

“© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

# A Fuzzy Q-Learning Approach for Enhanced Intercell Interference Coordination in LTE-Advanced Heterogeneous Networks

A. Daeinabi and K. Sandrasegaran

Centre for Real-time Information Networks, Faculty of Engineering and Information Technology,  
University of Technology, Sydney,  
Australia

**Abstract-** Since the transmission power of macro eNodeB (eNB) is higher than pico eNB in long term evolution -advanced heterogeneous network, the coverage area of picocell is small. In order to address the coverage problem, cell range expansion (CRE) technique has been recently proposed. However, CRE can lead to the downlink interference problem on both data and control channels when users are connected to pico eNB. In order to mitigate the downlink interference problem, a new dynamic almost blank subframe (ABS) scheme is proposed in this paper. In this scheme, a fuzzy q-learning approach is used to find the optimum ABS value. Simulation results show that the system performance can be improved through the proposed scheme.

**Keywords-** Heterogeneous networks, picocell, intercell interference, almost blank subframes (ABS), fuzzy q-learning.

## I. INTRODUCTION

Heterogeneous network (HetNet) consists of macrocells and low power nodes such as picocells to offload the traffic of macrocell and increase the spectral efficiency [1]. However, new challenges arise in network management when the picocell is deployed due to transmission power difference between macro eNodeB (eNB) and pico eNBs. In HetNet, the cell selection is performed using the cell range expansion (CRE) technique in which an offset value is added to downlink reference signal received power (RSRP) of the pico eNB. By this technique, the user equipment (UE) selects a pico eNB as the serving cell even if the pico eNB is not the strongest eNB as shown in Fig.1 (a) and (b). However, CRE leads to interference in downlink for the UEs located in the range expanded area (RE UE) due to decrease of their signal to interference plus noise ratio (SINR). As a result, RE UEs suffer from interference on both their data and control channels because they are not planned for too low SINR. Consequently, picocells may become underutilized due to severe interference. Therefore, one of the important aspects of HetNets is downlink intercell interference (ICI) mitigation.

In time domain, the interference problem can be mitigated using subframe utilization, called enhanced intercell interference coordination (eICIC), when macrocell and picocells share the bandwidth. This utilization is performed across different cells through almost blank subframes (ABSs) which are subframes without any activity or only transmitting the reference signals from macro eNB [2]. Based on the ABS configuration, RE UEs are scheduled within subframes that overlap with the ABSs of the macrocell. Moreover, pico eNB can also transmit during non-ABS periods. However, if the number of ABSs is not precisely selected, the throughputs of UEs connected to macro eNB and pico eNB are significantly

degraded because the number of ABS is not corresponding to number of RE UEs. Since ABSs value directly affects system performance then it is important which ABS value is suitable for a system such that the requirement of system can be satisfied and the system performance is improved.

In order to mitigate interference in time domain, several eICIC schemes have been proposed. Reference [3] has provided a good survey on ICIC schemes in long term evolution (LTE) and LTE-Advanced. Reference [4] proposed an algorithm in which the number of the ABS is defined as the number of the UEs connected to the picocells without CRE offset value, and the number of the UEs connected to the picocells with CRE offset value. In [5], the number and location of ABSs are determined using macro and pico loads and a minimizing approach. However, the system requirements such as the required data rate are not considered. Reference [6-7] proposed a method to find the optimal amount of ABS by a utility function. A cell selection technique has been presented in [8] to improve the system performance but the user association probability was not considered. Reference [9] introduced a decentralized interference management and cell selection using reinforcement learning approach. However, each macro and pico eNB store q-values in a table which is quite complex and impracticable for large state-action space and continuous state space. This paper proposes a dynamic and decentralized eICIC technique for downlink

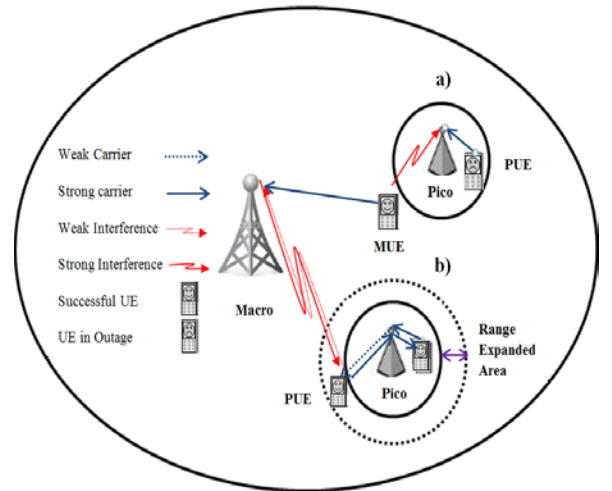


Fig.1.a) Macro UE (MUE) interferes on the uplink of a nearby picocell;  
b) Using cell range expansion to mitigate pico uplink interference present in Fig.1 (a).

