waveforms/modulation schemes depending on the design target. Generally, it includes the single carrier RSMA and the multi-carrier RSMA. The single carrier RSMA is optimized for battery power consumption and link budget extension by using single carrier waveforms. It allows grant-less transmission and potentially allows asynchronous access. The grant-less transmission using RSMA reduces the signaling overhead, while the single carrier waveform further reduces peak-to-average power ratio (PAPR) and achieves higher power amplifier efficiency. The pulse shaping block can further enhance the PAPR (e.g. potentially leading to constant envelope waveform), reducing out-of-band emission simultaneously. The multi-carrier RSMA is optimized for low latency access, where reducing access delay is the design priority. It is suitable for the scenario where a connected state device is already synchronized to the base station and not link budget limited (e.g., close to the base station). Such a device can use RSMA with OFDM-based multi-carrier waveform for grant-less transmission to reduce overall access delay.

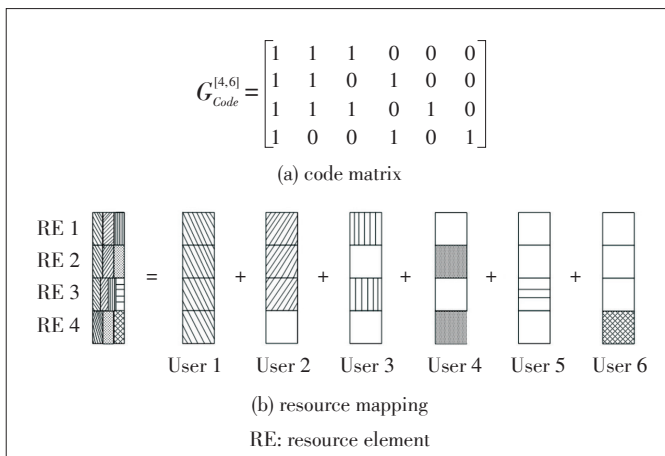
2.2.5 NCMA

NCMA is a multiple access scheme based on the resource spreading by using non-orthogonal codewords, which is composed of the codewords obtained by Grassmannian line packing problem [28]. To minimize the MUI theoretically, the spreading codes are designed with the minimum correlation.

The non-orthogonal codebook is defined by

$$C = [c^1 \cdots c^K] = \begin{bmatrix} c_1^1 & \cdots & c_1^K \\ \vdots & \ddots & \vdots \\ c_N^1 & \cdots & c_N^K \end{bmatrix}, C \in \mathbb{C}^{N \times K}, \text{ where } N \text{ is the spreading factor and } K \text{ is the superposition factor. Then, the codebook design problem can be posed in terms of maximizing the minimum chordal distance between codeword pairs } \min_c \left(\max_{1 \leq k \leq j \leq K} \sqrt{1 - |(c^k)^* \cdot c^j|} \right) \text{ where } (c^k)^* \text{ is the conjugate codeword of } c^k.$$

NCMA can provide the additional throughput or improved connectivity with a small loss of block error rate (BLER) in specific environments, by exploiting additional layers through the superposed symbol, while satisfying QoS constraints. Since the receiver of NCMA system is available for parallel interference



▲ Figure 5. Users sharing on four resource elements [5].

Unified Framework Towards Flexible Multiple Access Schemes for 5G

SUN Qi, WANG Sen, HAN Shuangfeng, and Chih-Lin I

cancellation (PIC), the multiuser detection can be implemented with low complexity. In addition, the MUI level between codeword pairs is always similar due to the correlation characteristics mentioned above. Consequently, NCMA provides the potentials in terms of throughput or connectivity under special scenarios, e.g., huge connections with small packet in mMTC scenarios without changing the transmission block size, or for reducing the collision probability in contention based multiple access.

2.3 Interleaver Based Solution

IGMA is an interleaver-based MA scheme, The typical transmitter system structure using IGMA is shown in Fig. 6. Basically, the IGMA scheme distinguishes different users based on

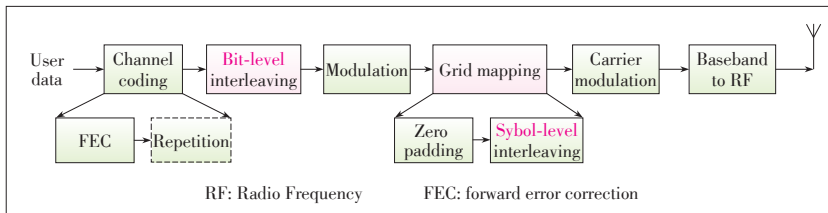


Figure 6. The IGMA transmitter [8].

different bit-level interleavers, different grid mapping patterns and different combinations of bit-level interleaver and grid mapping pattern.

