Multi-User collaborative scheduling in 5G massive MIMO heterogeneous networks

Marie Masson, Zwi Altman, Eitan Altman

Marie Masson, Zwi Altman, Eitan Altman. Multi-User collaborative scheduling in 5G massive MIMO heterogeneous networks. IFIP Networking 2020, Jun 2020, Paris / Online, France. hal-02566253

Multi-User collaborative scheduling in 5G massive MIMO heterogeneous networks

Marie Masson
Orange Labs, LIA, Inria
Châtillon, France
marie.masson@orange.com

Zwi Altman
Orange Labs
Châtillon, France
zwi.altman@orange.com

Eitan Altman
INRIA Sophia-Antipolis, LIA Avignon University and LINCS Paris
France
eitan.altman@sophia.inria.fr

***Abstract*—Macro cell densification with Small Cells (SCs) is an effective solution to cope with traffic increase. To fully benefit from the additional SCs capacity, interference mitigation techniques are needed. Densification in 5G networks with Massive Multiple Input Multiple Output (M-MIMO) deployment needs to rethink interference mitigation to account for highly focused beams and Multi-User (MU) scheduling. This paper presents a low complexity collaborative Proportional Fair (PF) based scheduling that maximizes the throughput and improves fairness of the heterogeneous network. The solution is based on the calculation of a loss factor indicator that each SC provides to the macro cell at each scheduling period. These indicators allow the macro cell MU scheduler to efficiently select the set of users for scheduling, leading to a significant improvement in performance. Numerical results illustrate the interest of the collaborative solution.**

***Index Terms*—Collaborative scheduling, Small Cells, densification, Massive MIMO, Multi-User MIMO, 5G**

I. INTRODUCTION

5G targets a significant leap forward in user experience in terms of data rates, latency, support of very high speed and massive connectivity [1]. Enhanced user throughput and network capacity can be achieved thanks to 5G technologies such as M-MIMO and SC densification [2] [3].

M-MIMO enhances Signal to Interference plus Noise Ratio (SINR) using beamforming. It allows better focusing of the radiated power on the user, reduced interference generated on neighboring cells, and spacial multiplexing of users known as Multi-User MIMO (MU-MIMO). However, significant interference can occur at cell edge in macro cells with M-MIMO deployment, or when SCs are deployed within the macro cell coverage. In this case, interference mitigation techniques are necessary, some of which are briefly summarized presently.

In 4G, the enhanced Inter Cell Interference Coordination (eICIC) feature was introduced to enhance capacity and mitigate interference in heterogeneous networks with SCs deployment [4] [5]. The interference reduction is obtained by means of the Almost Blank Subframe (ABS) feature that mutes almost all macro Base Station (BS) transmissions on a certain portion of the subframes, thus letting SCs serve their users with reduced interference. However with M-MIMO, only part of the users are served simultaneously by the focused beams, thus making ABS of limited use.

Interference mitigation can be fully performed in the physical layer via more complex transceivers with advanced precoding. The BS can then precode the Down Link (DL) signals for its users while being orthogonal to the dominating subspace of the interfering signals from SCs [6].

Interference between a macro cell and SCs can be minimized using Joint Processing (JP) Coordinated Multipoint transmission (CoMP) [7]. This scheme has been studied in the framework of heterogeneous networks with M-MIMO deployment [2]. The CoMP approach imposes strict requirements on the network such as synchronization between cooperating cells, high backhaul capacity and low latency transmissions which impact network cost.

The purpose of this paper is to propose a low complexity collaborative scheduling for a M-MIMO cell with SCs in its coverage area. It is assumed that the macro cell utilizes Grid of Beams (GoB) [8] for the control channels. At each Transmission Time Interval (TTI), each SC sends a signal comprising a single number to the macro cell. The received signals are used by the macro cell MU scheduler to maximize a global PF utility of the heterogeneous cell and to mitigate the interference that the macro cell produces. It is noted that the virtualization of the 5G Radio Access Network (RAN) motivates this approach. The Media Access Control (MAC) layer is expected to be virtualized within a centralized data center or a Mobile Edge Computing (MEC), thus simplifying the exchange of information between network nodes and alleviating requirements on the backhaul and associated costs.

