Non-Orthogonal Multiple Access (NOMA) for Future Radio Access


*Radio Access Network Development Department, NTT DOCOMO, INC.
**DOCOMO Beijing Communications Laboratories Co., Ltd.
***Graduate School of Science and Technology, Tokyo University of Science
E-mail: yuuya.saitou.fa@nttdocomo.com

Abstract—This paper presents a non-orthogonal multiple access (NOMA) concept for future radio access (FRA) towards the 2020s information society. Different from the current LTE radio access scheme (until Release 11), NOMA superposes multiple users in the power domain although its basic signal waveform could be based on the orthogonal frequency division multiple access (OFDMA) or the discrete Fourier transform (DFT)-spread OFDM the same as LTE baseline. In our concept, NOMA adopts a successive interference cancellation (SIC) receiver as the baseline receiver scheme for robust multiple access, considering the expected evolution of device processing capabilities in the future. Based on system-level evaluations, we show that the downlink NOMA with SIC improves both the capacity and cell-edge user throughput performance irrespective of the availability of the frequency-selective channel quality indicator (CQI) on the base station side. Furthermore, we discuss possible extensions of NOMA by jointly applying multi-antenna/site technologies with a proposed NOMA/MIMO scheme using SIC and an interference rejection combining (IRC) receiver to achieve further capacity gains, e.g., a three-fold gain in the spectrum efficiency representing a challenging target for FRA.

Keywords — NOMA, non-orthogonal, interference cancellation, multiple access, future radio access

I. INTRODUCTION

Radio access technologies for cellular mobile communications are typically characterized by multiple access schemes, e.g., frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA), and OFDMA. In the 3.9 and 4th generation (4G) mobile communication systems such as Long Term Evolution (LTE) [1] and LTE-Advanced [2, 3], standardized by the 3rd Generation Partnership Project (3GPP), orthogonal multiple access based on OFDMA or single carrier (SC)-FDMA is adopted. Orthogonal multiple access was a reasonable choice for achieving good system-level throughput performance in packet-domain services with simple single-user detection. However, considering future radio access (FRA) in the 2020s, further enhancement to achieve significant gains in capacity and system throughput performance is a high priority requirement in view of the recent exponential increase in the volume of mobile traffic, e.g., beyond a 500 fold increase in the next decade [4], and the need for enhanced delay-sensitive high-volume services such as video streaming and cloud computing. Thus, the 3GPP recently has initiated discussions on further evolution of LTE towards the future, i.e., Release 12 and onwards. In order to continue to ensure the sustainability of 3GPP radio access technologies over the coming decade, new solutions must be identified and provided that can respond to future challenges [4, 5]. Also, recent trends in research activity for the next generation of mobile and wireless communication systems for 2020 and beyond have emerged such as the Mobile and wireless communications Enablers for the 2020 Information Society (METIS) project [6].

To accommodate such demands, a combination of multiple approaches, i.e., technologies for spectrum efficiency enhancement, spectrum extension with efficient use of higher frequency bands, and network densification deploying small cells would be required. In this sense, innovative radio access technologies to enhance significantly the spectrum efficiency and the small-cell enhancement now on-going in the 3GPP are very important [7]. Although it may be very challenging, our target for FRA toward 2020s is to achieve a further 3-fold enhancement in the spectrum efficiency compared to the LTE baseline. Since LTE has already achieved a 3-4 fold enhancement in the spectrum efficiency compared to 3G High-Speed Packet Access (HSPA), our target gain would be more than 10 fold compared to the 3G HSPA. Thus, for instance, the gain in the total capacity such as beyond 500 fold can be achieved if other ways, i.e., spectrum extension and network densification achieve a capacity gain of 50 fold.

