



HAL
open science

Waterfilling-based Resource Allocation Techniques in Downlink Non-Orthogonal Multiple Access (NOMA) with Single-User MIMO

Marie-Josépha Youssef, Joumana Farah, Charbel Abdel Nour, Catherine Douillard

► **To cite this version:**

Marie-Josépha Youssef, Joumana Farah, Charbel Abdel Nour, Catherine Douillard. Waterfilling-based Resource Allocation Techniques in Downlink Non-Orthogonal Multiple Access (NOMA) with Single-User MIMO. ISCC 2017: The 22nd IEEE Symposium on Computers and Communications, Jul 2017, Crete, Greece. pp.499 - 506, 10.1109/ISCC.2017.8024578 . hal-01617336

HAL Id: hal-01617336

<https://hal.science/hal-01617336>

Submitted on 10 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Waterfilling-based Resource Allocation Techniques in Downlink Non-Orthogonal Multiple Access (NOMA) with Single-User MIMO

Marie-Josepha Youssef⁽¹⁾, Joumana Farah⁽¹⁾, Charbel Abdel Nour⁽²⁾, Catherine Douillard⁽²⁾

⁽¹⁾Department of Electricity and Electronics, Faculty of Engineering,
Lebanese University, Roumieh, Lebanon

⁽²⁾IMT Atlantique, Department of Electronics, Lab-STICC - UMR 6285
Technopôle Brest Iroise, CS 83 818 - 29238 Brest Cedex, France

Abstract—Non-Orthogonal Multiple Access (NOMA) has proven itself as a serious candidate for multiple access schemes in the 5th generation of wireless communication systems. In this paper, we propose new resource allocation techniques that allow the coupling of NOMA with multiple-input-multiple-output (MIMO) systems. As in most of downlink NOMA systems, the proportional fairness (PF) scheduler is used. As for power allocation (PA), most previous works assume an equal PA between antennas and subbands, which is sub-optimal. In this paper, we first propose a technique to reduce the complexity of the PF scheduler, and then we introduce an iterative waterfilling-based power allocation to the Single-User MIMO (SU-MIMO) case. Extensive simulations show that the novel PA technique improves the quality of experience of users suffering from bad channel conditions, as well as the user fairness.

Index Terms—downlink transmission, MIMO, non-orthogonal multiple access, power allocation, scheduling, waterfilling.

I. INTRODUCTION

Mobile traffic is predicted to reach a volume 10 times larger than today's by 2020. Future communication systems must accommodate the needs of machine-type communications involving a massive number of automated devices, in addition to conventional human-type-communications. Therefore, significant increases in the system capacity, the quality of service and the quality of user experience are required. In addition to that, future communication systems must deal with the spectrum crunch, and hence must use the available spectrum more efficiently.

One way to improve the spectral efficiency and to respond to the needs of a growing number of connected devices is to adopt novel multiple access schemes. In the 3.9 and 4th Generation (4G) of mobile communication systems, such as Long-Term Evolution (LTE) [1] and LTE-Advanced [2], Orthogonal Multiple Access (OMA) based on Orthogonal Frequency Division Multiplexing (OFDM) or Single Carrier Frequency Division Multiple Access (SC-FDMA) has been adopted, respectively for the downlink and uplink transmissions. OMA techniques are known to achieve good system performance and present a good potential for practical implementations due to a simple receiver design. However, if 5G systems are to support the required high throughputs, there is a need to diverge from the conventional approach of OMA schemes.

From an information-theoretical point of view, it is well known that non-orthogonal user multiplexing using superposition coding at the transmitter and proper decoding techniques at the receiver not only outperforms orthogonal multiplexing, but is also optimal in the sense of achieving the capacity region of the downlink broadcast channel [3]. Therefore, a new technique based on non-orthogonal multiple access (NOMA) has recently emerged. It relies on further multiplexing different users on the same subband in the power domain [4], [5], on top of the OFDM layer. The efficient use of NOMA requires the implementation of pertinent allocation algorithms that seek to strike a better trade-off between system throughput and user fairness, compared to OMA [4]–[7].

Furthermore, multiple-input-multiple-output (MIMO) antenna systems constitute one of the major components of current communication systems, since they can fundamentally increase system capacity. By exploiting the spatial multiplexing abilities of MIMO channels, it is possible to increase the achieved data rates while using the same amount of bandwidth. Therefore, using MIMO can greatly help overcome the challenges of the future generations of mobile communications.

In this paper, we focus on the design of scheduling techniques for a downlink single-user (SU) MIMO system based on NOMA. Most of the papers dealing with scheduling for NOMA systems consider the proportional fairness (PF) scheduler which aims at striking a balance between system capacity and user fairness through the maximization of a proper scheduling metric [4], [5], [8], [9]. The application of the PF scheduler in the SU-MIMO NOMA context is rather complex since it requires testing a large number of candidate user set combinations. In this work, we adopt an improved version of the PF scheduler which was introduced recently in [10]. We further propose to decrease the complexity of the PF scheduler by decreasing the number of candidate user sets to test during each allocation.

When it comes to power allocation, most previous works assume an equal repartition between the subbands [4], [5]. In the case of a MIMO system, they also assume equal power repartition between the transmit antennas [8], [9]. However, an improvement in the system performance is expected by

applying the waterfilling principle for the inter-subband and the inter-antenna power allocation. Recently, [11] proposed an iterative waterfilling technique for the PF scheduler in a single-input-single-output (SISO) system. In this work, we aim to generalize the proposed technique to the case of a SU-MIMO system.

The remainder of the paper is organized as follows. Section II summarizes the system model. In section III, we present the proposed enhancement to the scheduling algorithm and the proposed power allocation method in SU-MIMO. In section IV, the performance of the proposed techniques is evaluated, while section V concludes the paper.