The paper is organized as follows. Section II presents the system model. Section III describes the scheduling algorithms including the macro cell MU scheduling under beam-skipping constraints, and the collaborative PF scheduling. Numerical results are presented in Section IV followed by concluding remarks in Section V. Throughout the paper, the terms pico cell and SC have the same meaning and are used interchangeably.

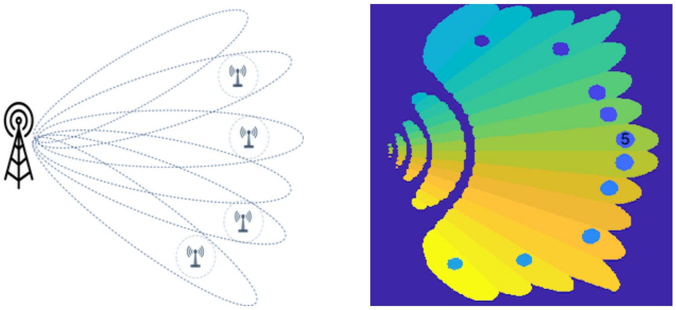
Notation: Scalars are denoted by lowercase italic letters a, vector columns are presented in bold lowercase letters $\mathbf{a}$ and matrices are denoted by bold uppercase letters $\mathbf{A}$. $\overline{a}$ denotes a time average and H the hermitian transpose. $card()$ denotes the cardinality.

II. SYSTEM MODEL

Consider a network with tri-sectorial sites with M-MIMO system in the DL. GoB is used for control, including initial access and synchronization and defines the macro cell area.

BS data is served using Time Division Duplex (TDD) with Maximum Ratio Transmission (MRT) beamforming [9]. It is noted that the collaborative scheduling is independent of the beamforming implementation, and can be directly applied also to Frequency Division Duplex (FDD).

Each sector m is equipped with a Large Scale Antenna System (LSAS) with $N = N_x \times N_z$ radiating elements and is serving mobiles, each with one receiving antenna [10]. Pico cells with omnidirectional antennas are deployed at hotspot zones in order to offload the macro cell traffic and thus increase the cell capacity, as shown in Fig. 1. Neighboring interfering cells are omitted. Fig. 1b presents the coverage area, namely the best server map. The color code is used for visualization purpose only.



(a) Pico cell deployed close to cell edge

(b) Macro cell and SCs coverage

Fig. 1: Network model

Denote by the subscripts m, p and b, the macro-, the pico BSs and the beam of the GoB (denoted hereafter as *control beam*), respectively. The number of users of a cell c is denoted by N_c. We assume that the beams are ordered. Two beams b_i and b_j are adjacent if $|i-j| = 1$. k is the maximum number of users that can be scheduled simultaneously by the MU scheduler and is smaller or equal to the number of beams in GoB. p_c denotes the transmission power of a cell c. Denote by $\mathbf{h}_u^m$ the channel response from the macro cell m to the user u, and by g_u^p the gain from a pico cell p to a user u. $\mathbf{c}_u$ is the precoding matrix, defined as:

$$\mathbf{c}_u^m = \mathbf{h}_u^m \Gamma, \tag{1}$$

where Γ is a normalizing factor such that $||\mathbf{c}_u^m||^2 = \frac{1}{k}$. We denote by σ the thermal noise.

Users with elastic traffic arrive in every cell according to a Poisson process. Each user downloads a file and leaves the cell afterwards. The size of the file varies according to a Gaussian distribution.