To enhance the spectrum efficiency, we have proposed a non-orthogonal multiple access (NOMA) scheme using interference cancellation, which is a promising candidate as a new multiple access scheme for FRA [8-12]. Here, FRA may include future LTE enhancements, i.e., beyond Release 12, and potential new systems beyond 4G. In the paper, we present our NOMA concept for FRA with some technical considerations and proposals with results from system-level evaluations. The remainder of the paper is organized as follows. Section II presents our motivations in proposing the NOMA concept. Section III describes the principle behind the basic NOMA scheme using a successive interference cancellation (SIC) receiver and its power allocation scheme with some system-level evaluations. In Section IV, we discuss possible extensions of NOMA by applying the multi-antenna/site technologies with a proposed NOMA/Multiple-Input Multiple-Output (MIMO) scheme [10] to achieve a further capacity gain. Finally, Section V concludes the paper.

II. MOTIVATIONS FOR NOMA CONCEPT

Table I shows the multiple access schemes for 3G (W-CDMA/HSPA), 3.9/4G (LTE/LTE-Advanced), and our expected FRA for the 2020s. In 3G, non-orthogonal user multiplexing based on direct sequence (DS)-CDMA is used. The receiver uses simple single-user detection such as the Rake receiver. Fast transmission power control (TPC) is adopted to address the well-known near-far problem in cellular deployments due to its non-orthogonal property. In 3.9/4G, orthogonal user multiplexing based on OFDMA, or SC-FDMA is adopted to achieve higher throughput performance in packet-domain services. Its signal waveform such as orthogonal frequency division multiplexing (OFDM), including discrete Fourier transformation (DFT)-spread OFDM [1], provides important benefits, i.e., robustness against multipath interference and good affinity to MIMO technologies, that dramatically increase the achievable data rate. Furthermore, fast TPC becomes non-essential because there is no near-far problem in orthogonal user multiplexing. Adaptive...
TABLE I. MULTIPLE ACCESS FOR 3G, 3.9/4G, AND FRA (EXPECTED)

<table>
<thead>
<tr>
<th>User multiplexing</th>
<th>3G</th>
<th>3.9/4G</th>
<th>FRA (our view)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-orthogonal (CDMA)</td>
<td>Orthogonal (OFDM)</td>
<td>Non-orthogonal with SIC (NOMA)</td>
<td></td>
</tr>
<tr>
<td>Signal waveform</td>
<td>Single-carrier OFDM (or DFT-s-OFDM)</td>
<td>OFDM (or DFT-s-OFDM)</td>
<td></td>
</tr>
<tr>
<td>Link adaptation</td>
<td>Fast TPC</td>
<td>AMC</td>
<td>AMC + Power allocation</td>
</tr>
<tr>
<td>Image</td>
<td>Non-orthogonal assisted by power control</td>
<td>Orthogonal between users</td>
<td>Superposition &amp; power allocation</td>
</tr>
</tbody>
</table>

modulation and coding (AMC) is introduced as a basic link adaptation scheme assuming static or slow TPC, in which one-cell frequency reuse for cellular network operation is supported similarly as it is supported in 3G systems by adaptively controlling the channel coding rate and repetition/spreading gain.

In FRA, we expect that non-orthogonal user multiplexing, i.e., NOMA, can again be a promising candidate as a multiple access scheme according to the following motivations.

- **Evolution of device processing capabilities**
  To make NOMA promising, it should be used with advanced transmission/reception techniques such as dirty paper coding (DPC) or a SIC receiver [13–15], which is different from the CDMA in 3G systems. In our view, NOMA with a SIC receiver is the baseline scheme as mentioned later. The fact that NOMA requires a SIC receiver may be a drawback in terms of receiver complexity. This is because the SIC receiver requires demodulation and decoding for other sets of user equipment (UEs) in addition to those for its own UE, which may increase the processing delay. Thus, the feasibility of the NOMA will highly depend on the evolution of device processing capabilities expected in the 2020s. However, such evolution has been typically observed in existing developed technologies such as channel coding technologies, e.g., the Turbo decoder.