II. SYSTEM MODEL

This section describes the concept of NOMA, including user pairing at the transmitter and signal separation at the receiver in a SU-MIMO downlink system.

We consider a downlink system consisting of N_t and N_r transmit and receive antennas, respectively. The system bandwidth B is partitioned into S subbands and the total transmit power of the base station (BS) is set to P_{max} . The number of users per cell is K . While the entire network consists of multiple cells, we assume that there is proper resource partitioning or resource management to address inter-cell interference, and hence we focus our study on a single cell. Over each subband, the BS simultaneously transmits signals with different power levels to N_s users, where N_s is the number of users non-orthogonally multiplexed over each subband. In our work, we assume that two users are multiplexed via NOMA on each subband, i.e. $N_s = 2$; in fact [5], [12] have shown that greater values of N_s present minimal performance gains at the expense of an increased complexity. Furthermore, since SU-MIMO transmission is assumed, the BS transmits to the same user set on a considered subband on all transmit antennas. Hence, the transmitted signal x_s on subband s is given by:

$$x_s = \sum_{r=1}^{N_t} \sum_{k=1}^{N_s} \sqrt{P_{k,r,s}} x_{k,r,s}, \quad (1)$$

where $P_{k,r,s}$ is the power allocated to the k^{th} non-orthogonally multiplexed user on subband s over the r^{th} antenna, and $x_{k,r,s}$ the transmitted signal relative to that user.

The received signal of user k on subband s is given by:

$$y_{k,s} = H_{k,s} x_s + w_{k,s}. \quad (2)$$

In (2), $H_{k,s}$ denotes the channel matrix which contains the coefficients characterizing the frequency-selective multipath fading channel. $w_{k,s}$ denotes the additive white Gaussian noise including inter-cell interference relative to user k over subband s .

Dealing with a MIMO system, the channel matrix of each user over each subband consists of $N_t \times N_r$ channel coefficients, significantly complicating the scheduling of users over subbands. [3], [13] propose to decompose the channel matrix into an equivalent set of parallel, independent scalar

Gaussian sub-channels through singular value decomposition (SVD). The SVD allows transforming the MIMO system into $\min(N_t, N_r)$ independent channels modeled by $\min(N_t, N_r)$ singular values. Therefore, the scheduling and power allocation can be approached in a more structural manner and methods inspired from those used in SISO systems can be employed.

It has been shown that the best performance in NOMA-based systems is achieved when users with large channel gain differences are paired together, i.e. in the case where $N_s = 2$, a user with a good channel state is superposed with a user with a bad channel state (i.e. cell-edge user). Therefore, at the receiver side, the signal separation can be done according to a successive interference cancelation (SIC) process [3]. The optimal order of the SIC separation is in increasing order of the user channel gain. In other words, user j can properly decode user i 's signal if user j has a better channel gain than user i . When $N_s = 2$, if user 1 has a better channel state than user 2, the latter does not perform SIC but treats the signal of user 1 as noise and decodes its data from the received signal. On the other hand, user 1 performs SIC: it decodes the data of user 2 and then subtracts the transmit signal of user 2 from its received signal before decoding its own data. If we assume successful decoding and no error-propagation, the achieved data rates by user 1 and 2 over subband s , $R_{s,1}$ and $R_{s,2}$ are given by:

$$R_{s,1} = \sum_{r=1}^{\min(N_t, N_r)} \frac{B}{S} \log_2 \left(1 + \frac{P_{1,r,s} \lambda_{1,r,s}^2}{N_0 B / S} \right),$$

$$R_{s,2} = \sum_{r=1}^{\min(N_t, N_r)} \frac{B}{S} \log_2 \left(1 + \frac{P_{2,r,s} \lambda_{2,r,s}^2}{P_{1,r,s} \lambda_{2,r,s}^2 + N_0 B / S} \right). \quad (3)$$

$\lambda_{k,r,s}$ denotes the singular value describing the r^{th} channel of the k^{th} user on subband s , where $k = 1, 2$. N_0 is the power spectral density of the additive white Gaussian noise which is assumed to be constant over all subbands.

From (3), we notice that the choice of paired users as well as the amount of power allocated to each user greatly influences the rates achieved by all users superposed on a certain subband. For this purpose, proper multi-user scheduling and multi-user power allocation techniques will be proposed and evaluated in the sequel.

III. PROPOSED SCHEDULING AND POWER ALLOCATION TECHNIQUES

In this section, the scheduling and power allocation techniques employed in this work are presented. The aim of these algorithms is to decide on the choice of users to be superposed over each subband and each transmit antenna, as well as the amount of power that should be attributed to each one of these users. In addition to maximizing system throughput, this work pays particular attention at enhancing the fairness between users. Fairness measures to which extent users are receiving a fair share of system resources.

A. Proportional Fairness Scheduler

1) Conventional Scheduling Scheme:

The Proportional Fairness (PF) scheduler, extensively used in the literature [4], [5], [8], [9], aims at reaching a compromise between maximizing the achieved throughput of the entire cell and maximizing the fairness between users. To do so, the scheduler chooses the user set that maximizes the PF metric over a subband in a given time slot. In our MIMO system, after the SVD is performed, the PF algorithm schedules over subband s the user set U_s^* that maximizes the following metric:

$$U_s^* = \arg \max_U \sum_{k \in U} \frac{R_{s,k}(t)}{T_k(t)}. \quad (4)$$

$R_{s,k}(t)$ is the rate achieved by user $k \in U$ over subband s , where U denotes a candidate user set. $T_k(t)$ is the average throughput previously achieved by user k before time slot t . At each time slot, $T_k(t)$ is updated according to:

$$T_k(t+1) = (1 - 1/t_c)T_k(t) + 1/t_c \cdot R_k(t). \quad (5)$$

In (5), $R_k(t)$ is the total throughput achieved by user k during the previous time slot: $R_k(t) = \sum_{s=1}^S R_{s,k}(t)$. It can amount to zero if k was not granted any transmission in time slot t . Parameter t_c defines the time horizon in which we want to achieve fairness.