The interference created by a pico cell p on the user u is written as:

$$I_u^p = p_p |g_u^p|^2 \tag{2}$$

The interference generated by a macro cell m' on u reads:

$$I_u^{m'} = p_{m'} \sum_{u' \in m', u' \neq u} |(\mathbf{h}_u^{m'})^{\mathsf{H}} \mathbf{c}_{u'}^{m'}|^2 \tag{3}$$

We can express the SINR of a user u from a macro cell m [11] [6]:

$$S_u^m = \frac{p_m |(\mathbf{h}_u^m)^{\mathsf{H}} \mathbf{c}_u)|^2}{\sum\limits_{p \in \mathcal{P}} I_u^p + \sum\limits_{m'} I_u^{m'} + \sigma^2} \tag{4}$$

The SINR S_u^p of user u attached to the pico cell p can be written as:

$$S_u^p = \frac{p_p |g_u^p|^2}{\sum\limits_{p' \neq p} I_u^{p'} + \sum\limits_{m'} I_u^{m'} + \sigma^2} \tag{5}$$

III. MU SCHEDULING

In this section, we consider the PF scheduling framework for both pico cells and macro cells. At each time slot t_M the macro cell serves simultaneously up to k users. At most one user attached to a control beam b, denoted as *active user*, can be selected by the MU-scheduler. Denote by $\mathcal{B}_k$ the set of control beams having active users with $|\mathcal{B}_k| \leq k$. We first describe the concept of beam skipping. We then assume that $\mathcal{B}_k$ is given and derive the PF based scheduling utility function for the macro cell and the SCs [12]. Then we present the MU collaborative scheduling for the heterogeneous cell which relies on a selection algorithm for $\mathcal{B}_k$.

A. Beam skipping

Beam skipping is a beam selection technique used by the scheduler to optimize a Quality of Service (QoS) or a performance utility by reducing inter-beam interference. It is noted that beam skipping technique can be applied to both duplexing modes, TDD and FDD. Each User Equipment (UE) is attached to the control beam of the GoB providing it with the best SINR. The user attachment provides additional spatial information that is exploited by the scheduler to mitigate inter-beam interference: (i) by avoiding scheduling of two users attached to the same control beam, and (ii) by avoiding the scheduling of two users attached to adjacent beams of the GoB.

Given a performance utility, we can select a set of candidate UEs from a list while complying with a beam skipping constraint as follows. We select the top-ranked candidate UE, and remove the UEs attached to the same control beam and the adjacent ones from the candidate list. We repeat these two operations until we have selected k candidates or until the set of candidates is empty.

B. Macro cell utility

Denote by R_{u,t_M}^b the rate of user u attached to the beam b at time t_M, and by $\overline{R_{u,t_M}^b}$ the average rate of user u attached to beam b. The average rate at $t_{M+1}$is calculated using exponential moving average (or Abel average) with a small parameter ϵ [12].

$$\overline{R_{u,t_{M+1}}} = (1-\epsilon)\overline{R_{u,t_M}} + \epsilon R_{u,t_{M+1}} \tag{6}$$

Let $\mathcal{B}$ be the set containing all the beams b of the GoB of the macro cell m. We assume fading of different users is independent and for a given user, fading is independent at different allocation time t_M. While fading is assumed fixed during the transmission period, the coherence time is supposed not greater than $t_{M+1} - t_M$.

Denote by $U_m(t_M)$ the part of the utility function that depends on the past scheduling steps up to time t_M, and by ω_m the part that depends on the scheduling at time t_{M+1} and on the selected subset $\mathcal{B}_k$. We assume for the moment that $\mathcal{B}_k$ is given. The resulting utility is given by (7)-(9) [12]:

$$U_m(t_{M+1}) = U_m(t_M) + \omega_m(t_{M+1}) \tag{7}$$

where

$$U_m(t_M) = \sum_{b \in \mathcal{B}} \sum_{u \in b} (\log(\overline{R^b_{u,t_M}} + d) - \epsilon \frac{\overline{R^b_{u,t_M}}}{\overline{R^b_{u,t_M}} + d}) \tag{8}$$

$$\omega_m(t_{M+1}) = \epsilon \sum_{b \in \mathcal{B}_k} \max_{u \in b} (\frac{R^b_{u,t_{M+1}}}{\overline{R^b_{u,t_M}} + d}) \tag{9}$$