Compared to the need of well-designed codewords or code sequences, the sufficient source of bit-level interleavers and/or grid mapping patterns are able to provide enough scalability for different connection densities, and also provide flexibility to achieve good balance between channel coding gain and benefit from sparse resource mapping. By proper selection, the low correlated bit-level interleavers is achieved. In the grid mapping process, sparse mapping based on zero padding and symbol-level interleaving is introduced, which provides another dimension for user multiplexing. Moreover, the density ρ of the grid mapping pattern is defined as the occupied RE numbers N_{used} dividing the total assigned RE numbers N_{all} , i.e. $\rho = N_{used}/N_{all}$. Different densities could be flexibly configured. It should be noted that the symbol sequence order is randomized after the grid mapping process due to symbol-level interleaving, which may further bring benefit in terms of combating frequency selective fading and inter-cell interference, compared to resource mapping using direct code-matrices/codebooks.

At the receiver side, the low complexity multiuser detector (MUD) and the elementary signal estimator (ESE) that takes advantage of the special property of interleaving can be utilized with a simple de-mapping operation on the top. It should be noted that lower density of the grid mapping pattern further reduces detection complexity of ESE for IGMA. In addition, MAP and MPA detectors are also applicable for IGMA, which can improve the detection performance a lot comparing to ESE

at the cost of complexity. The complexity of MAP/MPA for IGMA probably can be alleviated when sparse grid mapping is used, due to the similar property of LDS.

Fig. 7 shows an example of the grid mapping process of IGMA. The sparse symbol-to-RE mapping is performed based on an assigned grid mapping pattern. An exemplary operation can be mathematically formulated as a process by permutation matrix α_{GM} . According to the symbol-level interleave $\theta_{k,2}$ associated with the grid mapping pattern β_k with density $\rho_k (0 < \rho_k \leq 1)$, the corresponding permutation matrix $\alpha_{GM} \in \mathbb{C}^{N \times L}$ can be obtained. Thus, the k_{th} user's symbol sequence s_k after zero padding and interleaving can be denoted by $s'_k = s_k \times \alpha_{GM} = [s'_{k,1}, s'_{k,2}, \dots, s'_{k,L}]$, where $L = N/\rho_k$ and ρ_k decides the number of zeros padded.

2.4 Summary of Multiple Access Techniques

The pros and cons of the multiple access techniques introduced above are summarized here in Table 1.

It's worth mentioning that some of these non-orthogonal schemes, such as SCMA MUSA and PDMA, can be implemented within a unified framework, and each of them corresponds to a different codebook mapping module. In this way, the air interface can handover between different multiple access schemes in a flexible way, and all the other modules can be reused. This helps to improve the resource utilization and reduce the cost. In the following section, we will provide a unified framework for the multiple access schemes.

3 Unified Framework of Multiple Access Schemes

Fig. 8 shows a unified framework of multiple access

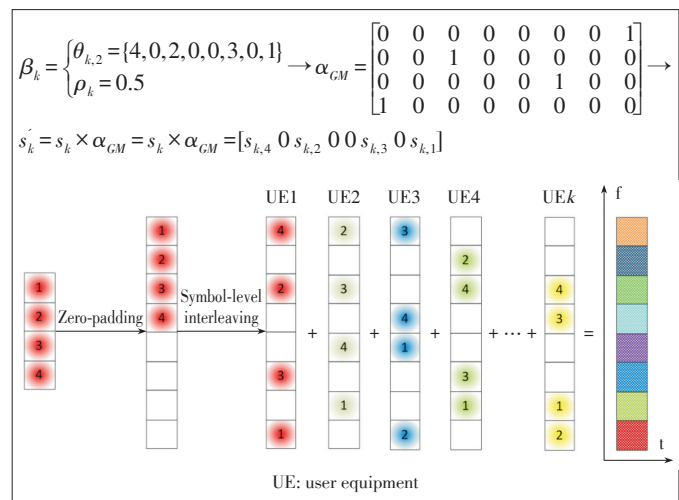


Figure 7. Example of the grid mapping process of IGMA when $N = 4$, $\rho_k = 0.5$ and $L = 8$ [8].