As stated in the previous section, NOMA achieves the best results when pairing together users with large channel gain differences. However, as was also observed in [9], we can no longer classify the users according to their channel gains because the received signal comprises $N_t \times N_r$ channel coefficients. In fact, the received signal is a superposition of multiple signals where each one is affected by a different channel coefficient. Hence, the pairing of users on the different subbands for the maximization of the PF metric is not straightforward in the case of NOMA. Consequently, as done in previous works [4], [5], [8], [9], the scheduling algorithm should test every possible user combination. This means that, in the case of NOMA with $N_s = 2$, when calculating the metric for users 1 and 2, the scheduler first computes the metric considering that user 1 is the strongest user and user 2 the weakest one, and then repeats the same computation after inverting the order. This ensures that all possible user pairing combinations are tested. Also, the scheduler calculates the PF metric for the case when a single user is scheduled for transmission over a subband, to allow an adaptive switching from NOMA to OMA. This leaves the scheduler with the following number of possible combinations to be tested at each allocation:

$$n_{poss} = K + P(K, N_s), \quad (6)$$

where $P(K, N_s)$ is the permutation of N_s users between a total number of K users, that accounts for the NOMA scheduling case, whereas the K term accounts for the OMA case.

After testing all possible combinations, the scheduler allocates resources to the user(s) that maximize(s) the PF metric.

Over each subband, the PF metric takes into account the current rate that would be achieved over that subband by the user(s) if granted transmission, in addition to the rate previously reached by the user(s) at time slot t .

2) Improved Scheduling Scheme:

Recently, [10] addressed the problem that the conventional PF scheduler does not take into consideration the rates achieved by users during the current allocation. The fairness in the conventional PF is achieved according to the last allocation. In other words, if a user is scheduled for transmission over subband s during time slot t , the rate achieved by this user on subband s is not considered during the calculation of the PF metric for the allocation of subband $(s+1)$ during the same time slot t , which might degrade the performance of the scheduling process. To compensate for this, [10] proposed to successively update the value of the average throughput T . Although this increases slightly the scheduling complexity, it significantly improves system performance. In this work, we take this improvement into consideration. Thus, the PF scheduling metric (4) becomes:

$$U_s^* = \arg \max_U \sum_{k \in U} \frac{R_{s,k}(t)}{T_k(t) + \frac{1}{t_c} \sum_{i=1}^{s-1} R_{i,k}(t)}. \quad (7)$$

In (7), the term $\sum_{i=1}^{s-1} R_{i,k}(t)$ accounts for the throughput achieved by user k on all previously allocated subbands during the current time slot t . It can amount to zero if this user has not been allocated any subband so far during the time slot t . Note that, contrary to [10], in our work this term is weighted by $1/t_c$ in order to average its contribution over the time window of length t_c .

3) Proposed Reduced-Complexity Scheduling Scheme:

In our work, we addressed the complexity problem that the scheduling process faces in a SU-MIMO-NOMA.

As stated in section III-A1, because the ranking of users according to their channel gain is not straightforward in a MIMO system, the PF scheduler must test n_{poss} candidate user sets, where n_{poss} is given by (6). This incurs a high number of user sets to test. For example, if the system consists of 10 users, and $N_s = 2$, a total of $n_{poss} = K + P(K, N_s) = 100$ user sets should be tested on each subband. If the system bandwidth is partitioned into $S = 64$ subbands, this incurs 6400 user sets to test at each time instant.

To reduce this complexity, we propose to reduce the number of user sets to be tested. Therefore, we propose to rank the users according to the channel gain of the first transmit antenna, in other words, according to the first singular value resulting from the SVD (since the SVD yields singular values in a decreasing order). This way, user i with a higher singular value than user j is considered stronger. Then, for each user i , we only test the user sets where he is considered to be the stronger user, i.e., the combinations in which user i is superposed with the users having a lower singular value on

the considered subband. According to this ranking, we are able to reduce the number of tested combinations to:

$$n'_{poss} = K + \frac{P(K, N_s)}{2}. \quad (8)$$

In fact, for the same parameters previously stated in this section, the number of combinations to test in each time instance reduces to 3520 instead of 6400, i.e. is almost halved.

B. Proposed Power Allocation Technique

In a MIMO system with a support for NOMA, the power allocation problem consists in dividing the total transmit power between all transmit antennas (inter-antenna power allocation), between the total number of subbands (inter-subband power allocation), and between the users non-orthogonally multiplexed over a certain subband (intra-subband power allocation). Most previous works [4], [5], [8], [9], [12], [13] assume an equal repartition of the total BS transmit power P_{max} between the transmit antennas and the subbands. In this case, the power allocated to each transmit antenna and to each subband, respectively, is given by:

$$P_r = \frac{P_{max}}{N_t}, \quad r = 1, \dots, N_t, \quad \text{and} \\ P_{r,s} = \frac{P_r}{S}, \quad s = 1, \dots, S. \quad (9)$$

For the intra-subband power allocation, multiple approaches have been proposed in the literature, including the full search power allocation (FSPA) [8], the static power allocation [7] and the fractional transmit power allocation (FTPA) [4], [6], [7]. In this work, we adopt the FTPA for the intra-subband power allocation since it is dynamic and much less computationally complex than FSPA. Hence, the power allocated to subband s , $P_{r,s}$, is further divided between the users non-orthogonally multiplexed over this subband according to:

$$P_{k,r,s} = \frac{P_{r,s} \left(\frac{\lambda_{k,r,s}^2}{N_0 \frac{B}{S}} \right)^{-\alpha_{FTPA}}}{\sum_{j \in U_s^*} \left(\lambda_{j,r,s}^2 / N_0 \frac{B}{S} \right)^{-\alpha_{FTPA}}}, \quad k \in U_s^*. \quad (10)$$

In (10), U_s^* is the set of users scheduled on subband s , $P_{k,r,s}$ is the power allocated to user k on that subband. α_{FTPA} ($0 \leq \alpha_{FTPA} \leq 1$) is a decay factor.