C. Pico cell utility

Similarly to the macro cell, we define the pico cell utility using the terms U_p and ω_p:

$$U_p(t_{M+1}) = U_p(t_M) + \omega_p(t_{M+1}) \tag{10}$$

where

$$U_p(t_M) = \sum_{u \in p} (\log(\overline{R^p_{u,t_M}} + d) - \epsilon \frac{\overline{R^p_{u,t_M}}}{\overline{R^p_{u,t_M}} + d}) \tag{11}$$

$$\omega_p(t_{M+1}) = \epsilon \max_{u \in p} (\frac{R^p_{u,t_{M+1}}}{\overline{R^p_{u,t_M}}}) \tag{12}$$

D. Collaborative scheduling

We define the global utility of the heterogeneous cell, U_{global} as the sum of utilities of the macro- and pico cells in its coverage area as defined in the previous two subsections. The collaborative scheduling aims at optimizing U_{global} under the beam skipping constraint at each time step. The optimization problem reads:

$$\begin{aligned} \underset{\mathcal{B}_k \in \mathcal{B}}{\text{maximize}} \; U_{global}(t_{M+1}) = U_m(t_M) + \sum_{p \in \mathcal{P}} U_p(t_M) \\ + \omega_m(t_{M+1}) + \sum_{p \in \mathcal{P}} \omega_p(t_{M+1}) \end{aligned} \tag{13}$$

$$\text{subject to } \forall (b_i, b_j) \in \mathcal{B}_k, |i - j| \geq 2$$

The problem (13) is a complex combinatorial optimization problem. The solution aims at maximizing the global throughput in a fair manner. A heuristic solution is proposed to allow real-time solution compatible with operational systems.

The scheduling decision for the time slot t_{M+1} impacts U_{global} via the terms $\omega_m(t_{M+1})$ and $\omega_p(t_{M+1})$. For sake of clarity, we denote by δ_b the utility gain term obtained by scheduling the macro cell user attached to the control beam b having the best PF criterion. The contribution of the beam b to the term $\omega_m(t_{M+1})$ is written as follows:

$$\delta_b = \epsilon \max_{u \in b} \frac{R^b_{u,t_{M+1}}}{\overline{R^b_{u,t_M}} + d} \tag{14}$$

Due to the small pico cells' size, we assume that a pico cell is located within the coverage area of one control beam of the GoB, and is interfered by the beam (denoted hereafter as *interfering beam*) scheduled to serve a macro cell UE attached to that control beam. The interfering beam comprizes the dominant contribution of the interference of a pico cell UE.

The interference experienced by pico cells' UEs from neighboring macro cells are taken into account while inter-pico cell interference is very low and is therefore neglected. When activated, the interfering beam reduces the rate of pico cell p and hence its utility given by (12). The idea of the collaborative scheduling is to take into account this utility reduction by the macro cell PF scheduler in order to optimize the global utility U_{global}.

The pico cell's UE u estimates the two following values: $R^{p,I}_{u,t_{M+1}}$, the rates received if its interfering beam is active and $R^{p,NI}_{u,t_{M+1}}$, the rate received otherwise (the superscripts I and NI stand for Interfered and Not Interfered, respectively). One of the two rates is calculated using actual received signals (which depend on the status of the interfering beam). The other one (denoted as the *complement rate*) is estimated as follows: samples of the interference signals in the presence and in the absence of the interfering beam are measured and stored once, and their average value is computed. Given the interference signal, the UE performs a simple test (e.g. comparison to a threshold or a Likelihood ratio test) and infers the status (active or inactive) of the interfering beam. The UE can then provide an estimate for the complement rate using the corresponding stored averaged interference value. These values are sent to the pico cell.