interference mitigation to improve the system throughput. In this scheme, ABS value is obtained based on the historical system performance and using fuzzy q-learning approach. By this technique, each macrocell can change the ABS configuration to reach the expected performance.

The rest of the paper is organized as follows. Section II provides a description of the network model. Section III explains the fuzzy q-learning and the proposed dynamic ABS scheme. The simulation results are shown in Section IV. The conclusion is given in the final section

## II. NETWORK MODEL

### A. Model Description

The network topology is comprised of a set of macro eNBs and pico eNBs (represented by  $J$ ) as well as UEs which are distributed within cells. Picocells are used with lower transmission power than macrocells [10-11]. Moreover, macrocells and picocells work in open access mode where all UEs can access them. In this work, a well-defined area is used for simulation where the eNBs and UEs are located. All UE movements and transmissions are simulated in this area. Moreover, macro eNB and pico eNB communicate using the X2 interface [12]. Release 10 already supports the ABS information exchanging over X2 interface.

### B. Cell Selection

CRE is used as the cell selection technique in this paper. By this technique, UEs close to picocell will be connected to the pico eNB. Therefore, the traffic can be efficiently offloaded from macrocell to picocells.

$$\text{Serving cell} = \arg \max_{j \in J} (RSRP_j + \text{offset\_value}) \quad (1)$$

Different offset values in range (0 20] dB [13] can be added to RSRP received by UE from eNB  $j$ . Note that the offset value is added only to RSRP received from pico eNBs.

## III. THE PROPOSED eICIC SCHEME BASED ON FUZZY Q-LEARNING

Q-learning approach is one type of reinforcement learning techniques [13] in which an agent creates a table of q-values. By exploring the environment, the q-function estimates the discounted future costs when the agent applies an action in the current state. Then, the q-value is stored in the table for each state and each possible action. However, each agent has to store q-values in a table which is quite complex and impracticable for large state-action space and continuous state space. In order to overcome these problems, the fuzzy logic system (FLS) [14] can be added to q-learning approach, known as FQL [13], to discretise the continuous variables. A FLS is an expert system using "IF...THEN" rules. It could simultaneously work with numerical data and linguistic information using a mapping between input data and output data. Since FLS deploys linguistic terms, the previous information can be gathered easier using the experience of an operator. The main difference between FLS and conventional rule base controllers is that FLS can simultaneously trigger several rules which lead to a smoother control [14].

In FQL, continuous states and actions are transformed into a finite number of fuzzy variables using fuzzy membership functions. A membership function consists of a curve to define the mapping from a given point in the input or output space with a membership degree within the real interval [0 1]. By combining the FLS and Q-learning, the best sequence for each rule can be optimally selected. Consequently, FQL not only solves the problem of continuous state space but also the human expertise can be used to determine the fuzzy rules. In the proposed scheme, the network is modelled as a multi-agent system where each macro eNB acts as an agent to find its optimum ABS value as shown in Fig.3. The required components of the proposed scheme and the FQLC algorithm used to find the optimum ABS value are explained as follows.

### A. Fuzzy Q-Learning Controller (FQLC) Components

During each iteration, each agent observes the complete state of the system and then applies an action which causes the system to move into a new state. After that, the action receives an instantaneous reward which indicates the quality of this transition. The FQLC's state input, action and rewards used in the proposed dynamic ABS scheme are described below:

#### 1. State Fuzzy Variables

If the number of ABS and non-ABS are not corresponding to number of macro and pico UEs, the throughputs of macro UEs and SINR of RE UEs will significantly decrease. Therefore, in order to indicate how the system is performing, important downlink metrics can be measured and used as an input state vector represented by  $S$ :

$$S = [N_{UE} \quad T_M \quad SINR_{RE} \quad V_{ABS}] \quad (2)$$

where  $N_{UE}$  indicates the user ratio which equals the numbers of RE UEs to macro UEs.  $T_M$  represents the average throughput of macro UEs. Moreover,  $SINR_{RE}$  is average of 5 % of cumulative distribution function (CDF) of SINR of pico UEs which indicates the minimum SINR achieved by 95% of pico UEs.  $T_M$  and  $SINR_{RE}$  ensure that not only the required throughput of macro UEs is satisfied but also the throughput of UEs with low SINR is improved.  $V_{ABS}$  is the current ABS value used by macrocell and its own picocells. In order to fuzzify the variables, finite numbers of fuzzy labels are defined for each input variable and each label is assigned to a membership function. In the proposed scheme, state variables are labelled using  $T_{s1}$ ,  $T_{s2}$ ,  $T_{s3}$ , and  $T_{s4}$  sets, respectively, as shown in Fig.3.

$$T_{s1} = \{Small, Medium, Large\}$$

$$T_{s2} = \{Low, Medium, High\}$$

$$T_{s3} = \{Low, Medium, High\}$$

$$T_{s4} = \{Low, Medium, High\}$$

Note that the dimension of the input state vector and the number of its corresponding fuzzy labels establishes a trade-off between the network accuracy and the convergence speed of the learning procedure. In order to mathematically explain of these linguistic components, the triangular, s-shape and z-shape functions are given by (3), (4) and (5), respectively. In