▼ Table 1. Summary of multiple access techniques

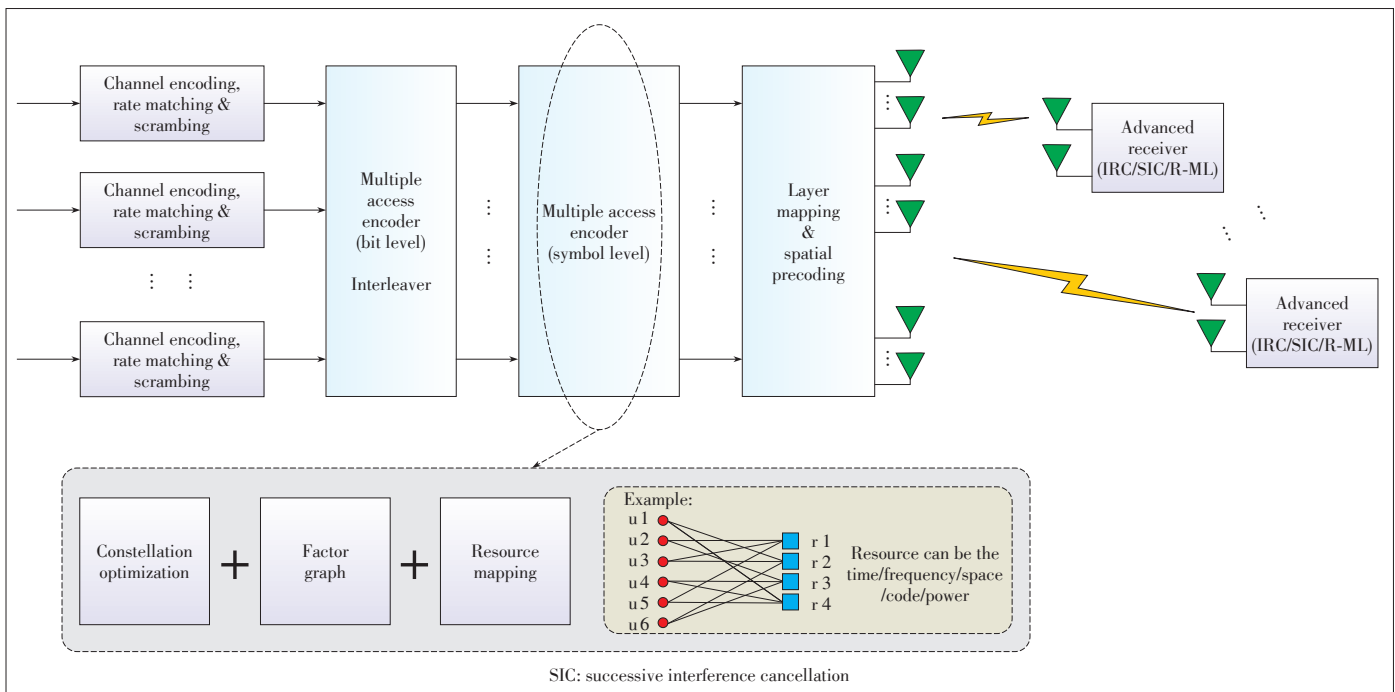
Category	Power domain based		Code domain based				Interleaver based
Scheme	SPC -NOMA	MUSA	SCMA	PDMA	RSMA	NCMA	IGMA
Scenario	DL: eMBB	UL: mMTC, URLLC DL: eMBB	UL: mMTC, URLLC DL: eMBB	UL: mMTC, URLLC DL: eMBB	UL: mMTC, URLLC	UL: eMBB, mMTC, URLLC	UL: eMBB, mMTC, URLLC
Multiplexing domain	Power	Code/Power	Code/Power	Code/Power/Spatial	Code/Power	code	Interleaver
Transmitter Overloading	Medium	High	High	High	High	High	High
Transmitter Spreading	No	Yes	Yes	Yes	Yes	Yes	Yes
Transmitter multi - dimension constellation	No	No	Yes	No	No	No	No
Receiver	SIC	SIC	MPA/SIC	SIC/MPA	SIC	PIC	MAP/MPA ESE MUD
DL: downlink eMBB: enhanced mobile broadband ESE: elementary signal estimator IGMA: interleave-grid multiple access mMTC: massive machine type communication		MPA: message passing algorithm MUD: low complexity multiuser detector MUSA: multi user shared access NCMA: non-orthogonal coded multiple access PDMA: pattern division multiple access		PIC: parallel interference cancellatio RSMA: resource spread multiple access SCMA: sparse code multiple access SIC: successive interference cancellation		SPC-NOMA: superposition coding based non-orthogonal multiple access UL: uplink URLLC: ultra-reliable low latency machine type communication	