We define the quantity δ_p as the loss factor of pico cell p associated with the activation of its interfering beam. It is calculated as the difference between the PF criteria with and without the activation of the interfering beam:

$$\delta_p = \epsilon \max_{u \in p} (\frac{R^{p,I}_{u,t_{M+1}} - R^{p,NI}_{u,t_{M+1}}}{\overline{R^p_{u,t_M}}}) \tag{15}$$

This loss factor is transmitted to the macro cell scheduler by the different pico cells. We define the quantity Δ_b in (16) that the BS calculates to assess the contribution of b to the global utility taking into account the corresponding loss factor δ_p:

$$\Delta_b = \delta_b + \delta_p \tag{16}$$

We calculate Δ_b for all the beams. We proceed and derive the set $\mathcal{B}_k$ using the Δ_b-s and the beam skipping constraint (Section III-A), as described in Algorithm 1. The algorithm modifies the macro cell PF selection criterion to account for the pico cell loss factors. Once the set $\mathcal{B}_k$ is known, the user selection for the macro cell becomes straightforward. The pico cells' scheduler remain unchanged and is performed independently. It is recalled that each beam b serves at most one user at a time which is selected using a standard PF criterion applied to the users covered by that beam. Finally, the complexity of the algorithm is $\mathcal{O}(N_{users})$, where N_{users} is the total number of users in the heterogeneous cell.

Algorithm 1 MU Beam Selection

```
Init: $\mathcal{B}_k = \{\}$, $\mathcal{B}_{candidates} = \{b \in \mathcal{B} | \exists u \in b\}$
for all $p \in \mathcal{P}$ do
    Compute loss factor $\delta_p$ using (15)
    Send $\delta_p$ to macro cell $m$
end for
for all $b \in \mathcal{B}_{candidates}$ do
    Compute gain factor $\delta_b$ using (14)
    if $\exists p \in \mathcal{P}$ interfered by user attached to $b$ then
        $\Delta_b \leftarrow \delta_b + \delta_p$
    else
        $\Delta_b \leftarrow \delta_b$
    end if
end for
while card$(\mathcal{B}_k) < k$ or $\mathcal{B}_{candidate} \neq \emptyset$ do
    $b_{select} \leftarrow \text{argmax}_{b \in \mathcal{B}_{candidates}} \Delta_b$
    $\mathcal{B}_k \leftarrow \mathcal{B}_k \cup b_{select}$
    Define $\mathcal{N}(b_{select})$ the set containing $b_{select}$ and its adjacent beams
    $\mathcal{B}_{candidates} \leftarrow \mathcal{B}_{candidates} \backslash \mathcal{N}(b_{select})$
end while
```

IV. NUMERICAL RESULTS

A. Simulation scenario

Consider the network model of Section II, comprising a central sector with 10 pico cells (see Fig. 1), surrounded by 18 interfering sectors placed over a hexagonal grid. The performance results are extracted from the central sector. We compare the collaborative MU scheduling presented in Section III-D to a baseline without collaboration, namely the macro- and pico cells schedule their users independently according to the PF criterion δ_b given by (14). The simulation parameters are summarized in Table I.

B. Performance evaluation

We first compare the Mean User Throughput (MUT) in the macro cell and the pico cells for a fixed λ_m and for different λ_p (see Fig. 2). The load of the pico cells varies between 75 and 85 percent for λ_p equals 3 and 4 users/sec/cell respectively. The MUT gain achieved by the collaborative scheduler with respect to the baseline varies from 4 to 20 percent, with bigger gain for higher pico cell traffic. As the pico cell traffic increases, so does the number of users in the pico cell. Therefore, the more users benefit from the muting of a beam. Furthermore, unlike in 4G network, the spatial granularity of the muting is smaller: in a 4G network, either the whole macro cell is muted or it is not. Here, muting a beam means that other users in the macro cell, attached to different control beams, can still be served, allowing higher gains.