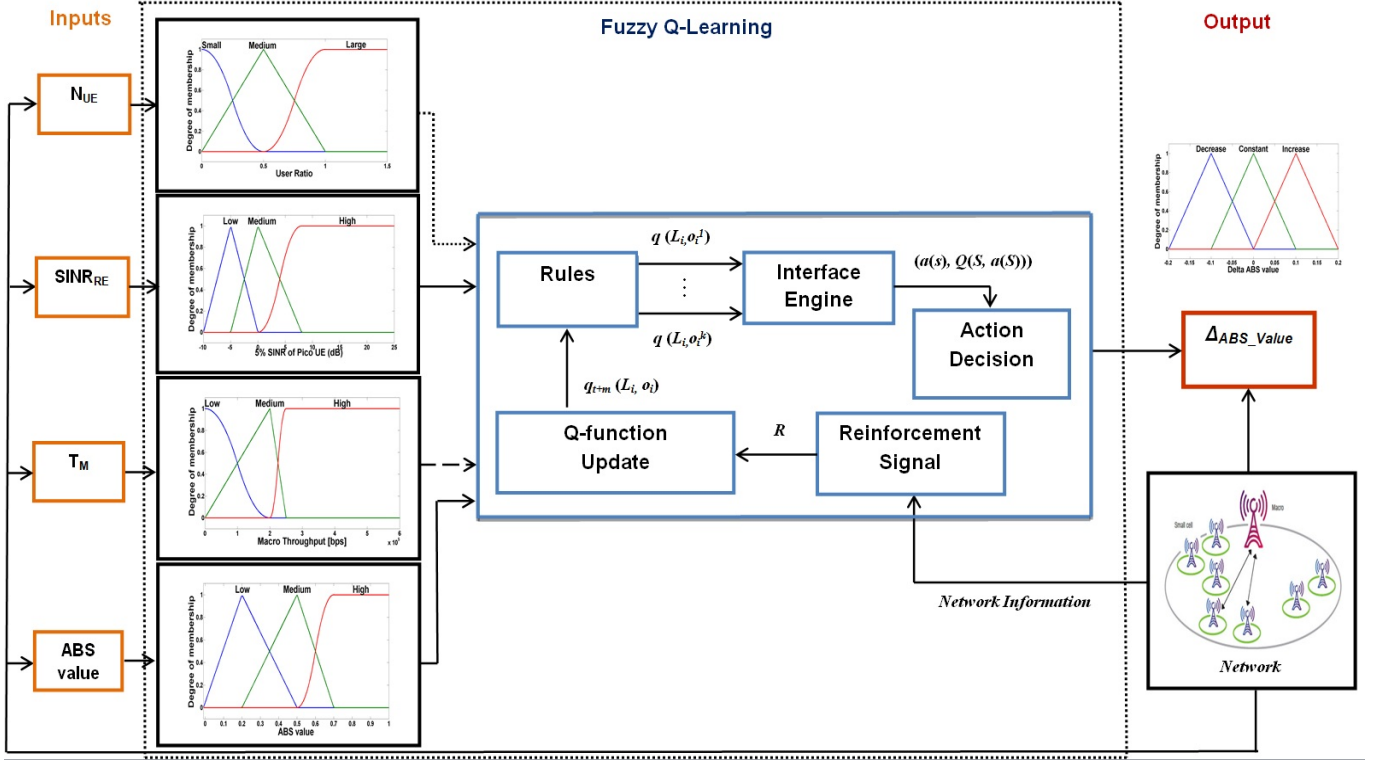


Fig.3. Block diagram of the proposed scheme

general, they are functions of a vector,  $x$ , and depend on scalar parameters  $a$ ,  $b$ , and  $c$  which are selected according to system requirements.

$$f(x; a, b, c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{b-x}{b-a}, & b \leq x \leq c \\ 0, & x \geq c \end{cases} \quad (3)$$

$$f(x; a, b) = \begin{cases} 0, & x \leq a \\ 2 \left( \frac{x-a}{b-a} \right)^2, & a \leq x \leq \frac{a+b}{2} \\ 1 - 2 \left( \frac{x-b}{b-a} \right)^2, & \frac{a+b}{2} \leq x \leq b \\ 1, & x \geq b \end{cases} \quad (4)$$

$$f(x; a, b) = \begin{cases} 1, & x \leq a \\ 1 - 2 \left( \frac{x-a}{b-a} \right)^2, & a \leq x \leq \frac{a+b}{2} \\ 2 \left( \frac{x-b}{b-a} \right)^2, & \frac{a+b}{2} \leq x \leq b \\ 0, & x \geq b \end{cases} \quad (5)$$

## 2. Action

The action of FQL controller is to find the  $\Delta_{ABS\_value}$  which is added to current ABS value for each macrocell. Each macro eNB performs local actions to find this value and then uses the optimum ABS value for itself and its own picocells.  $\Delta_{ABS\_value}$  is fuzzified with three membership functions as shown in Fig.3. The selected range ensures that the changes are not sharply performed.

## 3. Interface Stage

In the interface stage, mapping of states to actions are defined by a set of "IF-Then" rules to control the output actions. The fuzzy rules of FQL are represented by:

$$\begin{aligned} &\text{IF } (S^1 \text{ is } L_i^1) \text{ AND } (S^2 \text{ is } L_i^2) \dots \text{ AND } (S^N \text{ is } L_i^N) \\ &\text{THEN } a = o_i^1 \text{ with } q(L_i, o_i^1) \end{aligned}$$

$$\text{or } a = o_i^k \text{ with } q(L_i, o_i^k)$$

where  $L_i^n$  is a fuzzy label corresponding to a distinct fuzzy set defined in the domain of the  $n^{\text{th}}$  component  $S^n$  of the state vector  $S = [S^1, S^2, \dots, S^N]$  and  $o_i^k$  is the  $k^{\text{th}}$  output action for rule  $i$  which also is called the consequent part of rule  $i$ . The vector  $L_i = [L_i^1, L_i^2, \dots, L_i^N]$  is named the modal vector of rule  $i$  which specifies one state of controller.  $q(L_i, o_i^k)$  is the q-value function of the state  $L_i$  and action  $o_i^k$  of rule  $i$ . The initial values of all  $q$  are set to zero. Moreover, it is assumed that the set of actions is same for all eNBs. The number of actions is corresponding to the prior knowledge about the consequent part of the rules. If the rule is precisely acknowledged then one

action is enough to put in consequent part of that rule. When the knowledge about a rule is imprecise, a subset of possible actions is considered for that rule. For the system without any prior knowledge, all the possible actions are taken into account to define the rule with equally probability in all states. In this paper, it is assumed that there is no prior knowledge.

#### 4. Instantaneous Rewards

In order to inform the controller about the result of its previous actions, a reinforcement learning reward  $R$  is defined. If  $R$  is correctly defined, the best consequent can be selected for each rule.  $R$  is calculated when the controller transits from state  $S_t$  to state  $S_{t+m}$  to characterize the results of actions. If the selected action leads to a good result, a positive reward is considered while the negative reward is assigned to the poor result. In this paper,  $R$  is defined as follows:

$$\begin{cases} R = 1, & \text{If } M_{thr} \geq T_1 \text{ and } M_{SINR} \geq T_2 \\ R = -1, & \text{Otherwise} \end{cases} \quad (6)$$

where  $M_{thr}$  and  $M_{SINR}$  are the required macro UE throughput and 5% of CDF of pico UEs' SINR, respectively and  $T_1$  and  $T_2$  are their corresponding thresholds, respectively. Moreover,  $m$  is set to 10 to investigate the average effect of ABS value during each frame.

