Joint Resource and Power Allocation Technique for Downlink Power-Domain Non-Orthogonal Multiple Access

Marie-Rita Hojeij, Charbel Abdel Nour, Joumana Farah, Catherine Douillard

To cite this version:

HAL Id: hal-01868764
https://hal-imt-atlantique.archives-ouvertes.fr/hal-01868764
Submitted on 5 Sep 2018

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.
Joint Resource and Power Allocation Technique for Downlink Power-Domain Non-Orthogonal Multiple Access

Marie-Rita Hojeij, Charbel Abdel Nour, Joumana Farah, and Catherine Douillard

Abstract—In this paper, a novel resource and power allocation technique is proposed for a power-domain non-orthogonal multiple access scheme using successive interference cancellation in a downlink cellular system. It aims at providing a flexible balancing between throughput and fairness maximization and incorporates unequal power repartition among allocated subbands using waterfilling. Its main strength is the ability to achieve a high level of fairness in every scheduling slot, therefore improving the quality of experience for all users. Simulation results show that the proposed metric enhances fairness and improves system capacity, when compared to the conventional proportional fairness scheduler.

Index Terms—Power-domain non-orthogonal multiple access, user scheduling, power allocation, fairness.

I. INTRODUCTION

With the proliferation of smart and machine-to-machine devices, it is expected that by 2023 the mobile traffic volume will be almost 10 times larger than today’s [1]. To satisfy such constraints, the concept of Power-Domain Non-Orthogonal Multiple Access (PD-NOMA) has emerged a few years ago as a promising candidate for future radio access [2]. PD-NOMA allows the cohabitation of multiple users per subband at the transmitter side on top of the Orthogonal Frequency Division Multiplexing (OFDM) layer, using the superposition coding principle [3]. At the receiver side, signal decoding relies on Successive Interference Cancellation (SIC) [4].

The majority of existing works dealing with PD-NOMA have investigated the system-level performance in terms of system capacity and cell-edge user throughput, such as [5], [6]. They show that, by solving the user pairing, subband assignment and power allocation problems appropriately, PD-NOMA can provide a significant improvement compared to Orthogonal Multiple Access (OMA).

Aiming at further enhancing system capacity and user fairness, we propose several new resource allocation as well as power allocation techniques well suited for PD-NOMA systems. This paper is organized as follows: In Section II, we introduce the general system model and give a brief description of the conventional algorithm used for user scheduling. Section III details the proposed joint resource and power allocation schemes. Simulation results are given in Section IV, while Section V concludes the paper.

M. Hojeij, C. Abdel Nour and C. Douillard are with IMT Atlantique, Brest, France (marie.hojeij, charbel.abdelnour, catherine.douillard@imt-atlantique.fr). M. Hojeij is also with the Holy-Spirit University of Kaslik, Jounieh, Lebanon. J. Farah is with the Lebanese University, Roumieh, Lebanon (joumana farah@ul.edu.lb).

II. SYSTEM DESCRIPTION

A. Description of the Basic PD-NOMA Scheme with SIC

In this section, we describe the basic concept of PD-NOMA including user multiplexing at the base station (BS) transmitter and signal separation at the user terminal.

In this paper, a downlink air interface with a single transmitter and a single antenna receiver is considered. The system consists of $K$ users per cell, and a total bandwidth $B$ divided into $S$ subbands. Among the $K$ users, a set of users $U_s = \{k_1, k_2, ..., k_n, ..., k_{n(s)}\}$ is selected to be scheduled on each frequency subband $s$, $(1 \leq s \leq S)$. The $n$-th $(1 \leq n \leq n(s))$ user scheduled at subband $s$ is denoted by $k_n$, and $n(s)$ is the number of users non-orthogonally scheduled at subband $s$. At the BS transmitter side, the information sequence of each scheduled user at subband $s$ is independently coded and modulated, resulting into symbol $x_{s,k_n}$ for the $n$-th scheduled user. Therefore, the signal $x_s$, transmitted by the BS on a certain subband $s$, represents the sum of the coded and modulated symbols of the $n(s)$ scheduled users:

$$x_s = \sum_{n=1}^{n(s)} x_{s,k_n}, \text{ with } E[|x_{s,k_n}|^2] = P_{s,k_n}$$