- **Utilization of additional domain for user multiplexing**
  For FRA, good properties of 3.9/4G, i.e., LTE/LTE-Advanced, should be maintained as much as possible such as robustness against multipath interference, good affinity to MIMO technologies, and supportability for one-cell frequency reuse. Thus, we assume that the basic signal waveform for NOMA could be based on OFDM or DFT-spread OFDM as well as LTE radio access although alternative waveforms such as a filter bank multicarrier (FBMC) [16] could be investigated. However, different from the current LTE radio access scheme (until Release 11), NOMA superposes multiple users in the power domain (forming a superposition coding) so that its user separation is achieved through SIC and capacity-achieving channel codes such as the Turbo code and low-density parity check (LDPC) code. In this sense, NOMA is a scheme that utilizes an additional new domain, i.e., the power domain, which is not sufficiently utilized in 3.9/4G systems. The performance gain of NOMA compared to OFDMA increases when the difference in channel gains, i.e., path loss between UEs, is large [8].

- **Robust performance gain in practical wide area deployments**
  In LTE/LTE-Advanced, many technologies are adopted to enhance the system performance. Among them, there are some basic technologies such as OFDM, hybrid automatic repeat request (HARQ), Turbo coding, and open-loop MIMO, which can provide a robust performance gain irrespective of the UE mobility or feedback/processing latency. In other words, these basic technologies do not rely so much on the transmitter knowledge of instantaneous frequency-selective fading channels such as the frequency-selective channel quality indicator (CQI) or channel state information (CSI) that require fine feedback signaling from the UE. In practical cellular deployments, often UE feedback signaling cannot follow the variation in frequency-selective instantaneous fading due to mobility, feedback/processing latency, limited implementation capability of the base station (BS) scheduler, uplink coverage limitation for UE feedback, and so on. This becomes more challenging in higher/wider frequency bands. In this sense, it is preferable for NOMA to take a receiver cancellation approach and to specify the basic cancellation mechanism so that NOMA with a SIC receiver can be a basic technology that provides a robust performance gain in practical wide area deployments.

III. BASIC NOMA WITH SIC

In this section, we discuss basic NOMA with SIC. In the paper, we focus on the NOMA downlink applying SIC for UE receivers although NOMA can also be applied to the uplink with SIC for a BS receiver [8].

A. **Principle**

![Fig. 1. Basic NOMA scheme applying SIC for UE receivers in downlink.](image)

The basic NOMA scheme applying SIC for UE receivers in the cellular downlink. For simplicity, we assume a two-UE case, a single transmitter, and a single receiver antenna. The overall system transmission bandwidth is assumed to be 1 Hz. The base station transmits a signal for UE-i (i = 1, 2), x_i, where E[x_i^2] = 1, with transmission power P_i. The sum of P_i is restricted to P at maximum. In the NOMA, x_1 and x_2 are superposition coded as

\[ x = \sqrt{P} x_1 + \sqrt{P} x_2. \]  

(1)

The received signal at UE-i is represented as

\[ y_i = h_i x + w_i, \]  

(2)

where h_i is the complex channel coefficient between UE-i and the base station. Term w_i denotes the receiver Gaussian noise including inter-cell interference. The power density of w_i is N_0/2.

In the NOMA downlink, the SIC process is implemented at the UE receiver. The optimal order for decoding is in the order of the increasing channel gain normalized by the noise and inter-cell interference power, |h_i|^2/N_0. Based on this order, any user can correctly decode the signals of other users whose decoding order comes before that user for interference cancellation. Thus, UE-i can remove the inter-user interference from the j-th user whose |h_j|^2/N_0 is lower than |h_i|^2/N_0. In a 2-UE case, assuming that |h_1|^2/N_0 < |h_2|^2/N_0, UE-2 does not perform interference cancellation since it comes first in the decoding order. UE-2 first decodes x_2 and subtracts its component from received signal y_1. Therefore, UE-2 can decode x_1 without interference from x_2. The throughput of UE-i, R_i, is represented as

\[ R_i = \log_2 \left( 1 + \frac{P_i |h_i|^2}{N_0/2} \right). \]

(3)

