Improving NOMA multi-carrier systems with intentional frequency offsets

Mehdi Ganji, *Student Member, IEEE,* Hamid Jafarkhani, *Fellow, IEEE*

***Abstract*—In this letter, we investigate the possible benefits of asynchrony in the frequency domain for the non-orthogonal multiple access (NOMA) schemes. Despite the common perspective that asynchrony in transmission or reception of multi-stream signals is harmful, we demonstrate the advantages of adding intentional frequency offset to the conventional power domain-NOMA (P-NOMA). We introduce two methods which add artificial frequency offsets between different sets of sub-carriers destined for different users. The first one uses the same successive interference cancellation (SIC) method as the conventional P-NOMA except that it enjoys reduced inter-user interference (IUI) between interfering sub-carriers. The second scheme adopts a precoding at the base station and a linear preprocessing scheme at the receiving user. It decomposes the broadcast channel into parallel channels circumventing the need for SIC. As a result, it fully exploits the advantages provided by the frequency asynchrony and enables the interference-free transmission to the users. The numerical results show that both methods can outperform the conventional P-NOMA.**

I. Introduction

In most of the work in the literature, perfect synchronization in time and frequency is a common presumption. In fact, the asynchrony is mostly considered as an impairment [1], [2] where different synchronization methods are applied to eliminate it [3]. However, by using appropriate transceiver design, asynchrony can indeed be beneficial. The results in [4] show that time asynchrony can increase the capacity region in a single carrier MAC channel. The benefits of time asynchrony in single carrier transmission for different scenarios like MIMO systems, relay networks and NOMA framework have been discussed in [5]–[10]. While time asynchrony in single carrier transmission can be beneficial, unfortunately, in multi-carrier (MC) methods, the effect of time asynchrony turns into a phase shift and loses its benefits. However, the prevalence of MC methods in different communication standards like 3GPP Long Term Evolution (LTE), motivates the need for finding the counterpart of time asynchrony in MC transmission.

In [11], [12], non orthogonal multiple access (NOMA) is proposed as a candidate of future radio access. In NOMA, by relaxing the constraint of orthogonal radio resource allocation, the total number of users served as well as the overall capacity of the system can be greatly improved. The NOMA method in power domain is based on the superposition coding in which the signals intended for different users are superimposed [13]. However, the full advantage of NOMA can only be realized if complex detection algorithms like successive interference cancellation (SIC) are used at the receiving side.

The complexity of employing SIC method at each mobile device scales with the number of users in the cell as it needs to decode and cancel out the interfering signals. Besides, as the number of users increases, the effect of error propagation increases. One reasonable solution is to break the users in the cell into groups and within each group, superposition coding/SIC is performed, and across the groups, transmissions are kept orthogonal. For example, in [14], [15], authors have proposed multi-carrier-based NOMA (MC-NOMA) to enhance the spectrum efficiency of existing MC-based cellular networks. MC-based NOMA can bring us the simplicity of orthogonal multiple access (OMA) while enjoying the increased spectral efficiency of NOMA.

Intuitively, as time asynchrony could improve the single carrier transmission by proper modifications in transmission and reception, frequency asynchrony might be beneficial in MC transmission, as well. Thus, in this work, we investigate this intuition and analyze the benefits of frequency offset in a NOMA framework. We verify that by using proper transmission and receiver design, frequency asynchrony can indeed be beneficial in MC transmission. We introduce two schemes called frequency-asynchronous-power domain-NOMA (FAP-NOMA) and frequency-NOMA (F-NOMA) which use frequency asynchrony to improve and enlarge the downlink rate region.

II. Synchronous Transmission

We first provide the system model for the synchronous scenario, and then, provide that of the asynchronous cases. In this letter, we consider a downlink NOMA communication system where a BS serves two users. The transmitted symbols of users are superimposed on available N sub-carriers. Considering two users with superposition coding on each sub-carrier, the complex envelope of one MC block is expressed as:

$$x(t) = e^{j2\pi f_c t} p(t) \sum_{n=0}^{N-1} \left(\sqrt{P_1} a_n + \sqrt{P_2} b_n \right) e^{j2\pi f_n t}, \quad (1)$$

where f_c is the carrier frequency, f_n is the frequency of the nth sub-carrier, $p(t)$ is the pulse shaping function, and $a_n(b_n)$ is the data symbol transmitted on the nth sub-carrier for the user 1(2). P_k is the transmit power for the kth user with the total power constraint P at BS, i.e. $P_1 + P_2 \leq P$. To ensure sub-carrier orthogonality, it is assumed that $f_n - f_m = (n-m)/T$ where $1/T$ is the minimum sub-carrier frequency spacing required. $p(t)$ is usually assumed to be a pulse shape that has spectral nulls at frequencies $\pm 1/T, \pm 2/T, \cdots$, e.g., Raised Cosine (RC) pulse shape with roll-off factor of β.