Proposed Waterfilling Procedure For the Inter-Subband and Inter-Antenna Power Allocation

The waterfilling principle is known to boost system performance since it allocates power, in an optimized manner, according to the channel coefficients. Recently, [11] proposed a novel method for iterative waterfilling across subbands for a SISO system using the PF scheduler. The waterline level is recursively predicted from the previous level and from the channel gain of the considered candidate user to be scheduled on the current subband and having the highest channel gain on this subband. In this work, we present a generalization of this iterative method to the case of a MIMO system. For this purpose, we develop a joint iterative waterfilling-based power allocation across transmit antennas and subbands. Indeed, according to [3], the best performance of a MIMO

system is achieved when the power is divided between the transmit antennas according to the waterfilling algorithm.

In this method, we propose to perform a joint inter-subband and inter-antenna power repartition by a one-dimensional iterative waterfilling procedure. At the beginning of a time slot, for the first subband ($s = 1$), the whole transmit power P_{max} is divided between the transmit antennas according to the waterfilling principle. This subdivision is performed for each candidate user set. In other words, for every user set U where k^* is the user with the highest channel gain, the power allocated to transmit antenna r is given by:

$$P_{U,r,1} = W(1) - \frac{N_0 B/S}{\lambda_{k^*,r,1}^2}, \quad r = 1, \dots, \min(N_t, N_r). \quad (11)$$

$W(1)$ is the waterline at subband $s = 1$; it is estimated such that the total power constraint $\sum_{r=1}^{\min(N_t, N_r)} P_{U,r,1} = P_{max}$ is satisfied:

$$W(1) = \frac{1}{\min(N_t, N_r)} \left(P_{max} + \sum_{r=1}^{\min(N_t, N_r)} \left(\frac{N_0 B/S}{\lambda_{k^*,r,1}^2} \right) \right). \quad (12)$$

On subsequent subbands s , at each allocation stage, P_{max} is redistributed between the previously allocated subbands and the newly added one. The redistribution is only considered if subband s has a sufficiently high channel gain. In other terms, the waterline is updated (and hence the subband s is considered for allocation to the candidate user set) only if the inverse channel state of s , corresponding to the best state among the $\min(N_t, N_r)$ channels, is lower than the previous waterline, i.e. if:

$$\min_r \left(\frac{N_0 B/S}{\lambda_{k^*,r,s+1}^2} \right) < W(s), \quad r = 1, \dots, \min(N_t, N_r). \quad (13)$$

For every user set satisfying this condition, the total power is redistributed between the subbands previously allocated and the new added subband ($s + 1$) such that (14) is verified:

$$P_{max} = \sum_{i=1}^{s+1} \sum_{r=1}^{\min(N_t, N_r)} \left(W(s+1) - \frac{N_0 B/S}{\lambda_{k^*,r,i}^2} \right). \quad (14)$$

Since the total transmit power at the previous allocation stage for subband s was the same as the one at stage ($s + 1$), we can also write:

$$P_{max} = \sum_{i=1}^s \sum_{r=1}^{\min(N_t, N_r)} \left(W(s) - \frac{N_0 B/S}{\lambda_{k^*,r,i}^2} \right). \quad (15)$$

As (14) denotes the partitioning of the power between the ($s + 1$) allocated subbands, it can be re-written as:

$$P_{max} = \sum_{i=1}^s \sum_{r=1}^{\min(N_t, N_r)} \left(W(s+1) - \frac{N_0 B/S}{\lambda_{k^*,r,i}^2} \right) \\ + \sum_{r=1}^{\min(N_t, N_r)} \left(W(s+1) - \frac{N_0 B/S}{\lambda_{k^*,r,s+1}^2} \right). \quad (16)$$

By comparing (15) and (16), and if we denote by $N(s)$ the number of subbands allocated at the s^{th} allocation stage, we get the following expression for the new waterline $W(s+1)$ as a function of the previous one, $W(s)$:

$$W(s+1) = \frac{1}{\min(N_t, N_r)(N(s)+1)} \left(\min(N_t, N_r)N(s) \cdot W(s) + \sum_{r=1}^{\min(N_t, N_r)} \frac{N_0 B/S}{\lambda_{k^*, r, s+1}^2} \right). \quad (17)$$

After the new waterline is updated using (17), the power allocated to the newly added subband on every transmit antenna is calculated according to:

$$P_{U, r, s+1} = W(s+1) - \frac{N_0 B/S}{\lambda_{k^*, r, s+1}^2}, \quad r = 1, \dots, \min(N_t, N_r). \quad (18)$$

Note that the power on the previously allocated subcarriers does not need to be updated at this stage. The complete update can be moved to the end of the algorithm in order to limit the computational load.

Afterwards, the power $P_{U, r, s+1}$ is partitioned on each channel r between superposed users via FTPA, using (10). Then, the PF metric is calculated using (7) to determine the best candidate user set U_{s+1}^* .