From (3), it can be seen that power allocation for each UE greatly affects the user throughput performance and thus the modulation and coding scheme (MCS) used for data transmission of each UE. By adjusting the power allocation ratio, P_i/P_2, the BS can flexibly control the throughput of each UE. Clearly, the overall cell throughput, cell-edge throughput, and user fairness are closely related to the power allocation scheme; therefore, a flexible radio interface is needed to utilize the potential gain of the NOMA.
When we assume OFDMA with orthogonal user multiplexing, where the bandwidth of $\alpha$ (0 < $\alpha$ < 1) Hz is assigned to UE 1 and the remaining bandwidth, 1−$\alpha$ Hz, is assigned to UE 2, $R_i$ is represented as

$$R_i = \alpha \log_2 \left( 1 + \frac{P_i h_i^2}{\alpha N_0} \right),$$

$$R_i = (1-\alpha) \log_2 \left( 1 + \frac{P_i h_i^2}{(1-\alpha) N_0} \right).$$

(4)

In NOMA, the performance gain compared to OFDMA increases when the difference in channel gains, i.e., path loss between UEs, is large. For example, as shown in Fig. 2, we assume a 2-UE case with a cell-interior UE and a cell-edge UE, where $P_1 h_1^2 / N_0$ and $P_2 h_2^2 / N_0$ are set to 20 and 0 dB, respectively. In OFDMA, when equal bandwidth and equal transmission power are allocated to each UE ($\alpha = 0.5$, $P_1 = P_2 = 1/2P$) assuming the proportional fairness (PF) criteria, the user rates are calculated according to (4) as $R_1 = 3.33$ and $R_2 = 0.50$ bps, respectively. On the other hand, in NOMA, when the power allocation is conducted as $P_1 = 1/5P$ and $P_2 = 4/5P$, the user rates are calculated according to (3) as $R_1 = 4.39$ and $R_2 = 0.74$ bps, respectively. The corresponding gains of NOMA from OFDMA are 32% and 48% for UE 1 and UE 2, respectively.

To investigate further the performance gain of NOMA considering practical assumptions such as CQI feedback, AMC, and HARQ, a multi-cell system-level simulation is conducted. The major simulation parameters based on the existing LTE/LTE-Advanced specifications [3] are utilized. We employed a 19-hexagonal macrocell model with 3 sectors per cell. The cell radius of the macrocells is set to 289 meters. The locations of the UEs are assigned randomly with a uniform distribution. In the propagation model, we take into account distance-dependent path loss with the decay factor of 3.76, lognormal shadowing with the standard deviation of 8 dB and instantaneous multipath fading. The shadowing correlation between the cells (sectors) is set to 0.5 (1.0). The 6-ray typical urban (TU) channel model is assumed. The maximum Doppler frequency, $f_D$, is set to 5.55 Hz, which corresponds to 3 km/h at the carrier frequency of 2 GHz. The transmission power of the macrocells is 46 dBm. The antenna gain at the macrocell and UE is 14 dBi and 0 dBi, respectively. One-antenna transmission and two-antenna reception are assumed and a full buffer traffic model is used. The feedback delay is modeled such that the CQI is not available for scheduling until 4 subframes after the periodic report with a 2-ms interval. Outer-loop link adaptation (OLLA) based on acknowledgement (ACK)/negative ACK (NACK) feedback signaling [22] is applied to compensate for the fluctuation in the scheduling signal-to-interference plus noise power ratio (SINR). In NOMA, power allocation is utilized to determine the power assignment ratio among UEs for a given UE set so that a combination with a scheduling algorithm is needed in order to select the best UE set among all possible UE sets. In order to investigate the performance gain of NOMA, the cell throughput and cell-edge user throughput are evaluated based on the following definitions. The cell throughput is defined as the cell throughput for one macrocell while the cell-edge user throughput is defined as the 5% value of the cumulative distribution function (CDF) of the user throughput.

First, the performance gain of NOMA using the wideband scheduling and power allocation is evaluated. The reason why we evaluate the wideband scheduling case is that the system performance not relying on the frequency-selective channel information is important for the practical wide area deployments. Fig. 3 shows the CDF of the user throughput for OFDMA and NOMA with SIC. For NOMA, it is assumed that the fractional transmission power allocation (FTPA) described in [12] is used. The figure shows that the user throughput performance of NOMA is improved compared to that of OFDMA. An approximate gain of 30% from NOMA in both the cell throughput and cell-edge user throughput is obtained.