We assume an AWGN channel and perfect oscillators at the users to unfold the main points in this work, however, fading channel models and also system impairments will be analyzed in future works. Thus, the received signal at the kth user becomes $r_k(t) = e^{-j2\pi f_c t} x(t) + n_k(t)$ where $n_k(t)$ is the

M. Ganji and H. Jafarkhani are with the Center for Pervasive Communications and Computing, University of California, Irvine, CA, 92697 USA (e-mail: {mganji, hamidj}@uci.edu). This work was supported in part by the NSF Award CCF-1526780.

white Gaussian noise at User k with variance σ_k^2. Without loss of generality, we assume that User 1 has a better channel, i.e., $\sigma_1^2 < \sigma_2^2$ and this order is known by the transmitter and both users. Thus, the mth sub-carrier correlation demodulator at User k provides the decision variable for transmitted symbols a_m and b_m

$$r_k[m] = \int_{-\infty}^{\infty} r_k(t)p^*(t)e^{-j2\pi f_m t}dt, \tag{2}$$

which can be written as sum of the desired signals and inter-carrier interference (ICI):

$$\begin{aligned} r_k[m] &= \left(\sqrt{P_1}a_m + \sqrt{P_2}b_m\right)G(0) \\ &+ \sum_{\substack{n=0 \\ n\neq m}}^{N-1} \left(\sqrt{P_1}a_n + \sqrt{P_2}b_n\right) G\left(\frac{m-n}{T}\right) + n_k[m], \end{aligned} \tag{3}$$

where $G(f)$ is the Fourier transform of $g(t) = p(t)p^*(t)$. Due to the orthogonality between the sub-carriers, the ICI is eliminated and $r_k[m]$ is simplified to $r_k[m] = \sqrt{P_1}a_m + \sqrt{P_2}b_m + n_k[m]$. Then, the well-known SIC method is applied at the users. In particular, User 1 first decodes b_m and subtracts its component from the received signal $r_1[m]$. Then, User 1 decodes a_m without interference from the other user. On the other hand, User 2 directly decodes b_m while considering the other user as noise. Assuming successful decoding and no error propagation, the achievable rate-region R_i of the ith user can be represented as [11], [12]:

$$\begin{cases} 0 \leq R_1 \leq \frac{1}{2}\log_2\left(1 + \frac{P_1}{\sigma_1^2}\right) \\ 0 \leq R_2 \leq \frac{1}{2}\log_2\left(1 + \frac{P_2}{P_1+\sigma_2^2}\right) \end{cases}. \tag{4}$$

III. Asynchronous Transmission

A. System Model

In the asynchronous transmission, we introduce an intentional frequency offset, $0 < \delta f < 1/T$ between sub-carriers of different users. Thus, the transmitted signal in Eq. (1) turns into:

$$x(t) = e^{j2\pi f_c t}p(t)\sum_{n=0}^{N-1}\left(\sqrt{P_1}a_n e^{j2\pi f_n t} + \sqrt{P_2}b_n e^{j2\pi(f_n+\delta f)t}\right).$$

After performing matched filtering and demodulation, each user obtains two sets of samples as follows:

$$r_k^1[m] = \int_{-\infty}^{\infty} r_k(t)p^*(t)e^{-j2\pi f_m t}dt, \tag{5}$$

$$r_k^2[m] = \int_{-\infty}^{\infty} r_k(t)p^*(t)e^{-j2\pi(f_m+\delta f)t}dt. \tag{6}$$