#### B. FQLC Algorithm

At the first step, the current state of the agent is specified using the truth degree of each rule. The truth degree of rule  $i$  equals the product of membership degrees of each input state for the rule  $i$  as follows:

$$\alpha_i(S) = \prod_{n=1}^N \mu_{L_i^n}^n(S^n) \quad (7)$$

where  $N$  represents the number of FLC inputs and  $\mu_{L_i^n}$  is the membership function for the  $n^{\text{th}}$  FLC input and the rule  $i$ . At the next step an action is selected for each activated rule based on the greedy exploration/exploitation policy [13]:

$$\begin{aligned} a_i &= \arg \max_{k \in K} q(L_i, o_i^k) \quad \text{with probability } \varepsilon \\ a_i &= \text{random}\{o_i^k\}_{k \in K} \quad \text{with probability } 1-\varepsilon \end{aligned} \quad (8)$$

where  $\varepsilon$  determines a trade-off between exploration and exploitation in the algorithm. When  $\varepsilon$  increases over the time, a slow transition occurs from exploration to exploitation. After determining the action of each activated rule, the inferred action of FQLC is calculated by the following equation:

$$a(S) = \sum_{i \in P} \alpha_i(S) \cdot o_i \quad (9)$$

where  $P$  is the set of rules. The  $q$ -value for the input state vector is computed using an interpolation among the current  $q$ -values at the modals point of the activated rules:

$$Q(S, a(S)) = \sum_{i \in P} \alpha_i(S) \cdot q(L_i, o_i) \quad (10)$$

---

#### Algorithm 1. Fuzzy Q-learning

---

1. Initialize the  $q(L_i, o_i^k)$  for all  $i \in P$  and  $k \in K$
  2. Select an action for each activated rule based on  $\varepsilon$ -greedy policy via (8).
  3. Calculate the inferred action using (9).
  4. Estimate the  $Q$ -function from the current  $q$ -values and the degree of truth of the rules using (10).
  5. Apply the inferred action and observe new state  $S_{t+m}$ .
  6. Extract the reinforcement learning reward based on performance results.
  7. Compute the value of the new state represented by  $V(S_{t+m})$  in (11).
  8. Calculate the difference between the old and new values of  $Q(S, a(S))$  using (12).
  9. Update the  $q$ -values for each activated rule based on (13).
  10. Repeat the process starting from step 2 until the convergence is attained.
- 

When the inferred action is applied by FQLC the agent transits from the current state  $S_t$  to the new state  $S_{t+m}$ . The value of the new state is calculated as follows:

$$V(S_{t+m}) = \sum_{i \in P_{S_{t+m}}} \alpha_i(S_{t+m}) \cdot \max_{k \in K} q(L_i, o_i^k) \quad (11)$$

Then, the reward  $R$  is given to the agent because of this transition and the  $q$ -values are updated using a quantity  $\Delta Q$  defined as the difference between the old and new values of  $Q(S, a(S))$ .

$$\Delta Q = R_{t+m} + \gamma V(S_{t+m}) - Q(S_t, a(S_t)) \quad (12)$$

$$q_{t+m}(L_i, o_i) = q_t(L_i, o_i) + \eta \alpha_i(S_t) \Delta Q \quad (13)$$

The parameters  $\eta$  and  $\gamma$  are the learning rate and discount factor in range  $[0, 1]$ , respectively. By updating the  $q$ -values in each state, the best action is obtained for each rule and then the optimum ABS value can be resulted.

## IV. SIMULATION AND DISCUSSION

This section evaluates the proposed scheme using a system level simulation for 2GHz carrier frequency and 5 MHz bandwidth. The network consists of 7 macro eNBs deployed with a hexagonal layout and an inter-site distance of 400m. The transmission powers of macro eNBs are set to 43dBm. Moreover, There are 4 pico eNBs in each macrocell with transmission powers of 30dBm [10].  $T_1$  and  $T_2$  are set to 200 kbps and -6.5 dB [15-16]. Same as other literature work, the proposed scheme is benchmarked with respect to static ABS schemes. It is considered as a sanity check of results obtained by the proposed scheme from prior known results. The metrics used for evaluation are explained as follows:

- **ABS Value:** it shows the percentage of ABSs in each frame.
- **Percentage of RE UEs:** it indicates the percentage of RE UEs to total number of UEs.

