Satisfaction-Based Decentralized Interference Mitigation in Two-Tier Wireless Networks Using Equilibria

Pawel Sroka
Chair of Wireless Communications
Poznan University of Technology
Poznan, Poland
Email: pawel.sroka@put.poznan.pl

***Abstract*—In this paper the problem of interference mitigation in ultra dense heterogeneous cellular networks is considered, with the aim to maximize the satisfaction of users. Several game-theoretic approaches based on correlated equilibrium (CE) and satisfaction equilibrium (SE) are compared. Moreover, a new solution based on cooperative approach using CE and satisfaction-based optimization is proposed. Presented simulation results of an ultra dense 5G wireless system show, that the proposed scheme outperforms other solutions in terms of throughput satisfaction.**

I. Introduction

One of the main challenges in modern wireless networks is the interference, resulting from the densification of base stations (BSs) using the same frequency bands. Interference becomes a very severe limiting factor especially in case of multi-tier networks, where the macro-cells overlap the coverage of small access nodes, such as the pico or femto BSs. Interference management for heterogeneous networks has been studied as a part of 3GPP Long Term Evolution (LTE) Release 10 and beyond study items. The current 3GPP proposal adopts time-domain muting mechanisms known as almost blank sub-frames (ABSF) [1], [2]. The main idea of ABSF in the current standard is that macro-cell to small cell interference can be effectively reduced through muted time frames. However, the enhanced inter-cell interference coordination (eICIC) approach focuses mostly on improving the performance of cell-edge users, thus impacting the overall performance of the system. Therefore, an alternative approach based on game-theoretic solution has been proposed in [3], where a long-term interference mitigation scheme based on resource partitioning and power allocation has been proposed.
Game Theory has played a significant role recently in the analysis of many problems related to radio resource management (RRM) in wireless networks [4]–[7]. Such problems can be modelled using non-cooperative games when the network nodes act autonomously to maximize their benefits, such as the minimum throughput level. Commonly known example of such approach is the use Nash equilibrium (NE) [8], which represents the state where each network node cannot improve its utility by changing its actions when other nodes do not change their behaviour. However, in real communication systems the network nodes may be more interested in satisfying their quality of service (QoS) requirements, such as the minimum throughput or maximum delay, at a low cost rather than maximizing their data rate [7]. Therefore the concept of SE has been introduced, which represents the network state, where all nodes satisfy their QoS requirements, independently of the maximization of their utility [7]. However, in a large, ultra-dense wireless network, such as the one considered for 5G systems [9], SE may be unachievable, with the interference being the limiting factor.
An alternative approach to non-cooperative one may be to include nodes cooperation, using a concept of CE [10]. In this case all nodes choose their actions according to their observations of the environment, including the actions and benefits of other players. In case of CE approach, by taking into account the utility changes of neighbours, a joint optimization of the whole network performance is obtained. Depending on the definition of the payoff (utility) function, different approaches are possible. When using pure rate-based utility all nodes tend to maximize their throughput. On the other hand, when analysing the data rate or packet delays relative to the QoS requirements, SE behaviour can be achieved.
In this paper we compare the performance of a two-tier 5G wireless network using the interference mitigation based on the CE and SE approach. Moreover, we extend the CE solution proposed in [3] by defining the players payoffs based on their minimum throughput or maximum packet delay satisfaction. With the use of system-level simulations of a 5G wireless system we show that the CE approach outperforms the SE thanks to the use of cooperation.
The rest of the paper is organized as follows. Section 2 presents the considered system model and assumptions. In Section 3 the idea of SE is described, with game and payoffs definitions provided. Section 4 outlines the concept of CE and its applicability in terms of satisfaction provisioning. In Section 5 main simulation assumptions and results are presented, that indicate the advantages of CE approach. Finally, Section 6 draws the conclusions.