where $P_{s,k_n}$ is the power allocated to user $k_n$ at subband $s$. The received signal of user $k_n$ at subband $s$, $y_{s,k_n}$, can be written as:

$$y_{s,k_n} = h_{s,k_n} x_{s,k_n} + w_{s,k_n}$$

where $h_{s,k_n}$ is the channel coefficient between the BS and user $k_n$, at subband $s$. $w_{s,k_n}$ indicates the Gaussian noise plus inter-cell interference received by user $k_n$ on subband $s$. Let $P_{\text{max}}$ be the BS maximum transmit power; the sum power constraint is formulated as:

$$\sum_{s=1}^{S} \sum_{n=1}^{n(s)} P_{s,k_n} = P_{\text{max}}$$

At the receiver side, we consider SIC processing. The optimal order for user decoding is in increasing order of the channel gains observed by the users, normalized by the noise and inter-cell interference $h_{s,k_n}^2/n_{s,k_n}$, where $h_{s,k_n}^2$ is the equivalent channel gain and $n_{s,k_n}$ the average power of $w_{s,k_n}$. Therefore, user $k_n$ can remove the inter-user interference from user $k_j$ at subband $s$, provided that $h_{s,k_j}^2/n_{s,k_j}$ is lower than $h_{s,k_n}^2/n_{s,k_n}$, and the received signals from other users with higher $h_{s,k_j}^2/n_{s,k_j}$ are treated as noise.
Assuming successful decoding and no error propagation, the throughput of user $k_n$, at subband $s$, $R_{s,k_n}$, is given by:

$$R_{s,k_n} = \frac{B}{\log(2)} \left( 1 + \frac{h_{s,k_n,}^2 P_{s,k_n}}{\sum_{j \in U_s} h_{s,k_j}^2 P_{s,k_j} + n_{s,k_n}} \right)$$

(B. Conventional User Scheduling for PD-NOMA)

The Proportional Fairness (PF) scheduling algorithm is widely used in communication networks, since it achieves a balance between high throughput and user fairness. In [8], the operation of the PF scheduler is detailed: at the beginning of each scheduling slot, each user provides the base station with its channel state or its feasible rate. The PF algorithm keeps track of the historical rate of each user in a sliding window of length $t_c$. In scheduling slot $t$, a user is selected to be served based on:

$$\arg \max_k R_k(t) \frac{T_k(t)}{T_k(t)}$$

where $R_k(t)$ is the feasible rate of user $k$ in slot $t$.

An approximated version of the PF scheduler for multiple users transmission has been adopted in the majority of the works dealing with NOMA, such as [9], for selecting users to be non-orthogonally scheduled on the available resources. For each subband $s$, the set of scheduled users $U_s$ is chosen in such a way to maximize the PF scheduling metric :

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

$U$ is a possible candidate user set, $R_{s,k} U(t)$ is the instantaneous achievable throughput of user $k$ on subband $s$ at time $t$. $R_{s,k} U(t)$ is calculated based on Eq. 4, whereas $T_k(t)$ is updated according to:

$$T_k(t + 1) = \left( 1 - \frac{1}{t_c} \right) T_k(t) + \frac{1}{t_c} \sum_{s=1}^{S} R_{s,k}(t)$$

The throughput averaging time window, $t_c$, is chosen such that a tradeoff between system performance in terms of fairness and system capacity is guaranteed.

III. PROPOSED RESOURCE AND POWER ALLOCATION TECHNIQUE FOR AN ENHANCED PD-NOMA SYSTEM

The objective of the PF is to maximize the long-term averaged user rates, thus targeting long-term fairness. However, short-term fairness and fast convergence towards required performance are an important issue to be addressed in upcoming mobile standards [1]. Since the PF scheduler gives each user a priority inversely proportional to its average throughput, and since users instantaneous rates are not considered in the scheduling metric, it happens that some users do not gain access to available resources whenever their historical rates are high. This can be problematic for applications requiring a quasi-constant QoE, such as multimedia transmissions. Moreover, the PF scheduler does not take advantage of the fact that the sum rate of NOMA systems increases when increasing the difference in channel gains of paired users. The resource allocation technique described hereafter aims at addressing these issues.