Table I: Average Percentage of the Muted UEs (%)

Offset Value	Percentage of RE UEs	Average Percentage of the Muted UEs (%)								
		ABS value= FQL	ABS value= 10%	ABS value= 20%	ABS value= 30%	ABS value= 40%	ABS value= 50%	ABS value= 60%	ABS value= 70%	ABS value= 80%
9 dB	14%	8.97	7.91	10.7	12.3	14.4	16.5	18.7	20.8	23
12 dB	41%	7.14	2.05	4.09	6.08	8.01	9.92	11.83	13.76	15.77
15 dB	58%	5.38	1.02	2.03	3.03	4.02	5.01	6	6.98	7.97

- *Average percentage of the muted UEs*: number of macro UEs which physical resource blocks (PRBs) are not allocated to them to total number of UEs.
- *5% pico UE throughput*: it equals the 5% of CDF of the throughput for pico UEs.
- *Average macro UE throughput*: it equals the average throughput of UEs connected to macro eNB.
- *Average macrocell throughput*: it is defined as the throughput of macrocell without considering picocells.
- *Average picocell throughput*: it equals the sum of throughputs of picocells distributed in each macrocell.
- *Outage Probability*: it is the ratio of UEs which their average SINR is smaller than the SINR threshold [15-16].

#### A. Percentage of the Muted UEs

The ABS value resulted by the proposed dynamic scheme is depicted in Fig.8. Because of space limitation, only the ABS value for centre cell is shown. Table I shows when the offset value is low, a few numbers of UEs are located in range expanded area, while more UEs are offloaded to this area for larger offset values. Therefore, the proposed scheme selected higher ABS value for offset value of 15 dB than offset values of 12 dB and 9 dB. Moreover, the percentage of the muted UEs is presented for different offset values and ABS values in this table. When the offset value is small (e.g., 9 dB), more UEs are connected to macro eNB. In this case, number of PRBs is not enough for macro UEs' buffered data and hence some macro UEs are not allocated PRB. It will be worse when the ABS value increases because the number of subframes that macro UEs have to be muted increases. However, when the offset value increases (e.g., 15 dB), number of UEs offloaded to picocell increases which yields to decrease of number of the muted UEs. Since the proposed dynamic ABS scheme determines ABS value based on the defined inputs, the percentages of the muted UEs are lower than or close to the corresponding static ABS values.

#### B. Outage Probability

The outage probability of the proposed scheme is compared for different CRE offset values with different ABS

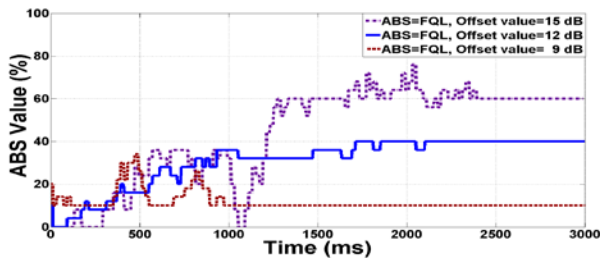


Fig.8. The ABS value obtained by FQLC for the center cell

values in Fig.9. It is obvious that the outage probability is an increasing function of offset value. The reason is when the offset value increases UEs far away from pico eNBs are offloaded to picocell. On the other hand, using ABS value leads to decrease of outage probability because UEs with high received interference can be scheduled on ABSs. Since the proposed scheme considers the SINR value, it can keep the outage probability in the range of [0.22 0.25] as shown by the red ellipse in Fig.9. By changing the SINRs' membership functions, the outage probability can be changed with respect to the system requirements.