II. System model

In this work we assume a cellular wireless system using orthogonal frequency division multiple access (OFDMA) with M BSs, comprising both macro-cell BSs and micro-cell BSs operating in the same frequency band. For the purpose of this investigation we assume $M = 21$, with 3 macro-BSs and 18 micro-BSs, as shown in Fig.1, however, the considered solutions are applicable to a general case of heterogeneous network [3]. All BSs can exchange control and signalling

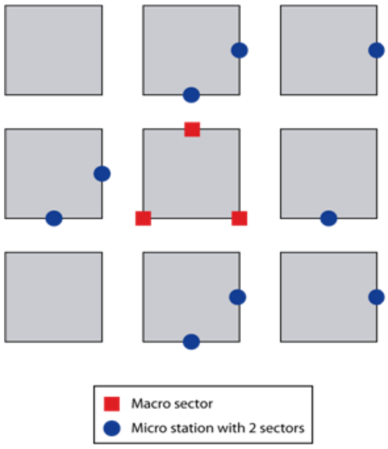

Fig. 1: BSs layout

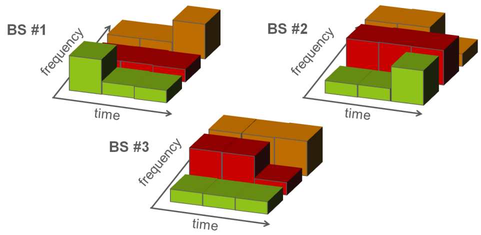

Fig. 2: Example of considered strategies

information using a dedicated interface, e.g. an optical fibre based backhaul network. The bandwidth available to each of the BSs is divided into time-frequency blocks, with the BSs transmitting in each of the blocks using one of the selected power per sub-carrier levels, selected out of the set $P = \{p_{\text{low}}, p_{\text{high}}\}$, as shown in Fig. 2. We assume that every BS selects a time-frequency pattern of powers from a discrete set, that comprises the so-called strategies (actions), with a given time interval. Let us denote the set of users as J, that are deployed uniformly on streets in the considered environment, with J_i denoting the set of users served by BS i. At each time interval, each BS divides the available resources among at most 10 user equipments (UEs) according to the proportional fairness (PF) rule. Let $|h_{i,j}^{(s)}|^2$ denote the channel gain between the i-th BS and j-th UE on sub-carrier s ($h_{i,j}^{(s)} \in \mathbb{C}$), and σ_j^2 be the noise variance at receiver j. The signal to interference plus noise ratio (SINR) for UE j served by BS i on sub-carrier s is given as follows:

$$\gamma_{i,j}^{(s)} = \frac{|h_{i,j}^{(s)}|^2 p_i^{(s)}}{\sigma_j^2 + \sum_{l \in M \setminus i} |h_{l,j}^{(s)}|^2 p_l^{(s)}}, \tag{1}$$

where $p_i^{(s)}$ denotes the transmit power of BS i on sub-carrier s. Let us assume that all BSs are interested in achieving at least the minimum throughput T_{min}. Let us define the throughput achieved by UE as follows:

$$T_j = \sum_t \sum_s R_j^s(t) \cdot x_j^s(t) \tag{2}$$

where $x_j^s(t)$ denotes the allocation of a sub-carrier s at time t to UE j, with $x_j^s(t) \in \{0,1\}$ and the rate of UE j on sub-carrier s - R_j^s, depends on the SINR values $\gamma_{i,j}^{(s)}$ through the use of adaptive modulation and coding (AMC). In this paper we consider the AMC mechanism proposed for LTE [1];

III. Satisfaction Equilibrium

A. Game and payoff definition

The process of learning SE can be described using the elements of the following game:

$$\mathbb{G} = (M, A, \{f_i\}_{i \in M}),$$

where M represents the set of players (BSs), A denotes the set of available actions, with $|A| = N$, and f_i is the satisfaction (QoS constraints) correspondence of player i, which indicates whether player is satisfied. According to [7], the correspondence can be defined as $f_i(\alpha_i, \boldsymbol{\alpha}_{-i}) = \{\alpha_i \in A : U_i(\alpha_i, \boldsymbol{\alpha_{-i}}) \geq \Gamma_i\}$ with $U_i(\alpha_i, \boldsymbol{\alpha_{-i}})$ representing player's observed utility when playing action α_i and Γ_i denoting the minimum utility level required by player i. A state of the game when all players satisfy their individual constraints simultaneously is referred to as satisfaction equilibrium (SE), that is defined as follows [7]:
An action profile α^+ is an equilibrium for the game $\mathbb{G} = (M, A, \{f_i\}_{i \in M})$ if