schemes. The differences among these multiple access schemes lie in the different realization of interleaver, constellation optimization, factor graph and multiplexing domain. The detailed explanations are listed in Table 2.

4 Flexible Multiple Access in 5G

The above discussed advanced multiple access schemes as well as the traditional orthogonal multiple access scheme, e.g. OFDMA are all identified as potential candidates for 5G. There is no individual scheme can fulfill the requirements of all applications and scenarios in 5G system. A flexible adaptation of these multiple access schemes is needed to support the diverse deployment scenarios and traffic requirements. For ex-

ample, in the case of massive connections, how to accommodate more users with limited resources has become a critical problem for next generation access network. With non-orthogonal multiple access schemes, e.g., SCMA, MUSA, PDMA and RSMA, the same resources are shared and reused by multiple users, thus the number of connections increases. To support the traffic with low latency requirement, non-orthogonal multiple access schemes help to realize grant-free multiple access, with which the latency is much lower, and the power consumption of the devices can be reduced. In other scenarios, such as downlink machine type traffic, the simple orthogonal multiple access schemes are better due to the device cost and implementation complexity. OFDMA can be utilized for the cell-center user with high data rate transmission applications.



▲ Figure 8. Unified framework of multiple access schemes.

Unified Framework Towards Flexible Multiple Access Schemes for 5G

SUN Qi, WANG Sen, HAN Shuangfeng, and Chih-Lin I

Table 2. Configuration methods of different multiple access schemes based on a unified framework

	Interleaver	Constellation mapping	Factor graph	Resource mapping (multiplexing domain)
OMA	Identity matrix	Gray-mapped legacy constellation	Identity matrix	Time/frequency/code/space
SPC -NOMA	MUST Cat 1 [9]	Identity matrix	non-Gray-mapped superposed constellation	Power
	MUST Cat 2 [9]	Constraint permutation matrix	Gray-mapped superposed constellation	Power/bit
	MUST Cat 3 [9]	Permutation matrix	Gray-mapped legacy constellation	Bit
MUSA (uplink)	Identity matrix	Legacy constellation	Matrix composed of low cross-correlation and non-binary spreading sequence	Code
SCMA	Identity matrix	Joint optimization (multi -dimensional modulation + Sparse matrix ³)	Sparse matrix	Code/power
PDMA	Identity matrix	Legacy modulation	Sparse matrix with unequal diversity order	Code/power/space
RSMA	Optional	Legacy modulation	Matrix composed of scrambling code with good correlation properties	Code
NCMA	Identity matrix	legacy modulation	Matrix obtained by Grassmannian line packing problem	code
IGMA	Permutation matrix	legacy modulation	Sparse matrix	code

IGMA: interleave-grid multiple access
 MUSA: multi user shared access
 MUST: Downlink Multiuser Superposition Transmission
 NCMA: non-orthogonal coded multiple access
 OMA: orthogonal multiple access
 PDMA: pattern division multiple access
 RSMA: resource spread multiple access
 SCMA: sparse code multiple access
 SPC-NOMA: superposition coding based non-orthogonal multiple access

Besides, the multiple access scheme should also be properly selected, taking the tradeoff of multiple conflicting objectives into account, e.g., complexity vs. performance, energy efficiency (EE) vs. spectral efficiency (SE) and coverage. In addition, because the channel conditions and service load may also dynamically vary, the multiple access schemes and their related parameters such as the number of codewords, length of codeword, spreading factor, max number of layers, need to be optimized based on the instant services and the link conditions. In the following, we provide two potential adaptive multiple access schemes in 5G.