While subband $(s+1)$ is being allocated, one or more channels may have very low coefficients, resulting in negative power values. In such a case, the waterline at this stage is recalculated after setting the power for the corresponding antennas to zero. In other terms, these antennas do not transmit any information on this subband and the transmit power allocated to the subband is partitioned between the remaining antennas. If we denote by $S_{s+1, n-all}$ the set of non-allocated channels on subband $(s+1)$, the total power at $(s+1)$ must be partitioned according to:

$$P_{max} = \sum_{i=1}^s \sum_{r=1}^{\min(N_t, N_r)} \left(W(s+1) - \frac{N_0 B/S}{\lambda_{k^*, r, i}^2} \right) + \sum_{r=1, r \notin S_{s+1, n-all}}^{\min(N_t, N_r)} \left(W(s+1) - \frac{N_0 B/S}{\lambda_{k^*, r, s+1}^2} \right). \quad (19)$$

In this case, (17) must be re-written to reflect the change depicted by (19). Noting by $|S_{s+1, n-all}|$ the number of non-allocated channels at subband $(s+1)$, (17) becomes:

$$W(s+1) = \frac{(\min(N_t, N_r)N(s)W(s))}{\min(N_t, N_r)(N(s)+1) - |S_{s+1, n-all}|} + \frac{\sum_{r=1, r \notin S_{s+1, n-all}}^{\min(N_t, N_r)} \frac{N_0 B/S}{\lambda_{k^*, r, s+1}^2}}{\min(N_t, N_r)(N(s)+1) - |S_{s+1, n-all}|}. \quad (20)$$

Note that the removed channels will be accounted for in the computation of every next waterline, during the current time slot. When one or more channels, on one or more subbands,

have been turned off before reaching subband s , the waterline at this subband is calculated according to:

$$W(s) = \frac{1}{\min(N_t, N_r)(N(s-1)+1)} \left((\min(N_t, N_r)N(s-1) - n_{not-all}(s))W(s-1) + \sum_{r=1}^{\min(N_t, N_r)} \frac{N_0 B/S}{\lambda_{k^*, r, s}^2} \right). \quad (21)$$

In (21), $n_{not-all}(s)$ is the number of non-allocated channels before subband s .

When all subbands have been considered for allocation and the last waterline level has been reached, any channel having a state worse than that dictated by the last waterline is freed. The final power values assigned to each transmit antenna on each allocated subband in the current time slot are calculated according to:

$$P_{U^*, r, s} = W_{final} - \frac{N_0 B/S}{\lambda_{k^*, r, s}^2}, \quad s \in S_{allocated}, r = 1, \dots, \min(N_t, N_r). \quad (22)$$

In (22), U^* is the user set chosen for transmission on subband s , W_{final} is the final waterline level, and $S_{allocated}$ is the set of allocated subbands, i.e. those having an inverse channel gain lower than the final waterline. Then, each $P_{U^*, r, s}$ is divided between the non-orthogonally superposed users within each subband using FTPA, and the actual rate achieved by each scheduled user is updated.

IV. NUMERICAL RESULTS

A. Performance Evaluation

In this paper, we focus on two important system-level performance indicators: the achieved throughput per user and the user fairness. User fairness is estimated by the Gini Fairness index given by [14]:

$$G = \frac{1}{2K^2 \bar{r}} \sum_{x=1}^K \sum_{y=1}^K |r_x - r_y|, \quad \text{with } \bar{r} = \frac{1}{K} \sum_{k=1}^K r_k, \quad (23)$$

where r_k is the total throughput achieved by the k^{th} user. G varies between 0 and 1, a value of zero (resp. one) expresses the maximum (resp. lowest) level of fairness between users in terms of achieved throughput.

B. System Parameters

The simulation parameters are summarized in Table I. The channel model adopted in our numerical evaluations is the Spatial Channel Model Extended (SCME) [15] which is a standardized model developed by 3GPP-3GPP2 for evaluating MIMO system performance in outdoor environments. It is an extension of the 3GPP Spatial Channel Model (SCM) and allows higher values of the system bandwidth. The scenario adopted for the SCME is "Urban-Micro" with a varying cell size. Perfect channel estimation is assumed in this work.

TABLE I
SIMULATION PARAMETERS

| Parameter | Value Range |
|---|--|
| Channel Model | SCME |
| Cell Radius | 500 m to 10 km |
| Carrier Frequency | 2 GHz |
| Overall Transmission Bandwidth | 10 MHz |
| Number of subbands | 8, 16, 32, 64 |
| Number of users in the cell | 5, 10, 15, 20 |
| Number of Transmit Antennas | 2 |
| Number of Receive Antennas | 2 |
| Maximum Transmission Power | 40 Watts (46 dBm) |
| Distance Dependent Path Loss | $128.1 + 37.6 \log_{10}(d)$ (dB), d in Km |
| Receiver Noise Density | -174 dBm/Hz |
| User Speed | 15 m/s (54 Km/h) |
| Maximum number of users multiplexed with NOMA | 2 |
| $\alpha_{FTP A}$ | 0.7 |

C. Simulation Results

In this subsection, we analyze the performance of the described scheduling methods. Our proposed scheduling technique will be referred to as Waterfilling-PF in the graphs. It relies on the waterfilling-based power allocation presented in section III-B, with the incorporation of both the successive update of the average throughput T described in III-A2, and the complexity reduction technique described in III-A3. Two other versions of the PF are considered for comparison:

- The equal-power PF that takes into account the actual achieved rates [10] as described in III-A2 (EP-PF-AR).
- A version of the latter including the proposed complexity reduction strategy (EP-PF-AR-CR).

In these simulations, we considered a system consisting of 10 users, each one assumed to be moving on a circle in a cell of radius 500 m. The ten circles are equally spaced on the cell, i.e. the radius of the circle on which the first user is moving is 50 m, the second user is moving on a circle of radius 100 m, and so on, until the last one which is 500 m (corresponding to a single cell-edge user).

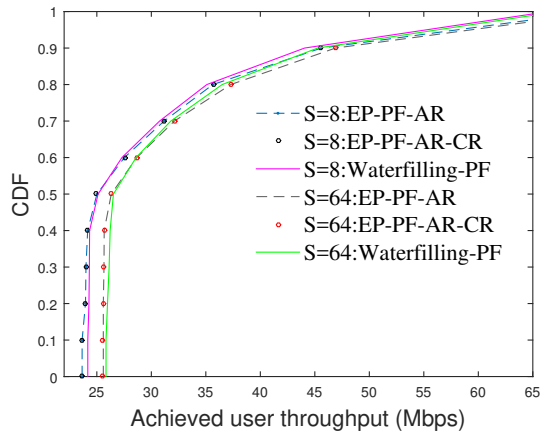


Fig. 1. CDF of the achieved user throughput for a system consisting of 10 users, with $S = 8$ and $S = 64$ subbands.