TABLE I: Network and Traffic characteristics

Network parameters	
Number of macro BSs m	1
Number of beams b in the GoB	16
Number of pico BSs p	10
Number of interfering macros	3x6 sectors
Intersite distance	500 m
Bandwidth	20 MHz
Channel characteristics	
Thermal noise	-174 dBm/Hz
Macro Path Loss (d in km)	128.1 + 37.6$\log_{10}(d)$ dB
Pico Path Loss (d in km)	140.7 + 36.7$\log_{10}(d)$ dB
Traffic characteristics	
Macro users' arrival rate λ_m	3.5 users/sec/cell
Pico users' arrival rate λ_p	$[3, 3.4, 3.8, 4]$ users/sec/cell
Macro/pico traffic spatial distribution	Uniform
Service Type	FTP
Average file size	5 Mbits

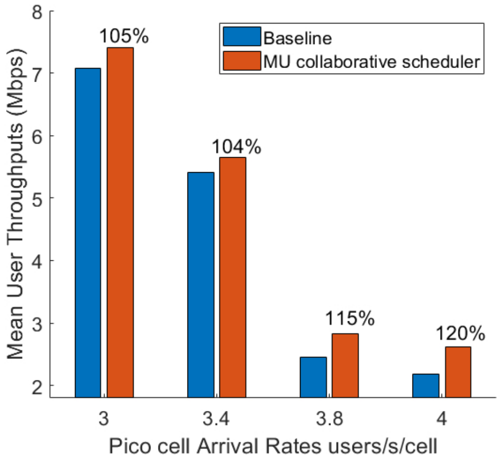


Fig. 2: MUT in the macro cell and its attached pico cells

Fig. 3 presents the CETs in cell 5 (see Fig. 1) as a function of the users' arrival rate in this cell. Here as well, the CET gain increases with the users' arrival rate. A throughput gain of around 17 percent with respect to the baseline is achieved, from 3.4 users/sec/cell and above.

The time variations of the filtered MUT and the CET are depicted in Figures 4 and 5 respectively. Performance gain is achieved for both MUT and CET using the collaborative scheduling, while the latter is particularly important: around 2 Mbps for the baseline and 2.4 Mbps and above for the collaborative scheduling.

V. CONCLUSION

This paper has presented a collaborative scheduling technique to manage 5G heterogeneous network with M-MIMO