In addition, in the majority of papers dealing with PD-NOMA, such as [4], an equal repartition of power among subbands is considered to simplify the scheduling problem. In our work, power is assigned to the subbands based on the waterfilling principle, since it allows the achieved system throughput to be maximized [5].

The design steps of the proposed resource and power allocation technique are detailed in the following subsections.

A. Initialization and Priority Assignment

At the beginning of the process, allocated powers and rates are all set to zero. Therefore, users are first chosen based on the priority constraint proposed in previous works [7], [10]. Priorities are assigned in the BS based on the channel coefficients experienced by users on available subbands. For each user $k$, its highest channel coefficient, $h_{s_{max},k}$, is selected among the channel gains experienced over all subbands. Then, the users are processed in the increasing order of $h_{s_{max},k}$ values: user $k_{min}$ with the lowest value of $h_{s_{max},k}$ is selected first while user $k_{max}$ with the highest value $h_{s_{max},k}$ is assigned last.

B. First User Selection and Subband Assignment

After the initialization phase, unlike for PF scheduling, user rates already granted during the current allocation step are also taken into consideration. User $k_1$ is selected among the set of users that need to communicate as the user experiencing the lowest total achieved throughput. User selection can be written as follows:

$$k_1 = \min_k \left( \left( 1 - \frac{1}{t_c} \right) T_k(t) + \frac{1}{t_c} \sum_{s=1}^{S} R_{s,k}(t) \right)$$

$R_{s,k}(t)$ is calculated based on Eq. 4, and can be equal to zero if user $k$ is not scheduled on subband $s$.

User $k_1$ is then attributed the most favorable subband, i.e., the one with the highest channel coefficient, denoted by $s_f$.

C. User Pairing and Power Allocation

We have limited our study to the case where the number of scheduled users per subband is equal to 2 ($n(s) = 2$), since it was shown in [11] that the improvement in cell throughput is only 1% when the number of scheduled users per subband is equal to 3, compared to the case where it is equal to 2. However, the proposed pairing technique described in the following has the ability to consider more than 2 users per subband if needed.

Having determined the subband $s_f$ attributed to selected user $k_1$, the following step consists in selecting the best user $k_2$ to be paired with $k_1$. The selection of this second user is
based on metric calculations requiring rate estimates, which in turn include power estimates.

The implementation of an optimal Power Allocation (PA) scheme based on a waterfilling process jointly performing intra and inter-subband PA is quite complex since it requires the resolution of a non-convex optimization problem [7]. For this purpose, we have proposed a sub-optimal PA technique using an iterative waterfilling process [12], where only the user with the highest channel coefficient among scheduled users is considered in the inter-subband PA step. This technique was applied in the current work, combined with the Fractional Transmit Power Allocation (FTPA) technique [13] for intra-subband PA between the two paired users.

As for the second user assignment, we have proposed in [14] several weighted PF metrics that aim at (i) enhancing the user capacity, thus increasing the total achieved user throughput, (ii) reducing the convergence time towards required fairness performance, (iii) enhancing fairness among users (both long-term and short-term fairness), (iv) limiting the fluctuations of user data rates, and (v) incorporating the delivery of different levels of QoS. Later, we have also introduced a new metric called Flexible Throughput vs. Fairness Maximization Metric (FTFMM) [15] in order to cope with applications where throughput versus fairness maximization is essential. FTFMM aims at striking a flexible balancing between throughput and fairness, since it is based on the instantaneous rate on the considered subband \(s_f\), the current user rates on the so far attributed subbands within the current allocation slot and on the historical rates of each possible candidate. A user \(k_2\) is selected to be paired with user \(k_1\) on the considered subband \(s_f\) based on:

\[
k_2 = \max_k \frac{R_{s_f,k} + R_{s_f,k_1}}{aR_{k_1,tot} + bR_{k_1,tot}}, \quad a + b = 1
\]  

where \(R_{s_f,k}\) is the data rate achieved by user \(k\) over subband \(s_f\), calculated from Eq. 4.