Fig. 1 shows the cumulative distribution function (CDF) of the achieved user throughput with the different methods. The curves corresponding to EP-PF-AR and EP-PF-AR-CR

demonstrate that the proposed technique aiming at reducing the complexity of the scheduling does not entail any penalty in performance. The proposed Waterfilling-PF leads to a performance improvement for 5 users among 10, when compared to the EP-PF-AR-CR, with a gain ranging from 0.2 Mbps to 0.6 Mbps. Fig. 1 shows that the users achieving relatively small data rates exhibit an improvement in rate, whereas those achieving relatively high data rates exhibit a rate loss. This is due to the fact that when the power is allocated using our iterative waterfilling approach, the probability of scheduling users having a small historical throughput is increased. Besides, the power repartition is done in a way to boost the allocated power of disfavored or far from BS users, compared to an equal-power repartition strategy. This comes at the expense of a decrease ranging from 0.2 Mbps to a little more than 4 Mbps in the throughput of 5 users having already achieved rates higher than 30 Mbps, corresponding typically to close-to-BS users. It should be noted that in a MIMO system, the users close to the BS have generally good channel conditions and hence benefit more from the spatial diversity provided by the MIMO configuration. This fact is further validated in Fig. 2

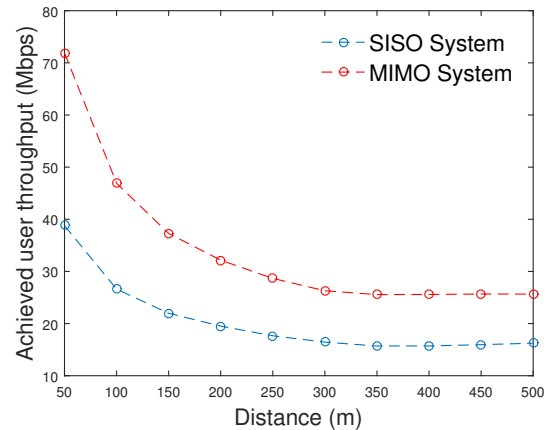


Fig. 2. Comparison of the achieved user throughput between a MIMO and a SISO system in terms of the users' distances, for $K = 10$ and $S = 64$.

which displays the achieved user throughput for a MIMO and a SISO system when the EP-PF-AR-CR is used, in terms of their distance from the BS. This was also observed in [16], where it was shown that in a high SNR regime, the ergodic capacity achieved by a MIMO system is much larger than the one achieved by a SISO system, in contrast to the low SNR regime where both systems perform similarly. This fact results in a decrease of the user fairness when the conventional PA is adopted. The main benefit of our approach is the improved user fairness.

Fig. 3 shows the Gini Fairness index in terms of the number of subbands (from 8 to 64), for a system consisting of 10 users and 20 users respectively. We can observe that the waterfilling-based power allocation significantly improves the fairness of the system when compared to the equal power repartition. We also see that user scheduling using EP-PF-AR-CR gives almost the same performance as EP-PF-AR, with a lower complexity, which justifies its use.

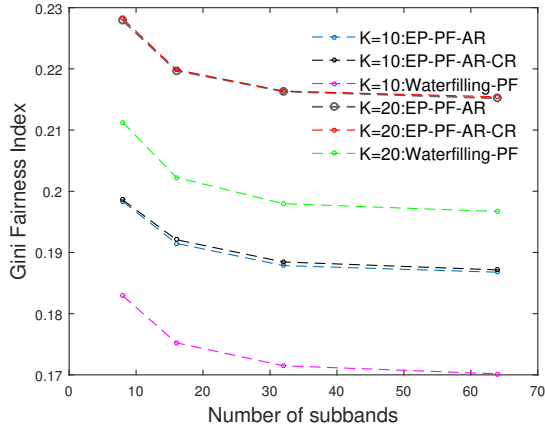


Fig. 3. Gini Fairness index in terms of the number of subbands, in a system consisting of 10 and 20 users.

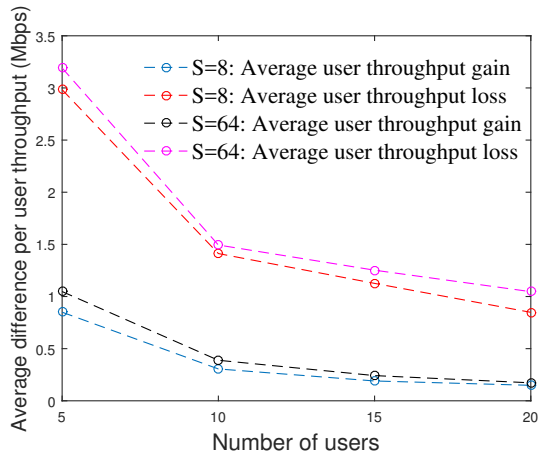


Fig. 4. Average gain and loss of the achieved user throughput of Waterfilling-PF over EP-PF-AR-CR in terms of the number of users in the system.

As already mentioned, the throughput gain achieved by Waterfilling-PF for the distant users in the cell comes at the cost of a loss experienced by the users close to the BS. In Fig. 4, we display the absolute value of the average difference in achieved user throughput between Waterfilling-PF and EP-PF-AR-CR, when the number of users in the system varies from 5 to 20. As expected, both the average rate loss of close-to-BS users and the average rate gain of far-from-BS users decrease when K increases, because a larger number of users are competing for system resources. In addition, both increase when S increases to 64, while the user fairness is better for $S = 64$, compared to $S = 8$. Although the average throughput loss of close-to-BS users is generally larger than the average gain of far-from-BS users, the users suffering from this loss already have very high throughputs as previously mentioned, whereas the users benefiting from Waterfilling-PF have relatively low achieved throughputs, hence will be affected positively.