In [21], the adaptive multiple access scheme is studied from EE-SE co-design perspective, taking the detection complexity into consideration. The SCMA and OFDMA schemes are taken as the candidate uplink multiple access schemes in the study. The problem is formulated to choose the optimal multiple access scheme and the related parameters simultaneously to maximize the EE under the total transmit power constraint, the quality of service (QoS) constraints and other specific requirements. The considered power consumption includes the transmit power consumption, the static circuit power consumption, and the SCMA decoding power consumption related which is proportional to the SCMA decoding complexity order $O(M^{d_f})$, where M is the constellation size, $d_f = \frac{N}{K}J$, K and N denotes the codeword length and non-zero entries in each codeword in SCMA, respectively, and $J = \binom{K}{N}$ is the maximum number of access users. Fig. 9 shows the EE performance comparison of SCMA, OFDMA and the proposed link adaptation schemes with various cell radiuses. When the cell radius is small, the SCMA scheme has better EE performance; when the cell radius is large, the OFDMA scheme performs better than SCMA

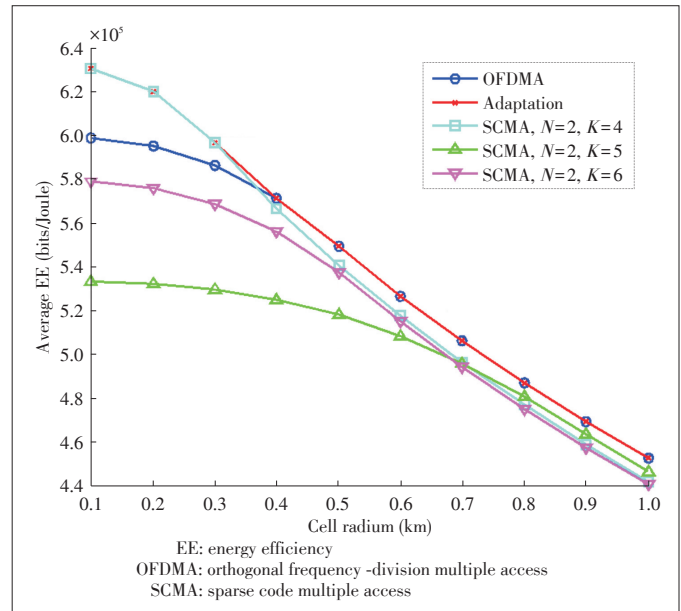


Figure 9. Average EE v.s. Cell Radiuses.

scheme. The reason is that the SCMA can access more users than OFDMA, and the increment of the number of access users per resource can improve the system EE when the cell radius is small since the user transmit power efficiency is large when the path loss is small. When the cell radius increases, the user transmit power efficiency decreases and the increment of the number of access users per resource will decrease the system EE. The adaptation scheme can obtain the overall good EE performance for all the cell radiuses (Fig. 9).

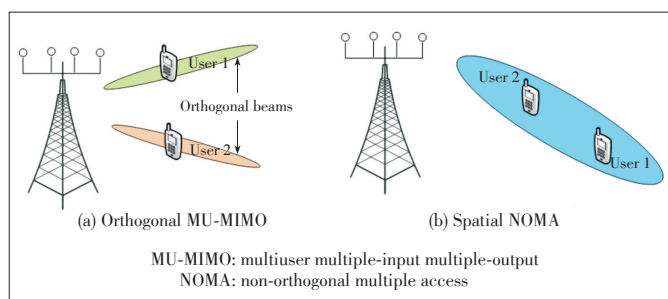
Another example of the adaptive multiple access is between the spatial NOMA (also known as MIMO NOMA) scheme and orthogonal the multiuser MIMO (MU - MIMO). Fig. 10 shows

the concept of the orthogonal MU-MIMO and the spatial NOMA. In spatial NOMA, each two users can be served via one beam, and the interference between these two users will be large but can be cancelled via SIC decoding at the stronger user. This is unlike the orthogonal MU-MIMO precoding (e.g. zero-forcing MU-MIMO), in which there is no interference between users or it is usually small. Owing to this feature, when the channel has large correlations, spatial NOMA will have significantly higher throughput over orthogonal MU-MIMO since the MU-MIMO precoding needs to reduce the interference between users and will suffer large gain loss in that case. In addition, when the user channel gain difference is large, the spatial NOMA will also have better performance compared to orthogonal MU-MIMO due to the near-far effects. While in the high SNR regimes with low transmit correlation, the orthogonal MU-MIMO is preferred since it can approach the capacity bound of MIMO broadcast channel in high SNR regimes. Considering the time varying fading characteristics of MIMO channels and the random distribution feature of active users, in a multiuser scenario, the adaptation between orthogonal MU-MIMO and spatial NOMA is desired for both the cell average and cell edge throughput enhancement.