Sample sets of $r_k^1[m]$ and $r_k^2[m]$ can be rewritten as:

$$\begin{aligned} r_k^1[m] &= \sqrt{P_1}a_m G(0) + \sqrt{P_1}\sum_{\substack{n=0 \\ n\neq m}}^{N-1} a_n G\left(\frac{m-n}{T}\right) \\ &+ \sqrt{P_2}\sum_{n=0}^{N-1} b_n G\left(\frac{m-n}{T} - \delta f\right) + n_k^1[m], \end{aligned} \tag{7}$$

$$\begin{aligned} r_k^2[m] &= \sqrt{P_2}b_m G(0) + \sqrt{P_2}\sum_{\substack{n=0 \\ n\neq m}}^{N-1} b_n G\left(\frac{m-n}{T}\right) \\ &+ \sqrt{P_1}\sum_{n=0}^{N-1} a_n G\left(\frac{m-n}{T} + \delta f\right) + n_k^2[m]. \end{aligned} \tag{8}$$

Defining $\boldsymbol{r_k^{1(2)}} = \left(r_k^{1(2)}[0], \cdots, r_k^{1(2)}[N-1]\right)^T$, $\boldsymbol{r_k} = \left(\boldsymbol{r_k^1}^T, \boldsymbol{r_k^2}^T\right)^T$, $\boldsymbol{n_k^{1(2)}} = \left(n_k^{1(2)}[0], \cdots, n_k^{1(2)}[N-1]\right)^T$, $\boldsymbol{n_k} = \left(\boldsymbol{n_k^1}^T, \boldsymbol{n_k^2}^T\right)^T$, $\boldsymbol{a} = (a_0, \cdots, a_{N-1})$ and $\boldsymbol{b} = (b_0, \cdots, b_{N-1})$, the output samples can be written in matrix form as:

$$\begin{aligned} \boldsymbol{r_k} &= \begin{pmatrix} \boldsymbol{I_N} & \boldsymbol{G_{12}} \\ \boldsymbol{G_{21}} & \boldsymbol{I_N} \end{pmatrix}\begin{pmatrix} \boldsymbol{P_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P_2} \end{pmatrix}\begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{pmatrix} + \boldsymbol{n_k} \\ &= \boldsymbol{G^{asyn}P s} + \boldsymbol{n_k}, \end{aligned} \tag{9}$$

where $\boldsymbol{I_N}$ is the $N \times N$ identity matrix, $\boldsymbol{P_i} = \sqrt{P_i}\boldsymbol{I_N}$ and the elements of each sub-matrix $\boldsymbol{G_{ij}}$ is defined as:

$$\boldsymbol{G_{ij}}(m,n) = G\left(\frac{m-n}{T} + (\delta f_i - \delta f_j)\right), \tag{10}$$

where $\delta f_1 = 0$, $\delta f_2 = \delta_f$ and $G(f)$ is the Fourier transform of the pulse shape $g(t)$. For a fair comparison between the asynchronous system model and the synchronous one, we present the synchronous system model with an extra set of samples at an arbitrary frequency shift Δf (Note that shift of samples in time domain will change the samples just by multiplication of a phase rotation).

$$\begin{aligned} \boldsymbol{r_k^{syn}} &= \begin{pmatrix} \boldsymbol{I_N} & \boldsymbol{I_N} \\ \boldsymbol{G} & \boldsymbol{G} \end{pmatrix}\begin{pmatrix} \boldsymbol{P_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P_2} \end{pmatrix}\begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{pmatrix} + \boldsymbol{n_k^{syn}} \\ &= \boldsymbol{G^{syn}P s} + \boldsymbol{n_k^{syn}}, \end{aligned} \tag{11}$$

where

$$\boldsymbol{G}(m,n) = G\left(\frac{m-n}{T} + \Delta f\right). \tag{12}$$

It can be easily verified that the effective channel matrix in the asynchronous scenario, $\boldsymbol{G^{asyn}}$, is Hermitian and positive definite, however the synchronous channel matrix, $\boldsymbol{G^{syn}}$, is rank deficient and results in performance loss. In the next sections, the achievable rates by our asynchronous transmission schemes are analyzed.