$$\forall i \in M, \alpha_i^+ \in f_k\left(\alpha_i^+, \boldsymbol{\alpha}_{-i}^+\right) \tag{3}$$

The existence of SE mainly depends on the set of imposed constraints on the utility function, with the necessary condition being the feasibility of the constraints.
For the considered scenario, where BSs act as game players, the satisfaction correspondence has to be modified to account for the satisfaction levels of all users served by the BS, as given below:

$$f_i(\alpha_i, \boldsymbol{\alpha_{-i}}) = \frac{1}{|J_i|} \sum_{j \in J_i} s_{i,j}(\alpha_i, \boldsymbol{\alpha_{-i}}), \tag{4}$$

where $s_{i,j}(\alpha_i, \boldsymbol{\alpha_{-i}})$ is the satisfaction of UE j when BS selects action α_i. The individual UE satisfaction can de defined using the binary representation:

- for rate-based satisfaction:

$$s_{i,j}(\alpha_i, \boldsymbol{\alpha_{-i}}) = \begin{cases} 1 & if\ T_j \geq T_{min} \\ 0 & otherwise \end{cases}, \tag{5}$$

  where T_j is the throughput of UE j and T_{min} is the minimum rate constraint.

- for packet delay-based satisfaction:

$$s_{i,j}(\alpha_i, \boldsymbol{\alpha_{-i}}) = \begin{cases} 1 & if\ D_j < D_{max} \\ 0 & otherwise \end{cases}, \tag{6}$$

  where D_j is the head-of-line packet delay of UE j and D_{max} is the maximum delay constraint (the deadline).

Alternatively, one can consider a relaxed version of individual UE satisfaction:

- for rate-based satisfaction using the sigmoid function:

$$s_{i,j}\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) = \begin{cases} 1 & if\, T_j \geq T_{min} \\ \frac{\exp(\beta\cdot(T_j-\epsilon\cdot T_{min}))}{1+\exp(\beta\cdot(T_j-\epsilon\cdot T_{min}))} & otherwise \end{cases}, \quad (7)$$

where β and ϵ are constants influencing the shape of the sigmoid.

- for packet delay-based satisfaction using the relaxed z-shaped function:

$$s_{i,j}\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) = \begin{cases} \frac{1}{1+\exp(\beta\cdot(D_j-\epsilon\cdot D_{max}))} & if\, D_j < D_{max} \\ 0 & otherwise \end{cases} \quad (8)$$

B. Learning satisfaction equilibrium

We assume that the game players undertake actions in consecutive time intervals, with only one action selected per interval. At each time interval player also observes whether it is satisfied or not. The selection of actions at each time interval is done based on probability distribution;

$$\pi_i(t) = \left(\pi_i^{(\alpha_i^{(1)})}(t), \pi_i^{(\alpha_i^{(2)})}(t), \ldots, \pi_i^{(\alpha_i^{(N)})}(t)\right),$$

which is known as probability distribution of exploration [7]. Under such assumptions the SE can be found using the behavioural rule which states that the next action taken by player i is as follows:

$$\alpha_i\left(t\right) = \begin{cases} \alpha_i\left(t-1\right) & if\, f_i(t-1) = 1 \\ \alpha_i\left(t\right) \sim \pi_i\left(t\right) & otherwise \end{cases} \quad (9)$$

The choice of probability distribution $\pi_i(t)$ may impact the convergence time and should also allow for exploration of all actions (thus all actions should have non-zero probability). A simple choice may be to use uniform probability distribution $\pi_i^{\left(\alpha_i^{(k)}\right)}(t) = \frac{1}{N}$. On the other hand, more sophisticated probability distribution update methods may be used that increase the convergence speed, e.g. based on the number of times an action has been selected previously [7].
The main problem with the learning solution presented above is that it neglects the utilities observed by players in the process of update of probability distribution. An alternative approach, where the decentralized optimization is performed using the modified behavioural rule that accounts for observed utilities has been proposed in [11]. This approach, known as the satisfaction equilibrium search algorithm (SESA) algorithm, utilizes the knowledge of individual utilities to increase the probability of selection of actions that provide higher payoff.