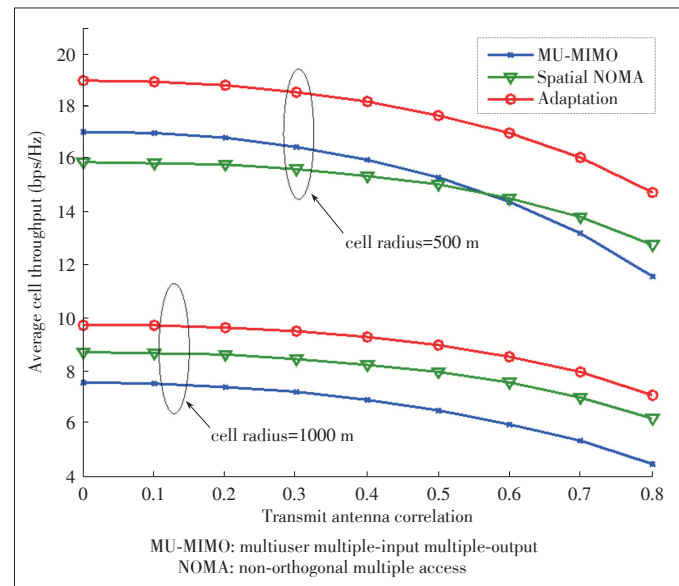
Fig. 11 shows the gain of the adaptation between orthogonal MU-MIMO based on zero-forcing scheme and spatial NOMA. When the transmit antenna correlation or the cell radius grows large, the spatial NOMA will have better performance than zero-forcing based MU-MIMO, and the adaptation gain will increase. The reason is that the higher transmit antenna correlation will lead to the higher probability that the user channels have high correlation, and the larger cell radius will lead to the higher probability that the user channels have large gain difference.

5 Conclusions

All the typical candidate NOMA solutions for 5G have different strength points and weakness points. None of them can surpass other schemes on all aspects. To fully exploit the advantages of these candidate technologies and traditional orthogonal multiple access solutions, a unified framework and a flexible multiple access schemes are required. Flexible switch among different NOMA schemes and the orthogonal multiple



▲ **Figure 10. Orthogonal MU-MIMO and Spatial NOMA.**



▲ **Figure 11. Average cell throughput comparisons of various multiuser MIMO schemes (32 antenna base station at 6 GHz with four transceiver chains and four single antenna users)**

access technologies is expected to efficiently enhance the data rate and accommodate the necessary scalability for massive IoT connectivity and drastic reduction in access latency, and then to fully meet the diversified needs of 5G wireless communication systems. Some challenging problems need to be solved before NOMA schemes are put into use in 5G. In future, the impact of these candidate schemes on the existing systems, e.g. the grant free access procedure, reference signal, channel estimation and network assisted signaling, need to be carefully designed. What's more, the performance tradeoff of the code mapping manners in these schemes and their implementation complexity may need further evaluated. The adaptive mechanisms for part of these candidate schemes are also worth further study to meet the diversified requirements of 5G.

References

- [1] FuTURE Mobile Communication Forum 5G SIG. (Nov. 2014). Rethink mobile communications for 2020+ [Online]. Available: <http://www.future-forum.org/dl/141106/whitepaper.zip>
- [2] A. Benjebbovu, Y. Saito, Y. Kishiyama, et al., "Concept and practical considerations of non-orthogonal multiple access for future radio access," in *IEEE ISPACS 2013*, Naha, Japan, Nov. 2014, pp. 770–774. doi: 10.1109/ISPACS.2013.6704653.
- [3] 3GPP, "Discussion on multiple access for new radio interface," TSG-RAN WG1 #84bis, Busan Korea, R1-162226, Apr. 11–15, 2016.
- [4] H. Nikopour and H. Baligh, "Sparse code multiple access," in *IEEE PIMRC 2013*, London, England, Sept. 2013, pp. 332–336. doi: 10.1109/PIMRC.2013.6666156.