Second, the performance gain of NOMA for the frequency-selective scheduling and power allocation is evaluated. To do this, the system bandwidth is divided into multiple subbands. In each subband, the BS transmitter performs downlink transmission to multiple UEs simultaneously with different transmission powers for different UEs. Fig. 4 shows a performance comparison between OFDMA and NOMA with different numbers of UEs per cell. Here, two power allocation schemes are simulated for NOMA. One is FTPA and the other is a tree-search based transmission power allocation (TTPA) [9]. In the simulation, the number of subbands is set to eight. In the OFDMA, only one UE is scheduled in each subband. As a result, NOMA achieves better performance than that for OFDMA. The performance gains in the overall cell throughput are shown in Table 1.

<table>
<thead>
<tr>
<th>User throughput (Mbps)</th>
<th>OFDMA</th>
<th>NOMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.44</td>
<td>17.97</td>
<td>23.34</td>
</tr>
<tr>
<td>0.57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In NOMA, the cell throughput and cell-edge user throughput are evaluated based on the following definitions. The cell throughput is defined as the cell throughput for one macrocell while the cell-edge user throughput is defined as the 5% value of the cumulative distribution function (CDF) of the user throughput.

The performance gains in the overall cell throughput are shown in Table 1.
applying the same precoding weights. By applying the same proposed scheme, the SIC is used for intra-beam user multiplexing precoding or beamforming is applied to each UE. Therefore, in the However, this property is not guaranteed when different MIMO order of the increasing channel gain as already explained. and thus the optimal order of decoding for the NOMA downlink is in the spectrum efficiency as discussed in the previous section. Here we discuss possible extensions of NOMA by applying advanced multi-antenna/site cooperation technologies to achieve further capacity gain.

IV. NOMA WITH MULTI-ANTENNA/SITE EXTENSIONS

Although it may be very challenging, our target for FRA toward the 2020s is to achieve a 3-fold further enhancement in the spectrum efficiency compared to the LTE baseline. However, the gain from basic NOMA using SIC is an improvement of approximately 30-40% in the spectrum efficiency as discussed in the previous section. Here we discuss possible extensions of NOMA by applying advanced multi-antenna/site cooperation technologies to achieve further capacity gain.

A. Proposed NOMA/MIMO Scheme

Fig. 5 shows the proposed NOMA/MIMO scheme in the downlink [10]. In this scheme, the BS transmitter generates multiple beams like a multi-user (MU)-MIMO, and superposes multiple UEs within each beam. In the UE receiver side, two interference cancellation approaches, i.e., SIC and interference rejection combining (IRC) [17], are jointly used based on the following criteria:

✓ SIC is used for intra-beam user multiplexing, i.e., interference cancellation among the UEs belonging to a group applying the same precoding weights. The multiple access scheme within the group is basically the same as basic NOMA with SIC.
✓ IRC is used for inter-beam user multiplexing, i.e., interference suppression among the UE groups applying different precoding weights.

The motivation for the above receiver criteria is based on the following characteristics of each cancellation approach.

SIC – In SIC, it is required that the receiver decode the received signal sent to other UEs for interference cancellation, and thus the SINR of the signal to be cancelled should be sufficiently high so that it is decodable at the receiver. In the downlink, if we do not consider the MIMO precoding or beamforming, it is possible for the receiver to decode a signal sent to other UEs that has a lower channel gain, and thus the optimal order of decoding for the NOMA downlink is in the order of the increasing channel gain as already explained. However, this property is not guaranteed when different MIMO precoding or beamforming is applied to each UE. Therefore, in the proposed scheme, the SIC is used for intra-beam user multiplexing applying the same precoding weights. By applying the same precoding for a group of UEs, overhead for demodulation reference signal (DM-RS) also can be conserved.