IV. Correlated equilibrium approach

A. Game and payoff definition

The problem of inter-cell interference mitigation can be described using the following normal-form game definition

$$\mathbb{G} = (M, A, \{U_i\}_{i\in M}),$$

where M represents the set of players.
Let us assume that at each time instant t BS i selects its action from a finite set A following a probability distribution

$$\pi_i(t) = \left(\pi_i^{(\alpha_i^{(1)})}(t), \pi_i^{(\alpha_i^{(2)})}(t), \ldots, \pi_i^{(\alpha_i^{(N)})}(t)\right),$$

where $\pi_i^{(\alpha_i^{(n)})}(t)$ denotes the probability that BS i plays action $\alpha^{(n)}$.
In general, at each time instant each BS plays one of N strategies $\alpha^{(n)}, 1 \leq n \leq N$. Therefore, assuming the set A is discrete and finite, at least one equilibrium exists that represents the system state when a player cannot improve its payoff (utility) when other players do not change their behaviour. Such a state is know as correlated equilibrium (CE), which is defined as follows

$$\sum_{\boldsymbol{\alpha}_{-i}\in A} \pi\left(\alpha_i^*, \boldsymbol{\alpha}_{-i}\right)\left(U_i(\alpha_i^*, \boldsymbol{\alpha}_{-i}) - U_i(\alpha_i', \boldsymbol{\alpha}_{-i})\right) \geq 0, \quad (10)$$
$$\forall \alpha_i', \alpha_i^* \in A, \forall i \in M,$$

In (10) $\pi\left(\alpha_i^*, \boldsymbol{\alpha_{-i}}\right)$ is the probability of playing strategy α_i^* in a case when other BSs select their own strategies $\alpha_j, j \neq i$. The probability distribution π is a joint point mass function of the different combinations of BSs strategies. As in [4], the inequality in correlated equilibrium definition means that when the recommendation to BS i is to choose action α_i^*, then choosing any other action instead of α_i^* cannot result in higher expected payoff for this BS.
Let us formulate the set of actions selected by all BSs as $\boldsymbol{\alpha} = \{\alpha_i \cup \boldsymbol{\alpha_{-i}}\}$, where $\boldsymbol{\alpha_{-i}}$ is the set of actions selected by all other BSs than i. We can introduce rate-dependent The Vickrey-Clarke-Groves (VCG) [5] auction mechanism design, where each of the BSs aims to maximize the utility U_i, $\forall i$, defined as:

$$U_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) \triangleq R_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) - \zeta_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right), \quad (11)$$

where ζ_i denotes the cost (rate loss) introduced by BS i to all other BSs, which is evaluated as follows:

$$\zeta_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) = \sum_{l\neq i} R_l\left(\boldsymbol{\alpha_{-i}}\right) - \sum_{l\neq i} R_l\left(\boldsymbol{\alpha}\right). \quad (12)$$

The use of VCG auction mechanism based on rate leads to maximization of the overall performance of the system by exploiting cooperation among nodes. However, in modern wireless systems the UEs are more interested in fulfilling their minimum QoS requirements rather than maximizing their rate or minimizing packet delay. Therefore, as an alternative one can consider a satisfaction-based VCG auction mechanism, with the satisfaction v_i defined as in (4) and (7), that can be formulated as:

$$U_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) \triangleq f_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) - \psi_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right), \quad (13)$$

where ψ_i denotes the satisfaction based cost evaluated as follows:

$$\psi_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) = \sum_{l\neq i} f_l\left(\boldsymbol{\alpha_{-i}}\right) - \sum_{l\neq i} f_l\left(\boldsymbol{\alpha}\right). \quad (14)$$