Parameters \(a\) and \(b\) can take values between 0 and 1 and are used to provide flexibility between throughput and fairness maximization. If \(a\) is set to 1 and \(b\) to 0 in Eq. 9, user \(k_2\) is selected in a fair manner by taking the instantaneous rates as well as the total user rates of the possible candidates into consideration. On the contrary, if \(a\) is set to 0 and \(b\) to 1, user \(k_2\) is selected to be paired with \(k_1\) according to its instantaneous rate achieved over \(s_f\), and therefore the choice is only based on the rate maximization on the current subband. Neither his historical rate nor his current rate have been taken into account. In such a situation, the notion of fairness is completely flawed. Between these two extreme cases, and depending on the requirement of the application, parameters \(a\) and \(b\) can be varied in such a way to favor either the fairness or the throughput or to ensure balance between them.

D. Adaptive Switching to Orthogonal Signaling

The estimation of power conducted within each stage of the allocation process is considered as temporary until the final stage is reached. At the final stage, once all subbands have been allocated, power is eventually distributed on each subband using the waterfilling principle and data rates achieved by users are computed accordingly.

IV. NUMERICAL RESULTS

A. System Model Parameters

Extensive system level simulations were carried out in order to evaluate the proposed allocation technique. The parameters considered in this work are based on existing LTE/LTE-Advanced specifications. The maximum BS transmission power is 46 dBm. The system bandwidth is 10 MHz divided into 8, 16, 32, 64, or 128 subbands, with a carrier frequency of 2 GHz. The noise power spectral density is \(4.10^{-18}\) mW/Hz. Users are randomly deployed in a cell of radius 500 meters, with a minimum distance of 35 meters between users and BS. Distance-dependent path loss is considered with a decay factor of 3.76. The Extended Typical Urban (ETU) channel model is used, with a mobile velocity of 50 km/h. We assume a \(t_c\) window of 100 ms (100 subframes of 1 ms). Perfect channel estimation is assumed. The maximum number of scheduled users per subband is 2 \((n(s) = 2)\). The FTPA technique is applied for intra-subband PA with coefficient \(a\) equal to 0.7 [13].

B. Performance Evaluation

In this study, we mainly consider three system-level performance indicators: achieved system capacity, long-term fairness, and short-term fairness. In order to evaluate the achieved fairness performance, we use the Gini fairness index defined in [16]:

\[
G = \frac{1}{2K^2\bar{r}} \sum_{k' = 1}^{K} \sum_{k'' = 1}^{K} |r_{k'} - r_{k''}|
\]

with

\[
\bar{r} = \frac{\sum_{k = 1}^{K} r_k}{K}
\]

Gini fairness index takes values between 0 and 1, where \(G = 0\) corresponds to the maximum level of fairness among users, while a high value value of \(G\) indicates that the resource allocation scenario is highly unfair. When fairness among users is to be evaluated within each scheduling slot, short-term fairness is considered and \(r_k\) is equal to \(R_k(t)\), the actual achieved throughput of user \(k\) during scheduling slot \(t\).

Fig. 1 compares the capacity of a NOMA system using the approach proposed in [15] with OMA and NOMA systems based on conventional PF scheduling, as a function of the number of subbands. When different couples of parameters \(a\) and \(b\) values are evaluated, as expected, the capacity is maximized for \(a = 0\) and \(b = 1\) (upper throughput bound), decreases progressively when \(a\) increases and \(b\) decreases, and reaches its lowest value for \(a = 1\) and \(b = 0\). A gain in performance is observed with the proposed metric, when compared to a NOMA-based classical PF, since the proposed metric tries to provide rates to all users, even those
when a number of subbands. The highest level of fairness is achieved in NOMA systems based on PF scheduling, as a function of the number of subbands. Using the proposed approach with conventional OMA and classical PF, depending on the number of subbands. This is in compliance with Eq. 9, where total achieved rates and current rates are both considered for user pairing, thus providing a high level of fairness.

V. Conclusion

In this paper, we described several low-complexity enhancements to the resource and power allocation techniques of PD-NOMA systems. Simulation results show that the proposed scheme allow a significant increase in the system capacity and long-term fairness, when compared to a conventional PF scheduler for PD-NOMA. Furthermore, although not shown in this paper, it was also observed that the proposed technique achieves a higher level of fairness within each scheduling slot, which improves the QoE of each user.

ACKNOWLEDGMENT

This work has been supported by the PHC CEDRE program.

REFERENCES