Unified Framework Towards Flexible Multiple Access Schemes for 5G

SUN Qi, WANG Sen, HAN Shuangfeng, and Chih-Lin I

- [5] 3GPP, "Candidate Solution for New Multiple Access", TSG -RAN WG1 #84bis, Busan, Korea, R1 -163383, Apr. 11-15, 2016.
- [6] 3GPP, "Candidate NR multiple access schemes", TSG -RAN WG1 #84bis, Busan Korea, R1 -163510, Apr. 11-15, 2016.
- [7] 3GPP, "Considerations on DL/UL multiple access for NR," TSG -RAN WG1 #84bis, Busan Korea, R1 -162517, Apr. 11-15, 2016.
- [8] 3GPP, "Considerations on DL/UL multiple access for NR," TSG -RAN WG1 #84bis, Busan Korea, R1 -163992, Apr. 11-15, 2016.
- [9] *Study on Downlink Multiuser Superposition Transmission (MUST) for LTE*, 3GPP TR36.859 v13.0.0, Jan. 2016.
- [10] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non -orthogonal multiple access," *IEEE Transactions on Wireless Communications.*, to be published.
- [11] Q. Sun, S. Han, C. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Communications Letters*, vol. 4, pp. 405-408, Aug 2015, doi: 10.1109/LWC.2015.2426709
- [12] Q. Sun, Chih -Lin I, S. Han, et al., "Sum rate optimization for MIMO non -orthogonal multiple access systems," accepted by *IEEE WCNC*, 2015.
- [13] S. Han, Chin -Lin I, Z. Xu, et al., "Energy efficiency and spectrum efficiency co -design: from NOMA to network NOMA," *IEEE MMTC E -Letter (Special Issue on 5G)*, vol. 9, no. 5, pp. 21-24, Sept. 2014.
- [14] M. Taherzadeh, H. Nikopour, A. Bayesteh, et al., "SCMA codebook design," in *IEEE VTC Fall 2014*, Vancouver, Canada, Sept. 2013, pp. 1-5. doi: 10.1109/VTCFall.2014.6966170.
- [15] K. Xiao, B. Xiao, S. Zhang, et al., "Simplified multiuser detection for SCMA with sum - product algorithm," in *IEEE WCSP 2015*, Nanjing, China, Oct. 2015, pp. 1-5. doi: 10.1109/WCSP.2015.7341328.
- [16] Y. Wu, S. Zhang, Y. Chen, et al., "Iterative multiuser receiver in sparse code multiple access systems," in *Proc. IEEE ICC 2015*, London, England, Jun. 2015, pp. 2918-2923. doi: 10.1109/ICC.2015.7248770.
- [17] A. Bayesteh, H. Nikopour, M. Taherzadeh, et al., "Low complexity techniques for SCMA detection," in *Proc. IEEE Globecom Workshops 2015*, San Diego, USA, Dec. 2015, pp. 1-6. doi: 10.1109/GLOCOMW.2015.7414184.
- [18] H. Mu, Z. Ma, M. Alhaji, et al., "A fixed low complexity message pass algorithm detector for uplink SCMA system," *IEEE Wireless Communications Letters*, vol. 4, no. 6, pp. 585-588, 2015. doi: 10.1109/LWC.2015.2469668.
- [19] Y. Du, B. Dong, Z. Chen, et al., "A fast convergence multiuser detection scheme for uplink SCMA systems," *IEEE Wireless Communications Letters*, to be published.
- [20] S. Zhang, X. Xu, L. Lu, et al., "Sparse code multiple access: an energy efficient uplink approach for 5G wireless systems," in *Proc. IEEE Globecom 2014*, Austin, USA, Dec. 2014, pp. 4782-4787. doi: 10.1109/GLOCOM.2014.7037563.
- [21] Q. Sun, Chih -Lin I, S. Han, et al., "Software defined air interface: a framework of 5G air interface," in *Proc. IEEE WCNC Workshop 2015*, New Orleans, USA, Mar. 2015, pp. 6-11. doi: 10.1109/WCNCW.2015.7122520.
- [22] K. Au, L. Zhang, H. Nikopour, et al., "Uplink contention based SCMA for 5G radio access," in *Proc. IEEE Globecom Workshop 2014*, Austin, USA, Dec. 2014, pp. 900-905. doi: 10.1109/GLOCOMW.2014.7063547.
- [23] A. Bayesteh, E. Yi, H. Nikopour, et al., "Blind detection of SCMA for uplink Grant -Free Multiple Access," in *Proc. IEEE ISWCS 2014*, Barcelona, Spain, Aug. 2014, pp. 853-857. doi: 10.1109/ISWCS.2014.6933472.
- [24] T. Liu, X. Li, and L. Qiu, "Capacity for downlink massive MIMO MU -SCMA system," in *Proc. IEEE WCSP 2015*, Nanjing, China, Oct. 2015, pp. 1-5. doi: 10.1109/WCSP.2015.7341100.
- [25] H. Nikopour, E. Yi, A. Bayesteh, et al., "SCMA for downlink multiple access of 5G wireless networks," in *Proc. IEEE Globalcom 2014*, Austin, USA, Dec. 2014, pp. 3940-3945. doi: 10.1109/GLOCOM.2014.7037423.
- [26] U. Vilaipornsawai, H. Nikopour, and A. Bayesteh, "SCMA for open -loop joint transmission CoMP," in *Proc. IEEE VTC Fall 2015*, Boston, USA, Sept. 2015, pp. 1-5. doi: 10.1109/VTCFall.2015.7391126.
- [27] Y. Li, X. Lei, P. Fan, et al., "An SCMA -based uplink inter -cell interference cancellation technique for 5G wireless systems," in *Proc. IEEE WCSP 2015*, Nanjing, China, Oct. 2015, pp. 1-5, doi: 10.1109/WCSP.2015.7341306
- [28] A. Medra and T. N. Davidson, "Flexible codebook design for limited feedback systems via sequential smooth optimization on the grassmannian manifold," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1305-1318, Mar. 2014. doi:10.1109/TSP.2014.2301137.