B. Achievable Rates for FAP-NOMA

In FAP-NOMA, the SIC detection is used with some modifications. Based on the assumption of User 1 being the stronger user, it can use the sample set of $\boldsymbol{r_1^2}$ to decode the intended symbols for User 2 and subtract them from the received signal. Then, it will be able to decode its own symbols with no interference from User 2. On the other hand, the weaker user,

User 2, will decode its own symbols while considering the signal from User 1 as noise. Thus, for User 1, where the interference from the other user is eliminated, the achievable rate is the same as the synchronous case:

$$R_1 \leq \frac{1}{2(1+\delta fT/N)} \log_2\left(1+\frac{P_1}{\sigma_1^2}\right). \tag{13}$$

The term $\delta fT/N$ indicates the spectral efficiency loss due to adding frequency offset which becomes negligible as N increases. Performance of User 2 is degraded by the interference from User 1. In the synchronous case, the interference power is equal to P_1, however, in the asynchronous case, the interference power is $\sum_{i=-\infty}^{\infty} G\left(\frac{i}{T}+\delta f\right)P_1$. Thus, the achievable rate for User 2 can be expressed as:

$$R_2 \leq \frac{1}{2(1+\delta fT/N)} \log_2\left(1+\frac{P_2}{\varepsilon P_1+\sigma_2^2}\right), \tag{14}$$

where $\varepsilon = \sum_{i=\infty}^{\infty} G\left(\frac{i}{T}+\delta f\right)$. It can be shown that for every δf and all even pulse shapes including the rectangular pulse shape, $\sum_{i=\infty}^{\infty} G\left(\frac{i}{T}+\delta f\right) \leq 1$ and the quality holds if $\delta f = 0$ [9]. As a result, for sufficiently large frame length, the achievable rate region of FAP-NOMA is larger than that of the conventional P-NOMA and it can be optimized by proper choice of the frequency offset.

C. Achievable Rates for F-NOMA

As shown before, unlike the synchronous case where the effective channel matrix is rank deficient, the asynchronous system model is full rank, thus provides higher degrees of freedom, i.e., non-zero eigenvalues. In this section, we introduce F-NOMA to exploit the increased asynchronous matrix rank and present its achievable rate-region. The full rank system model in the asynchronous case is exploited by using proper precoding and post-processing at the transmitter and receiver, respectively, in order to decompose the channel into interference-free sub-channels. F-NOMA can be described in three main steps:

- Noise whitening: due to loss of orthogonality, the noise samples are not independent and their covariance matrix is equal to $\boldsymbol{Q_n} = \boldsymbol{G^{asyn}}$. Using Cholesky decomposition, matrix $\boldsymbol{G^{asyn}}$ can be written as the product of a lower triangular matrix and its Hermitian, i.e., $\boldsymbol{G^{asyn}} = \boldsymbol{L_G}\boldsymbol{L_G}^H$. Multiplying the received samples by $\boldsymbol{L_G}^{-1}$ will result in a white noise with identity covariance matrix:

$$\boldsymbol{L_G}^{-1}\boldsymbol{r_k} = \boldsymbol{L_G}^H\boldsymbol{P}\boldsymbol{s} + \tilde{\boldsymbol{n}}_{\boldsymbol{k}}. \tag{15}$$

- Precoding: denoting $\boldsymbol{L_G}^H\boldsymbol{P}$ as $\boldsymbol{H}$ and considering its SVD decomposition, i.e., $\boldsymbol{H} = \boldsymbol{U_H}\boldsymbol{\Lambda_H}\boldsymbol{U_H}^H$, the transmitted vector can be precoded by a unitary matrix as $\boldsymbol{s} = \boldsymbol{U_H}\hat{\boldsymbol{s}}$. Thus:

$$\boldsymbol{L_G}^{-1}\boldsymbol{r_k} = \boldsymbol{U_H}\boldsymbol{\Lambda_H}\hat{\boldsymbol{s}} + \tilde{\boldsymbol{n}}_{\boldsymbol{k}}. \tag{16}$$

- Post-Processing: at the user side, the received samples can be decomposed into independent sub-channels by multiplying by $\boldsymbol{U_H}^H$:

$$\tilde{\boldsymbol{r}}_{\boldsymbol{k}} = \boldsymbol{U_H}^H\boldsymbol{L_G}^{-1}\boldsymbol{r_k} = \boldsymbol{\Lambda_H}\hat{\boldsymbol{s}} + \tilde{\boldsymbol{n}}_{\boldsymbol{k}}. \tag{17}$$

where matrix $\boldsymbol{\Lambda_H}$ is a $2N \times 2N$ diagonal matrix having eigenvalues of matrix $\boldsymbol{H}$ on its diagonal.