This observation is validated in Table II, which provides the achieved throughput per user in a system consisting of 10 users for $S = 8$ and $S = 64$. We can observe that the closest user to the BS suffers from the greatest loss; however, even with Waterfilling-PF its achieved throughput is still very high. As

TABLE II
ACHIEVED USER THROUGHPUT (Mbps) FOR EP-PF-AR-CR AND WATERFILLING-PF (WF-PF) WITH $S = 8$ AND 64, FOR $K = 10$ USERS

| Scheduling | U1 | U2 | U3 | U4 | U5 | U6 | U7 | U8 | U9 | U10 |
|-------------------------|------|------|------|------|------|------|------|------|------|------|
| $S = 8$ EP-PF-AR-CR | 70.6 | 45.5 | 35.7 | 31.1 | 27.5 | 24.9 | 24 | 24 | 23.9 | 23.6 |
| $S = 8$ WF-PF | 66.4 | 44.1 | 35.1 | 30.8 | 27.3 | 25.1 | 24.3 | 24.3 | 24.3 | 24.2 |
| $S = 64$ EP-PF-AR-CR | 71.8 | 46.9 | 37.2 | 32.1 | 28.6 | 26.3 | 25.5 | 25.5 | 25.6 | 25.6 |
| $S = 64$ WF-PF | 67.3 | 45.3 | 36.5 | 31.8 | 28.6 | 26.5 | 25.9 | 26 | 26.1 | 26.2 |

the achieved throughputs decrease when moving away from the BS, Waterfilling-PF outperforms EP-PF-AR-CR, and the farthest users benefit from an increase in throughput.

TABLE III
PERFORMANCE GAIN OF WATERFILLING-PF OVER EP-PF-AR-CR FOR DIFFERENT NUMBERS OF USERS AND SUBBANDS

| Number of users/ Subbands | % of users with gain in rate | % rate gain for far-from-BS users | % rate loss for close-to-BS users | % rate gain for cell-edge user | % gain in Gini user fairness |
|---------------------------|------------------------------|-----------------------------------|-----------------------------------|--------------------------------|------------------------------|
| $K = 5, S = 8$ | 60 | 2 | 3.5 | 3.1 | 11.3 |
| $K = 5, S = 64$ | 60 | 2.3 | 3.7 | 3.3 | 12.8 |
| $K = 10, S = 8$ | 50 | 1.3 | 2.6 | 2 | 7.9 |
| $K = 10, S = 64$ | 50 | 1.5 | 2.6 | 2.2 | 9.1 |
| $K = 20, S = 8$ | 50 | 1.2 | 2.6 | 2 | 7.4 |
| $K = 20, S = 64$ | 55 | 1.3 | 3 | 2 | 8.7 |

Table III gives the performance of Waterfilling-PF in comparison to EP-PF-AR-CR for different combinations of K and S . For all cases under consideration, at least half the users in the cell benefit from the proposed scheduling technique. Table III also shows the average percentage rate gain and loss for users who have their throughputs respectively increased and decreased with Waterfilling-PF. The percentage gain of the cell-edge user who normally has the lowest achieved throughput is also given. The values in the last column of Table III confirm that Waterfilling-PF achieves better fairness between the users when compared to EP-PF-AR-CR, for different values of K and S .

To further illustrate the advantage of the coupling of NOMA with the waterfilling based PA for the enhancement of the cell-edge user throughput, we show this metric in Fig. 5, for a system consisting of 10 users, employing NOMA (i.e. $N_s = 1$ or 2) and OMA (i.e. $N_s = 1$), respectively. The curves EP-PF-OMA and WF-PF-OMA denote the achieved cell-edge throughput by a MIMO system employing OMA, where the former adopts an equal power repartition between subbands and transmit antennas, while the latter is based on the waterfilling approach presented in III-B.

The use of NOMA greatly improves the throughput achieved by the cell-edge user. The gain can reach 26% for both PA methods. We can also see that using the waterfilling based power allocation benefits both the OMA and the NOMA schemes. As a side note, we also measured the additional computational load incurred by incorporating the waterfilling PA within the PF, in terms of the average execution time per allocation (i.e. in one time slot). When $S = 8$ (resp. $S = 64$), for the OMA allocation, the execution time per time slot increases by a factor of 1.8 (resp. 2.1) whereas for the NOMA case, it increases by a factor of 1.7 (resp. 1.8).

Finally, we tested the performance of Waterfilling-PF when

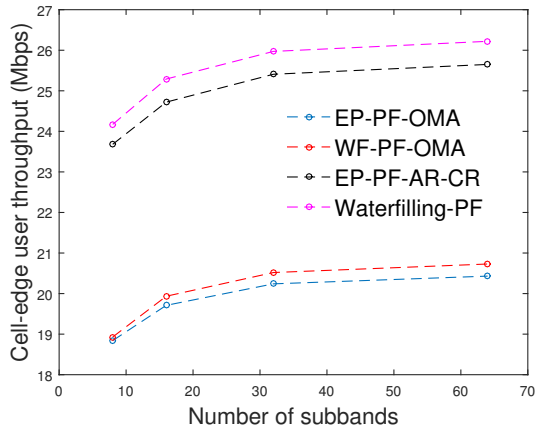


Fig. 5. Achieved throughput for the cell-edge user in terms of the number of subbands, where $K = 10$ users.