B. Regret-matching learning

To achieve the CE a centralized approach can be applied, which is, however, very complex [5]. Therefore, according to [12] the procedure of regret matching learning can be sued to iteratively achieve CE. In [4], [5], a modified regret-matching learning algorithm is proposed to learn in a distributive fashion how to achieve the correlated equilibrium set in solving the VCG auction, which aims at minimizing the regret of selecting certain action. The regret $REG^{(T)}$ of BS i at time T for playing action $\alpha^{(n)}$ instead of other actions is given as:

$$REG_i^{(T)}\left(\alpha_i^{(n)},\boldsymbol{\alpha_i^{(-n)}}\right)\triangleq\max\{D_i^{(T)}\left(\alpha_i^{(n)},\boldsymbol{\alpha_i^{(-n)}}\right),0\}, \tag{15}$$

where

$$\begin{aligned}D_i^{(T)}&\left(\alpha_i^{(n)},\boldsymbol{\alpha_i^{(-n)}}\right)=\\&=\max_{j\neq n}\frac{1}{T}\sum_{t\leq T}\left(U_i^t\left(\alpha_i^{(j)},\boldsymbol{\alpha_{-i}}\right)-U_i^t\left(\alpha_i^{(n)},\boldsymbol{\alpha_{-i}}\right)\right),\end{aligned} \tag{16}$$

where $U_i^t\left(\alpha_i^{(.)},\boldsymbol{\alpha_{-i}}\right)$ is the utility at time t. $D_i^T\left(\alpha_i^{(n)},\boldsymbol{\alpha_i^{(-n)}}\right)$ is the average payoff that BS i would have obtained if it had played other action than $\alpha_i^{(n)}$ every time in the past. Thus, positive value of $D_i^T\left(\alpha_i^{(n)},\boldsymbol{\alpha_i^{(-n)}}\right)$ means that BS i would have obtained higher average payoff when playing different action than n. Finally, given the regrets for all N actions, the probability of BS i selecting strategy n can be formulated as follows:

$$\pi_i^{(\alpha_i^{(n)})}(T)=1-\frac{1}{\mu^{(T-1)}}REG_i^{(T-1)}\left(\alpha_i^{(n)},\boldsymbol{\alpha_i^{(-n)}}\right), \tag{17}$$

where

$$\mu^{(T-1)}=\frac{\sum_n REG_i^{(T-1)}\left(\alpha_i^{(n)},\boldsymbol{\alpha_i^{(-n)}}\right)}{N-1}.$$

V. Simulations

A. General assumptions

To compare the performance of the system using interference mitigation based on CE and SE Monte-Carlo simulations of a Long Term Evolution - Advanced (LTE-A) like system with 3 macro BSs (forming 120°sectors) and 18 micro BSs (in 9 sites), as shown in Fig. 1, have been used. All BSs are assumed to transmit using the same frequency and bandwidth, equal to 2.6 GHz and 20 MHz, respectively. Full interference model has been implemented including links between each BS and each UE, with the channel propagation conditions determined based on models proposed in [13]. At each time instant there were 300 UEs uniformly distributed on the streets, hence only outdoor transmission has been considered. The maximum transmission power for macro BSs was set to 46 dBm, whereas for micro BSs it was 33 dBm. It was assumed that all BSs and UEs equip only single antenna and UEs are stationary.