Remarks 1: The noise whitening matrix is pre-calculated based on the pulse shape and frequency offset. In the precoding step, because the precoding matrix is unitary it does not change the power spectrum and thus the total power. Note that the precoding used in [9] is not unitary and despite satisfying the total power constraint, it suffers from slight spectrum broadening. The post-processing matrix at the receiver side depends on the pulse shape, frequency offset and power assignment and can be provided by the base station to remove the complexity from the user side. It can be observed that the interference channel is decomposed into $2N$ independent interference-free sub-channels whose strength are specified by eigenvalues of matrix $\boldsymbol{H}$.

Assuming the block length to be N, there will be $\binom{2N}{N}$ different assignments of the available sub-channels to users. Assume that the ith assignment associates N eigenvalues to User 1, i.e., $\boldsymbol{\lambda_1^i} : \{\lambda_{11}^i, \cdots, \lambda_{1N}^i\}$ and the rest of them, i.e., $\boldsymbol{\lambda_2^i} : \{\lambda_{21}^i, \cdots, \lambda_{2N}^i\}$ to User 2. Denoting the achievable rates for User 1 and User 2 corresponding to the ith assignment as R_1^i and R_2^i, then the total rate region is the union over all possible assignments, i.e., $(R_1, R_2) = \bigcup_i (R_1^i, R_2^i)$ and R_k^i is defined as:

$$R_k^i \leq \frac{1}{2N(1+\delta fT/N)} \sum_{n=1}^{N} \log_2\left(1+\frac{(\lambda_{kn}^i)^2}{\sigma_k^2}\right) \quad k = 1, 2.$$

Therefore, by exploiting the additional non-zero eigenvalues, offered by frequency asynchrony, interference-free detection is possible at both users. As the number of users increases, SIC will be complex for practical implementations. Besides, the authors in [16] showed that the optimal number of scheduled users on a sub-carrier is either two or three. Therefore, it is more practical to divide a group of users into multiple subgroups with two or three users. Nonetheless, the generalization to more number of users is straightforward and interested reader is refereed to [9] for similar generalization.

IV. Numerical Results

In this section, we present numerical results to show the effectiveness of using frequency offset in providing a larger rate region.

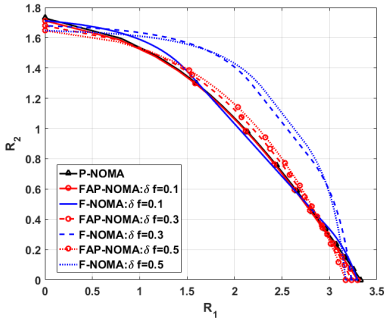

Fig. 1: Achievable rate-regions of the proposed methods in a Gaussian channel with $\sigma_1^2 = 0.1$, $\sigma_2^2 = 1$ and $N = 10$.

We assume that a transmit power of $P = 10\ dBW$ is available at the base station which is serving two users. We consider

the RC pulse shaping function ($\beta = 1$ in Figs. 1, 2 and $\beta = 0.5$ in Fig. 3) and assume that the symbol duration T is normalized to be 1. In Fig. 1, we show the achievable rate-regions of FAP-NOMA and F-NOMA with different frequency offsets in a Gaussian channel. The improvements in FAP-NOMA method is the result of IUI reduction due to frequency asynchrony and hence $\delta f = 0.1$ results in the worst performance close to the P-NOMA rate region while $\delta f = 0.5$ provides the least IUI and thus the largest rate region. Note that the loss in the corners of the rate region is due to spectral loss, i.e., $\delta f T/N$, which is negligible as N increases. F-NOMA method exploits the additional matrix rank provided by frequency offset and achieves significant rate region enlargement particularly with $\delta f = 0.5$. In Fig. 2, rate regions for different values of $N = 8, 10$ and 12 are shown which confirms the increase in the rate region as N increases. In Fig. 3, the symbol error rate (SER) performance of two users in a frequency selective channel is shown where User 1 and User 2 experience channel taps' powers of $[0, -1, -3]dB$ and $[0, -2, -4]dB$, respectively. It is assumed that User 1 with better channel uses QPSK with $P_1 = 10\ dBW$ transmit power and User 2 with worse channel uses BPSK with $P_2 = 13\ dBW$ transmit power. In FAP-NOMA method User 2 enjoys better SER performance due to decreased IUI compared with P-NOMA while the other user experiences the same performance. However, in F-NOMA method both users achieve improvement in SER performance, i.e., around $1dB$ and $3dB$ gains, for User 1 and User 2, respectively, at $SER = 10^{-6}$.