TABLE IV
PERFORMANCE GAIN OF WATERFILLING-PF OVER EP-PF-AR-CR FOR DIFFERENT VALUES OF THE CELL RADIUS

| Cell Radius | % rate gain for far-from-BS users | % rate loss for close-to-BS users | % cell-edge user gain |
|--------------------|-----------------------------------|-----------------------------------|-----------------------|
| $R = 500\text{m}$ | 1.5 | 2.6 | 2.2 |
| $R = 1\text{ km}$ | 2.2 | 4 | 5 |
| $R = 5\text{ km}$ | 8.5 | 9 | 14 |
| $R = 10\text{ km}$ | 32.8 | 22.3 | 42 |

the radius of the cell increases. In fact, as the cell gets larger, the channel conditions of certain users become worse and hence these users benefit more from Waterfilling-PF. Table IV gives the performance of Waterfilling-PF in comparison to EP-PF-AR-CR for different values of the cell radius, with $K = 10$ and $S = 64$. We can see that the application of Waterfilling-PF becomes more attractive as the cell gets larger, since the cell-edge user gain increases significantly with the cell radius and so does the percentage of average throughput gain for far-from-BS users. Although the loss experienced by certain users also increased, the average percentage gain greatly surpasses the percentage of loss for high values of the cell radius. Therefore, for scenarios where the cells are relatively large, the application of Waterfilling-PF is very beneficial to the users suffering from bad channel conditions. In particular, in certain scenarios, cooperative transmission between multiple BSs, employing a cell size adapting scheme, could be applied as was proposed in [17], to accommodate cell-edge users suffering from very bad channel conditions. For such scenarios, when a BS extends its cell radius, it becomes particularly appealing to use the waterfilling-based power allocation presented in this paper, since it enhances the cell edge user throughput.

V. CONCLUSION

In this paper, we have first proposed a method to reduce the complexity of the conventional equal-power PF scheduling algorithm. Simulation results show that the proposed complexity reduction method does not incur any degradation in performance, compared to the conventional PF. We have also proposed a generalization of the iterative waterfilling-based power allocation to the case of SU-MIMO NOMA, such that

to incorporate a joint inter-antenna and inter-subband unequal power allocation. This method allows the user fairness to be increased by improving the data rates achieved by the cell-edge users, either when compared to OMA signaling or to equal-power NOMA. The Waterfilling-PF technique was also shown to be more attractive for large cells. Future works consider an extension of the current study to the case of uplink transmission.

ACKNOWLEDGMENT

This work has been funded with support from the Lebanese University. Part of this work has been performed in the framework of the Horizon 2020 project FANTASTIC-5G (ICT-671660), which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in FANTASTIC-5G.

REFERENCES

- [1] LTE; TS 36.300 Version 8.9.0 Release 8, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description; Stage 2," 2009.
- [2] LTE; TR 36.913 Version 9.0.0 Release 9, "Requirements for further advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced)," 2010.
- [3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge University Press, 2005.
- [4] A. Benjebbour et al., "Concept and Practical Considerations of Non-orthogonal multiple access (NOMA) for future radio access," in *Proc. Int. Symp. on Intelligent Signal Process. and Commun. Syst. (ISPACS)*, Nov. 2013, pp. 770–774.
- [5] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2013, pp. 611–615.
- [6] M.R. Hojeij, J. Farah, C. Abdel Nour, and C. Douillard, "Resource Allocation in Downlink Non-Orthogonal Multiple Access (NOMA) for Future Communication Systems," in *Proc. IEEE 81st Vehicular Technology Conf. (VTC Spring)*, May 2015, pp. 1–6.
- [7] —, "New Optimal and Suboptimal Resource Allocation Techniques for Downlink Non-Orthogonal Multiple Access," *Wireless Pers. Commun.*, pp. 1–31, 2015.
- [8] X. Chen, A. Benjebbour, Y. Lan, A. Li, and H. Jiang, "Impact of Rank Optimization on Downlink Non-Orthogonal Multiple Access (NOMA) with SU-MIMO," in *Proc. IEEE Int. Conference on Communication Systems (ICCS)*, Nov. 2014, pp. 233–237.
- [9] —, "Evaluations of Downlink Non Orthogonal Multiple Access (NOMA) Combined with SU-MIMO," in *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2014, pp. 1887–1891.
- [10] E. Okamoto, "An Improved Proportional Fair Scheduling in Downlink Non-Orthogonal Multiple Access System," in *Proc. IEEE 82nd Vehicular Technology Conf. (VTC Fall)*, Sept. 2015, pp. 1–5.
- [11] M.R. Hojeij, C. Abdel Nour, J. Farah, and C. Douillard, "Waterfilling-based Proportional Fairness Scheduler for Downlink Non-Orthogonal Multiple Access," *IEEE Wireless Communications Letters*, Vol. 6, Issue 2, pp. 230–233, April 2017.
- [12] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System level performance of downlink non-orthogonal multiple access (NOMA) under various environments," in *Proc. IEEE 81st Vehicular Technology Conf. (VTC Spring)*, May 2015, pp. 1–5.
- [13] M.S. Maw and I. Sasase, "Resource Allocation Scheme in MIMO-OFDMA System for Users Different Data Throughput Requirements," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, March 2007, pp. 1706–1710.
- [14] M. Dianati, X. Shen, and S. Naik, "A new fairness index for radio resource allocation in wireless networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, March 2005, pp. 712–717.
- [15] J. Salo et al., "MATLAB implementation of the interim channel model for beyond-3G systems (SCME)," [Online]. Available: <http://www.tkk.fi/Units/Radio/scm/>, May 2005.
- [16] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Systems," [Online]. Available: http://web.stanford.edu/class/archive/ee/ee359/ee359.1062/cup_mimo.pdf.
- [17] X. Huang and N. Ansari, "Energy Sharing within EH-enabled Wireless," *IEEE Wireless Commun.*, vol. 2, pp. 144–149, 2015.