#### C. Throughput Performance

When a small ABS value is used, more non-ABS are assigned to macro UEs while the number of ABSs that can be allocated to RE UEs is low. Consequently, as shown in Fig.10, the throughput of macro UEs increases while the 5% throughput of pico UEs (see Fig.11) is degraded. In contrast, if the ABS value increases, the subframes that can be allocated to macro UE will decrease and the number of ABS assigned to RE UEs will increase. Therefore, the increase of ABS value can lead to increase of the 5% pico UE throughput at the expense of reduction of macro UE throughput. It can be concluded that there is a trade-off between macro UE and pico UE throughputs when ABS scheme is applied. On the other hand, the increase of pico UE throughput can increase the picocell throughput depicted in Fig.13. However, for larger ABS values, the throughput of macro UEs reduces significantly which lead to decrease of macrocell throughput as shown in Fig.12. In order to keep the trade-off between throughput of macro UEs and pico UEs as well as satisfying the required throughput for macro UEs, the proposed dynamic scheme set ABS value on 10% for offset value equals 9 dB (see Fig.8). This is because few numbers of UEs are located in range expanded area, and hence few numbers of ABSs is needed to support them. As the result, more non-ABSs can be allocated to macro UEs to satisfy their required throughputs.

By increasing the offset value to 15 dB, more UEs are located in range expanded area while they are far away from pico eNB. They suffer high interference from macro eNB and then the SINR of RE UE reduces. In this case, if the ABS value is set to small value, the number of ABS cannot support all of RE UE and then pico eNB has to schedule them on non-

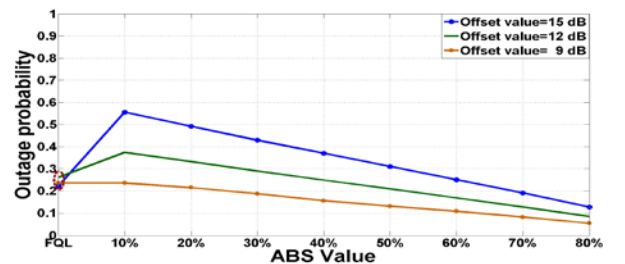


Fig.9. The outage probability for different CRE offset values with different ABS values



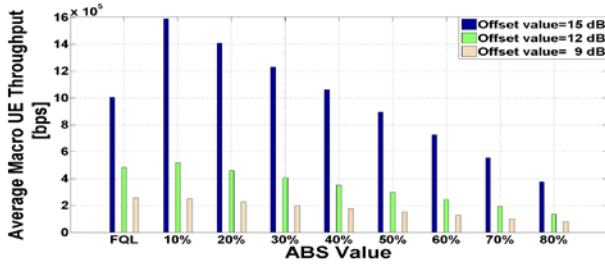


Fig.10. Average macro UE throughput

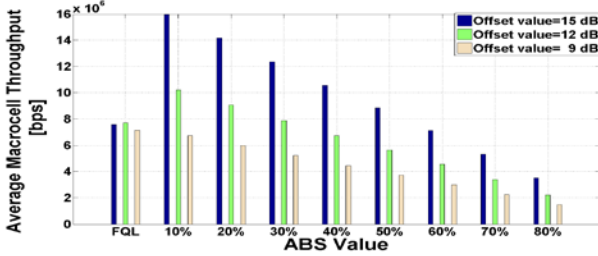


Fig.12. Average macrocell throughput

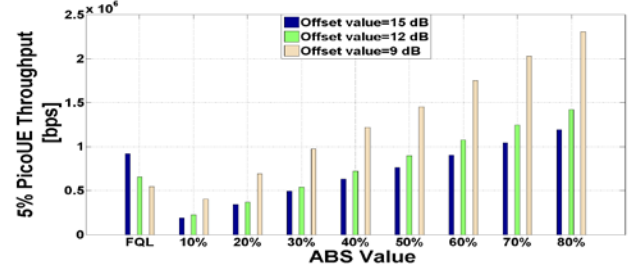


Fig.11. 5% pico UE throughput

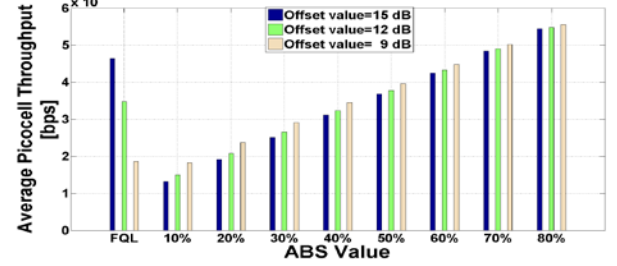


Fig.13. Average picocell throughput

ABS. Since the interference on non-ABS is high for RE UEs particularly for RE UEs far away from pico eNB, the 5% pico UE throughput and picocell throughput decrease as shown in Figs.11 and 13. To overcome this problem, the proposed dynamic ABS scheme dynamically converged to the ABS value equals 60% as shown in Fig.8. As shown in Figs 10 and 11, the average throughput of macro UEs is satisfied while the 5% throughput of pico UEs is still at the good level.