B. Full-buffer analysis

1) Assumptions: In the first considered setup it was assumed that all BSs have infinite amount of best effort (BE) data to transmit, which is known as the full-buffer model. The following approaches to interference mitigation have been considered in the evaluation:

- a system with no inter-cell interference coordination (ICIC),

TABLE I: System spectral efficiency

Scheme	PF scheduler		PFMR scheduler	
	spectral efficiency [bps/Hz]	gain vs. no ICIC [%]	spectral efficiency [bps/Hz]	gain vs. no ICIC [%]
no ICIC	1.40	-	1.15	-
LTE-A ICIC	1.43	2.1	1.13	-1.7
CE-rate	1.50	7.1	1.24	7.8
CE-satisfaction	1.50	7.1	1.24	7.8
SE-SESA	1.52	8.6	1.28	11.3
SE-behavioural	1.56	11.4	1.28	11.3

- LTE-A interference coordination mechanism based on ABSF with 4 ABSF possible within a radio frame,
- interference mitigation scheme proposed in [3] based on CE and rate based payoff - denoted as CE-rate,
- the proposed ICIC mechanism based on CE with throughput satisfaction based payoff - denoted as CE-satisfaction,
- SE based solution using the SESA algorithm and throughput-based satisfaction [11] - denoted as SE-SESA,
- a solution based on SE using behavioural learning and throughput-based satisfaction- denoted as CE-behavioural,

For all learning algorithms it was assumed that actions are selected every 10 ms, which is the duration of the radio frame in considered system. Moreover, a minimum throughput constraint has been considered for all UEs equal to 500 kbps. The considered schemes have been compared in terms of achieved system spectral efficiency, which represents the total throughput per unit of bandwidth, and achieved UE throughput distribution, with detailed analysis of the 5th, 50th and 95th percentile.
To provide more detailed analysis of the performance achieved with different solutions two scheduling algorithms based on PF criterion have been considered for time-frequency resource allocation:

- opportunistic proportional fair (PFs) scheduler
- proportional fair minimum rate (PFMR) scheduler that provides minimum rate guarantees to all UEs [14]

2) Numerical results: In Table I the achieved system spectral efficiency is presented, with the indication of gain vs. system with no ICIC. One can notice that higher spectral efficiency is achieved when using the approaches based on equilibria, with the solutions using SE providing the highest efficiency. Gains even over 10% can be observed for both considered schedulers. High spectral efficiency achieved with the use of SE may indicate that in such dense system the required minimum data rate is too high to be achieved for all UEs. Therefore, in process of SE learning the highest possible utility is obtained, although SE may not be reached.
A very interesting and important parameter in terms of UE satisfaction analysis is the UE throughput distribution. Fig. 3 shows the UE throughput distribution achieved for different solutions with PFs, with the detailed crucial points - the 5th, 50th and 95th percentile - also shown in Table II.

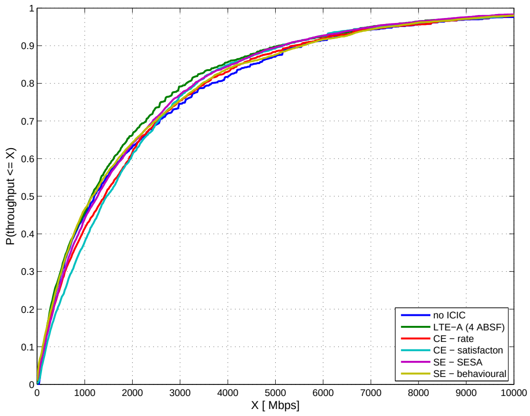

Fig. 3: UE throughput CDF with PFs scheduler

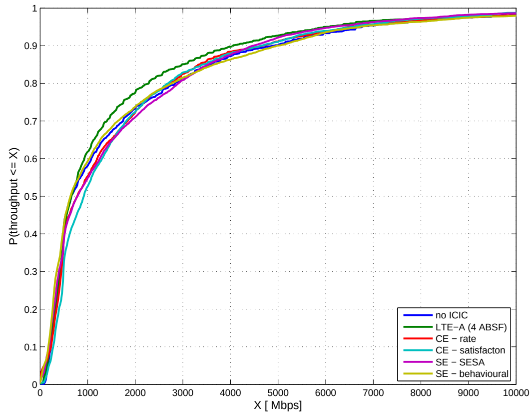

Fig. 4: UE throughput CDF with PFMR scheduler

TABLE II: UE throughput with PF scheduler

Scheme	5th percentile throughput [kbps]	50th percentile throughput [kbps]	95th percentile throughput [kbps]
no ICIC	99.55	1221.33	7474.53
LTE-A ICIC	95.07	1162.89	6996.60
CE-rate	96.32	1434.38	7348.0
CE-satisfaction	125.16	1480.71	7067.75
SE-SESA	84.59	1290.74	7039.50
SE-behavioural	54.42	1179.0	7457.0