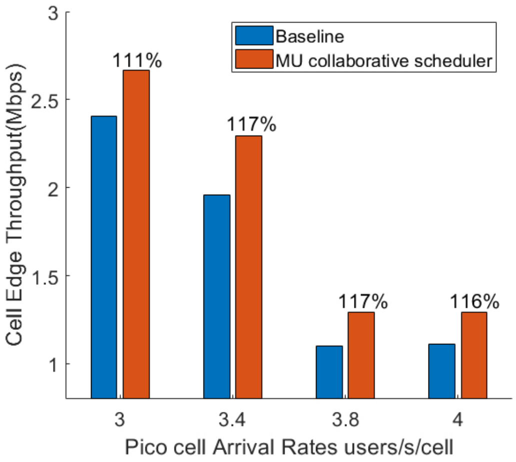


Fig. 3: Cell-Edge Throughput (CET) in pico cell 5

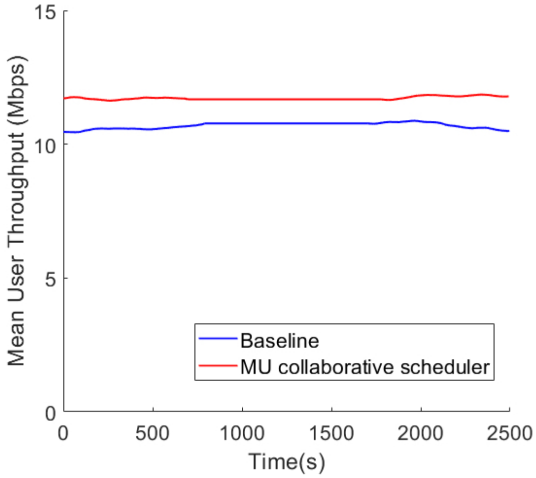


Fig. 4: MUT in pico cell 5 for $\lambda_p = 3.4$

macro cells and pico cells deployment. The solution is based on exchanging a loss factor that each pico cell provides to the macro cell, that estimates the performance utility loss due to the macro cell scheduling. While the optimization of MU scheduling is a combinatorial problem, it is shown that the the exchange of the loss factor allows to design a very low complexity scheduling algorithm.

The collaborative scheduling provides throughput gain that grows with the increase of traffic demand. In the simulations, the MUT gain of up to 20 percent is obtained for the heterogeneous cell comprising the macro- and pico cell. The proposed solution has been developed for TDD system. The solution is simpler for FDD system since the same beams are used for the transmission of both control and data signals.

REFERENCES

[1] S.-E. Elayoubi *et al.*, “5G innovations for new business opportunities,” 2017.

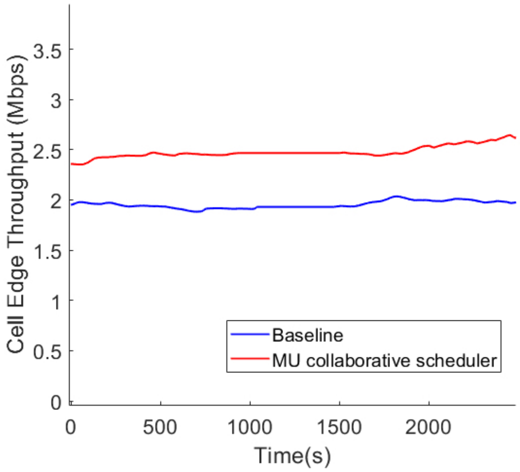


Fig. 5: CET in Pico Cell 5 for $\lambda_p = 3.4$

[2] V. Jungnickel *et al.*, “The role of small cells, coordinated multipoint, and massive MIMO in 5G,” *IEEE Communications Magazine*, vol. 52, no. 5, pp. 44–51, 2014.

[3] N. Bhushan *et al.*, “Network densification: the dominant theme for wireless evolution into 5G,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, 2014.

[4] Y. Wang and K. I. Pedersen, “Performance analysis of enhanced inter-cell interference coordination in LTE-Advanced heterogeneous networks,” in *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, 2012, pp. 1–5.

[5] A. Tall *et al.*, “Self organizing strategies for enhanced ICIC (eICIC),” *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2014 Proceedings of the 12th International Symposium on*, 2014.

[6] K. Hosseini *et al.*, “Massive MIMO and small cells: How to densify heterogeneous networks,” in *ICC 2013*, 2013, pp. 5442–5447.

[7] R. Irmer *et al.*, “Coordinated multipoint: Concepts, performance, and field trial results,” *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, 2011.

[8] F. W. Vook *et al.*, “MIMO and beamforming solutions for 5G technology,” in *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, 2014, pp. 1–4.

[9] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[10] A. Tall *et al.*, “Multilevel beamforming for high data rate communication in 5G networks,” *arXiv preprint arXiv:1504.00280*, 2015.

[11] A. Zappone *et al.*, “Online energy-efficient power control in wireless networks by deep neural networks,” in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018, pp. 1–5.

[12] R. Combes *et al.*, “Scheduling gain for frequency-selective Rayleigh-fading channels with application to self-organizing packet scheduling,” *Performance Evaluation*, Feb. 2011.