IRC – A key benefit of IRC is that it does not require decoding for other UEs since an interference signal is simply suppressed by combining the signals received at multiple UE receiver antennas. In other words, IRC can be applied independently irrespective of the channel gains and MIMO precoding/beamforming weights for other UEs. Therefore, in the proposed scheme, the IRC is used for inter-beam user multiplexing. However, the performance of the IRC receiver degrades when the spatial correlation between the own and interference users is high, which is similar to space division multiple access (SDMA) using an adaptive array antenna receiver [18]. Thus, an appropriate user selection scheme on the BS transmitter side is required to enhance the spectrum efficiency. One possible solution is to apply opportunistic beamforming or random beamforming technology [19, 20].

Fig. 6 shows the total user rate, $R_{\text{sum}}$, as a function of the number of UEs per cell, obtained via the system-level simulations for the proposed NOMA/MIMO scheme with random beamforming at the BS transmitter side and IRC-SIC receiver at the UE side in the downlink. Different from the evaluations in Section III, we apply a simple evaluation model using the Shannon capacity. We employed a 19-hexagonal macrocell model without sectorization. We evaluate two cases, i.e., a 1x2 non-MIMO case as evaluated in Section III and a 2x2 MIMO case with random beamforming. In addition to NOMA, we evaluate a case employing OFDMA, in which a single-stream transmission is applied per transmitter beam. Fig. 6 shows that NOMA increases $R_{\text{sum}}$ irrespective of the 2 antenna configuration cases, i.e., 1x2 non-MIMO or 2x2 random beamforming. One interesting thing observed here is that the performance of NOMA with non-MIMO is very similar to that of OFDMA with random beamforming. This may imply that the NOMA with SIC has a similar effect to spatial multiplexing using random beamforming, and NOMA can achieve a competitive level of performance to random beamforming with a smaller number of BS transmitter antennas. With this in mind, we find that there is more than a two-fold gain when comparing the NOMA with random beamforming and OFDMA with non-MIMO.
Fig. 6 showed that the proposed NOMA/MIMO scheme requires a relatively large number of UEs to obtain a sufficiently saturated throughput gain. When the number of UEs per cell is small, e.g., only a few UEs, it may be better to apply a closed-loop precoding or single-user (SU)-MIMO approach rather than the opportunistic or random beamforming. Therefore, support for multiple MIMO modes, e.g., closed-loop and open-loop, SU-MIMO and MU-MIMO and so on, should be further investigated from a total system perspective.

**B. NOMA with Multi-site Extension**

For NOMA, multi-site extension can also be investigated to achieve further spectrum efficiency enhancement especially to improve the user fairness. For instance, NOMA schemes combined with inter-cell fractional frequency reuse (FFR) are proposed in [11]. Furthermore, inter-site SIC using remote radio heads (RRHs) can be considered in the uplink [21]. Fig. 7 shows an example scenario applying a multi-site extended NOMA using RRHs in the uplink. The performance gain of the uplink multi-site NOMA would increase when the interference over thermal (IoT) at the BS receiver becomes higher. Such an interference-limited situation in the uplink may be often observed in dense small-cell deployments in high traffic areas due to a shorter distance between the UEs and small cells [7].

In the camera ready paper, we will add more discussion regarding NOMA with multi-site extension.

**V. CONCLUSION**

This paper presented our NOMA concept for FRA toward the 2020s information society. Different from the current LTE radio access scheme, NOMA superposes coding although its basic signal waveform could be based on OFDM or DFT-spread OFDM as well as the LTE baseline. In our concept, NOMA adopts a SIC receiver as a baseline receiver scheme for robust multiple access, with the expected evolution of device processing capabilities in the future. Based on the system-level evaluations, we demonstrated that the downlink NOMA with SIC improves both the capacity and cell-edge user throughput performance irrespective of the availability of the frequency-selective CQI at the BS transmitter side. Furthermore, we discussed possible extensions of NOMA by applying jointly multi-antenna/site technologies with the proposed NOMA/MIMO scheme using the SIC and IRC at the UE receivers to achieve a further capacity gain, e.g., a three-fold gain in the spectrum efficiency representing a challenging target for FRA.

**REFERENCES**