Manuscript received: 2016-06-30

Biographies

SUN Qi (sunqiyj@chinamobile.com) received the BSE and PhD degree in information and communication engineering from Beijing University of Posts and Telecommunications, China in 2009 and 2014, respectively. After graduation, she joined the Green Communication Research Center of the China Mobile Research Institute. Her research focuses on 5G key technologies, including non -orthogonal multiple access, new waveforms, flexible duplex and UDN.

WANG Sen (wangsenyj@chinamobile.com) received the PhD degree in information and communication engineering from Beijing University of Posts and Telecommunications, China in 2013. He joined the Green Communication Research Center of the China Mobile Research Institute after graduation. He is now working on the 5G key technologies and standardization. His research interests include massive MIMO, non -orthogonal multiple access, new waveforms and system level evaluation.

HAN Shuangfeng (hanshuangfeng@chinamobile.com) received his MS and PhD degrees in electrical engineering from Tsinghua University, China in 2002 and 2006 respectively. He joined Samsung Electronics as a senior engineer in 2006 working on MIMO, MultiBS MIMO etc. From 2012, he is a senior project manager in the Green Communication Research Center of the China Mobile Research Institute. His research interests are green 5G, massive MIMO, full duplex, NOMA and EE -SE codesign.

Chih-Lin I (icl@chinamobile.com) received her PhD degree in electrical engineering from Stanford University, USA. She has been working at multiple world -class companies and research institutes leading the R&D, including AT&T Bell Labs; Director of AT&T HQ, Director of ITRI Taiwan, and VPGD of ASTRI Hong Kong. She received the IEEE Trans. COM Stephen Rice Best Paper Award and is a winner of the CCCP National 1000 Talent Program. In 2011, she joined China Mobile as its Chief Scientist of wireless technologies, established the Green Communications Research Center, and launched the 5G Key Technologies R&D. She is spearheading major initiatives including 5G, C -RAN, high energy efficiency system architectures, technologies and devices; and green energy. She was an elected Board Member of IEEE ComSoc, Chair of the ComSoc Meetings and Conferences Board, and Founding Chair of the IEEE WCNC Steering Committee. She is currently an Executive Board Member of GreenTouch, a Network Operator Council Member of ETSI NFV, a Steering Board Member of WWRF, and a Scientific Advisory Board Member of Singapore NRF. Her current research interests center around "Green, Soft, and Open"