TABLE III: UE throughput with PFMR scheduler

Scheme	5th percentile throughput [kbps]	50th percentile throughput [kbps]	95th percentile throughput [kbps]
no ICIC	181.86	675.94	6784.80
LTE-A ICIC	165.32	676.62	5978.23
CE-rate	104.53	800.75	6505.75
CE-satisfaction	205.80	938.25	6582.50
SE-SESA	129.68	803.10	6149.0
SE-behavioural	101.91	657.73	6745.14

Analysing the UE throughput CDF one can notice the large gain achieved with the CE-satisfaction approach, which combines the gains form cooperation and satisfaction-based optimization. This is the only one solution that provides higher 5th percentile throughput than tor the system with no ICIC, what indicates a significant improvement of rate of UEs experiencing the strongest interference. On the other hand, the solutions based on SE provide much lower 5th percentile throughput, what indicates that the SE cannot be achieved. Moreover, due to the clipping effect [7], where a BS keeps playing the same action when it is satisfied, without consideration of other BSs satisfaction, several BSs cannot achieve satisfaction because of the strong inter-cell interference. Thus, the cooperative approach based on CE and satisfaction-based utility provides much better results, as the clipping effect is avoided.

Similar conclusions can be drawn when analysing the CDF of UEs throughput obtained when using the PFMR scheduler, shown in Fig. 4 and Table III. However, due to the minimum rate constraints of the PFMR scheduler, the 5th percentile UE throughput is much higher than for the PFs scheduler. A value of over 200 kbps is achieved with the use of ICIC employing CE and satisfaction-based optimization. This value is much higher than the ones achieved with SE which also indicate the infeasibility of the satisfaction constraint for SE.

In case of CE-satisfaction approach even over 75% of UEs are satisfied - they achieve the throughput of at least 500 kbps. This indicates that the solution based on CE and satisfaction-based utility combined with proper scheduling strategy can significantly improve performance of the system in ultra dense deployments.

C. *Non-full-buffer analysis*

1) Assumptions: In the second setup it was assumed that each UE is using a single service from the following list:

- real-time voice service - based on voice over Internet protocol (VoIP) specification,
- data service - represented by file download using file transfer protocol (FTP).

To represent the services in a realistic manner the models proposed in [13] have been used.
The following approaches to interference mitigation have been considered in the evaluation:

- a system with no ICIC,
- LTE-A interference coordination mechanism based on ABSF with 4 ABSF possible within a radio frame,
- interference mitigation scheme proposed in [3] based on CE and rate based payoff - denoted as CE-rate,
- the proposed ICIC mechanism based on CE with delay-based satisfaction - denoted as CE-satisfaction,
- the proposed ICIC mechanism based on CE with hybrid throughput- and delay-based satisfaction - denoted as CE-hybrid,
- SE based solution using the SESA algorithm and delay-based satisfaction [11] - denoted as SE-SESA,
- a solution based on SE using behavioural learning and delay-based satisfaction- denoted as CE-behavioural,

For all learning algorithms it was assumed that actions are selected every 10 ms, which is the duration of the radio frame in considered system. Moreover, a maximum delay constraint for VoIP services equal to 150 ms and a minimum throughput constraint for FTP services equal to 500 kbps has been considered.
The investigated schemes have been compared in terms of real-time traffic packet outage probability, achieved system spectral efficiency, which represents the total throughput per unit of bandwidth and achieved UE throughput distribution, with detailed analysis of the 5th, 50th and 95th percentile.

2) Numerical results: The numerical results for non-full-buffer analysis will be provided in the camera-ready version of the paper.