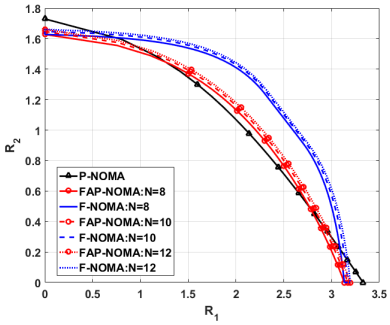

Fig. 2: Achievable rate-regions of the proposed methods in a Gaussian channel with $\sigma_1^2 = 0.1$ and $\sigma_2^2 = 1$.

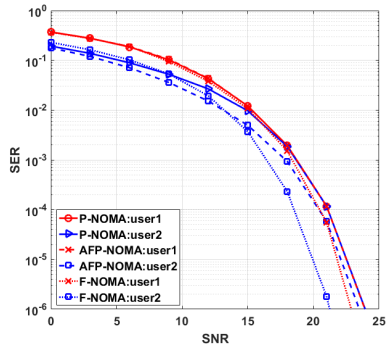

Fig. 3: Comparison of SER performance of proposed methods in a frequency selective channel with $N = 64$.

V. Conclusion

In this work, we propose novel frequency-asynchronous downlink NOMA schemes. In contrast to the conventional P-NOMA, we propose to intentionally add frequency offsets among superimposed sub-carriers. The receiver architecture in FAP-NOMA includes a SIC scheme similar to P-NOMA, however, frequency asynchrony reduces IUI and improves the performance. F-NOMA exploits the additional non-zero eigenvalues introduced by frequency asynchrony using novel precoding and simple post processing at users. In other words, F-NOMA decomposes the channel into independent sub-channels and eliminates the interference. Our analysis shows that the FAP-NOMA and F-NOMA can improve the performance including achievable rate-regions and SER.

References

[1] V. Kotzsch and G. Fettweis, "Interference analysis in time and frequency asynchronous network MIMO OFDM systems," *2010 IEEE Wireless Communication and Networking Conference*, pp. 1–6, Apr. 2010.

[2] S. Abeywickrama, L. Liu, Y. Chi, and C. Yuen, "Over-the-air implementation of uplink NOMA," *IEEE Global Communications Conference*, pp. 1–6, Dec. 2017.

[3] A. A. Nasir, S. Durrani, H. Mehrpouyan, S. D. Blostein, and R. A. Kennedy, "Timing and carrier synchronization in wireless communication systems: a survey and classification of research in the last 5 years," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, p. 180, Dec. 2016.

[4] S. Verdu, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Transactions on Information Theory*, vol. 35, no. 4, pp. 733–751, Jul. 1989.

[5] D. K. So and Y. Lan, "Virtual receive antenna for overloaded MIMO layered space-time system," *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1610–1620, Jun. 2012.

[6] M. Ganji and H. Jafarkhani, "Interference mitigation using asynchronous transmission and sampling diversity," *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec. 2016.

[7] X. Zhang, M. Ganji, and H. Jafarkhani, "Exploiting asynchronous signaling for multiuser cooperative networks with analog network coding," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, Mar. 2017.

[8] J. Cui, G. Dong, S. Zhang, H. Li, and G. Feng, "Asynchronous NOMA for downlink transmissions," *IEEE Communications Letters*, vol. 21, no. 2, pp. 402–405, Feb. 2017.

[9] M. Ganji and H. Jafarkhani, "Novel time asynchronous NOMA schemes for downlink transmissions," *arXiv preprint arXiv:1808.08665*, 2018.

[10] X. Zou, B. He, and H. Jafarkhani, "An analysis of two-user uplink asynchronous non-orthogonal multiple access systems," *IEEE Transactions on Wireless Communications*, Jan 2019.

[11] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[12] K. Higuchi and A. Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," *IEICE Transactions on Communications*, vol. 98, no. 3, pp. 403–414, Mar. 2015.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[14] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.

[15] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 611–615, Sep. 2013.

[16] A. Zafar, M. Shaqfeh, M.-S. Alouini, and H. Alnuweiri, "On multiple users scheduling using superposition coding over Rayleigh fading channels," *IEEE Communications Letters*, vol. 17, no. 4, pp. 733–736, Apr. 2013.