## V. CONCLUSION

In this paper, a dynamic ABS scheme has been proposed to overcome the downlink interference problem when macrocell and picocell share the bandwidth. In the proposed scheme, the ABS value is selected based on the system performance and using a FQL. The simulation results show that the proposed scheme can keep the trade-off between RE UE throughput and the required throughput of macro UEs compared to static ABS value schemes while the outage probability is kept around a constant value. As a future work, we will work on other types of rewards to obtain more improvements in terms of system throughputs.

## REFERENCES

- [1] A. Damnjanovic, J. Montojo, W. Yongbin, J. Tingfang, L. Tao, M. Vajapeyam and D. Malladi, "A Survey on 3GPP Heterogeneous Networks," IEEE Wireless Communications, vol.18, no.3, pp.10-21, February 2011,
- [2] 3GPP, TS 136.300, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 10)," 2011.
- [3] Ch. Kosta, B. Hunt, A. U. Qudus and R. Tafazolli, "On Interference Avoidance Through Inter-Cell Interference Coordination (ICIC) Based on OFDMA Mobile Systems," IEEE Communications Surveys & Tutorials, vol. 15, no. 3, Third Quarter, pp. 973-995, 2013.
- [4] M. Shirakabe, A. Morimoto, and N. Miki, "Performance Evaluation of Inter-Cell Interference Coordination and Cell Range Expansion in Heterogeneous Networks for LTE-Advanced Downlink," 8<sup>th</sup> International Symposium on Wireless Communication Systems (ISWCS), pp. 844-848, November 2011.
- [5] S.N.S. Kshatriya, S. Kaimalettu, S.R. Yerrapareddy, K. Milleth, and N. Akhtar, "On Interference Management based on Subframe Blanking in Heterogeneous LTE Networks," 5<sup>th</sup> International Conference on Communication Systems and Networks (COMSNETS), pp.1-7, 2013.
- [6] J. Pang, J. Wang, G. Shen, Q. jiang and J. Liu, "Optimized Time-Domain Resource Partitioning for Enhanced Inter-cell Interference Coordination in Heterogeneous Networks," IEEE Wireless Communication and Networking Conference: MAC and Cross-Layer Design, pp.1613-1617, 2012.
- [7] L. Jiang and M. Lei, "Resource Allocation for eICIC Scheme in Heterogeneous Networks," IEEE 23<sup>rd</sup> International Symposium on personal, Indoor and Mobile Radio Communications (PIMRC), pp.464-469, September 2012.
- [8] J. Oh, and Y. Han, "Cell Selection for range expansion with Almost Blank Subframes in Heterogeneous Networks," 23<sup>rd</sup> IEEE Personal Indoor and Mobile Radio Communications (PIMRC), pp.653-657, September 2012.
- [9] M. Simsek, M. Bennis and A. Czyliwki, "Dynamic Inter-Cell Interference Coordination in HetNets: A Reinforcement Learning Approach", IEEE Global Communications Conference (GLOBECOM), pp.5446 – 5450, December 2012 .
- [10] 3GPP, TR 36.931, "LTE; Evolved Universal terrestrial Radio Access (E-UTRA); Radio Frequency Requirements for LTE Pico Node B", May 2011.
- [11] 3GPP, TS 36.814, "Evolved Universal Terrestrial Radio Access (EUTRA); Further Advancements for E-UTRA Physical Layer Aspects", March 2010.
- [12] 3GPP, TS 36.423, "X2 application protocol (X2AP)," June 2011.
- [13] P.Y. Glorennec, "Reinforcement Learning: an overview" European Sym.
- [14] L. Zadeh, "Fuzzy sets", Information Control, vol. 8, no.3, pp.338–353, 1965.
- [15] Ch.Sh. Chiu, and Ch.Ch. Huang, "An Interference Coordination Scheme for Picocell Range Expansion in Heterogeneous Networks," 75<sup>th</sup> IEEE Vehicular Technology Conference (VTC Spring), pp.1-6, 2012.
- [16] 3GPP, R1-100350, "Downlink CCH performance aspects for co-channel deployed macro and HeNBs," 3GPP TSG RAN WG1 Meeting #59bis, Nokia Siemens Networks, Nokia, January 2010.