VI. Conclusion

In this paper we address the problem of interference mitigation in ultra dense heterogeneous cellular network with the minimum throughput constraints from a game theoretic perspective. Several approaches based on CE and SE are compared. Moreover, a new solution based on cooperative approach using CE and satisfaction-based optimization is proposed. The proposed scheme outperforms other solutions in terms of satisfying the QoS constraints in form of minimum throughput or maximum packet delay requirement. Moreover, the difference in overall system performance in terms of spectral efficiency is minimal in favour of SE-based approaches. Thus, in systems that allow for BSs coordination, the approach using CE and satisfaction constrained utility seems to be the most promising one for long-term ICIC. On the other hand, when the coordination interface is the limiting factor, solutions based on SE may be preferred as they operate in fully distributed manner.

Acknowledgment

This work has been supported by the Polish Ministry of Science and Higher Education funds for the status activity project DS-MK (DS/81/147/MK).
The author would also like to thank his colleague at PUT - Adrian Kliks- for his continuous support and suggestions.

References

[1] 3GPP TS 36.300, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overal description", v. 11.3.0, September 2012

[2] LG Electronics, "3GPP TSG RAN WG1 Meeting #62 R1-104661 Comparison of Time-domain eICIC Solutions", August 2010

[3] P. Sroka, A. Kliks, "Distributed Interference Mitigation in Two-Tier Wireless Networks Using Correlated Equilibrium and Regret-Matching learning", *Proc. European Conference on Networks and Communicatons 2014*, Bologna, Italy, June 23-26, 2014

[4] M. Charafeddine, Z. Han, A. Paulraj, and J. Cioffi, "Crystallized Rate Regions of he Interference Channel via Correlated Equilibrium with Interference as Noise "in *Proc. IEEE ICC 2009*, Dresden, Germany, June 14-18, 2009

[5] A. Kliks, P. Sroka, M. Debbah, "Crystallized Rate Regions for MIMO Transmission ", *Eurasip Journal on Wireless Communications and Networking*, 2010

[6] H. Li,"Multi-agent Q-Learning of Channel Selection in Multi-user Cognitive Radio Systems: A Two by Two Case", in *Proc. IEEE Conference on SYStem, Man and Cybernetics (SMC)*, 2009

[7] S. M. Perlaza, H. Tembine, S. Lasaulce, M. Debbah,"Quality-of-Service Provisioning in Decentralized Networks: A Satisfaction Equilibrium Approach", in *IEEE Journal on Selected Topics in Signal Processing*, vol. 6, pp. 104-116, 2012

[8] J. F. Nash,"Equilibrium points in n-person games", in *Proc. of National academy of Sciences of the United States of America*, vol. 36, no. 4, pp. 46-56, 1950

[9] METIS consortium, "D1.1 - Scenarios, requirements and KPIs for 5G mobile and wireless system", *project report*, Oct. 2013

[10] R. J. Aumann, "Subjectivity and correlation in randomized strategy", *Journal of Mathematical Economics*, vol. 1, no. 1, pp. 67-96, 1974

[11] S. M. Perlaza, H. Tembine, S. Lasaulce, M. Debbah,"Satisfaction Equilibrium: A general Framework for QoS Provisioning in Self-Configuring Networks", in *Proc. IEEE Global Communications Conference (GLOBECOM 2010)*, 2010

[12] S. Hart, A. Mas-Collel, "A simple adaptive procedure leading to correlated equilibrium", *Econometrica*, vol. 68, no. 5, pp. 1127-1150, September 2000

[13] METIS consortium, "D6.1 - Simulation guidelines", *project report*, Oct. 2013

[14] L. Qian, N. Song, D.R.Vaman, X. Li, Z. Gajic, "Power control and proportional fair scheduling with minimum rate constraints in clustered multihop TD/CDMA wireless ad hoc networks,"in *Proc. IEEE Wireless Communications and Networking Conference 2006 (WCNC 2006)*, vol.2, pp. 763-769, 3-6 April 2006