Load Balancing in Heterogeneous Networks Based on Distributed Learning in Potential Games

Mohd. Shabbir Ali, Pierre Coucheney, and Marceau Coupechoux

Abstract—We present a novel approach for distributive load balancing in heterogeneous networks that use cell range expansion (CRE) for user association. First, we formulate the problem as a minimisation of an $\alpha-$fairness objective function. Depending on α, different objectives in terms of network performance or fairness can be achieved. Next, we model the interactions among the base stations for load balancing as a potential game, in which the potential function is the $\alpha-$fairness function. The optimal Nash equilibrium of the game is found by using distributed learning algorithms. We use log-linear and binary log-linear learning algorithms for complete and partial information settings, respectively. By running extensive simulations, we show that the proposed algorithms converge within a few tens of iterations. The convergence speed in the case of partial information setting is comparable to that of the complete information setting. We also show that the best response algorithm does not necessarily converge to the optimal Nash equilibrium.

I. INTRODUCTION

Due to the ever increasing demand for improved quality of service in terms of higher data rates and improved coverage, the conventional cellular networks are becoming heterogeneous [1]–[25]. Heterogeneous networks consist of macro base stations (BSs) and small BSs that transmit with high and low power, respectively. Conventional user association rule is such that the users select a BS that provides the highest received power. This may however result in an imbalance between BSs loads because the macro BS transmits at higher power and thus associates with more users. This creates overload situation at the macro BSs and at the same time under-utilised resources at the small BSs. Therefore, a natural problem that arises is how to associate users to BSs such that the network resources are utilised efficiently and the load is shared among the BSs.

Load balancing has been extensively studied in the literature using various approaches. An overview can be found in [4], [26]. These can be broadly classified as centralised [15], [16], [27], [28] and decentralised optimisation approaches [14], [17], [18], [29], [30]. Centralised user association rules along with inter-cell interference avoidance for load balancing are proposed in [27], [28]. Basic idea is to schedule users across the BSs in the network so that the users do not severely interfere with each other. Similar approaches where the user association decisions are modelled as Markov decision process (MDP) are presented in [15], [16]. However, centralised

Mohd. Shabbir Ali and M. Coupechoux are with Telecom ParisTech and CNRS LTCI, 46 rue Barrault, 75013, Paris, France. Emails: mdshabbirali88@gmail.com, marceau.coupechoux@telecom-paristech.fr.

P. Coucheney is with PRiSM, UVSQ / CNRS, 45 avenue des Etats-Unis, 78035 Versailles, France. Email: pierre.coucheney@uvsq.fr,

solutions are computationally intensive, require huge information exchange overhead, and are not scalable. To overcome these limitations, decentralised approaches are followed. A distributed algorithm is proposed for the load balancing, which is formulated as a convex optimisation problem in [29]. A decentralised solution for convex optimisation approach for joint cell association and resource allocation problem is undertaken for the load balancing in [14]. In [30], an online algorithm for access points association in wireless local area network (WLAN) based on the $\mathbb{L}_p$ norm of the loads on access points in proximity is proposed. Game theoretical approaches are also proposed for decentralised solutions [17], [30]. Also heuristic approaches are studied in the literature. For example, in [31], an algorithm is proposed to find the optimal beacon power that minimises the load of the most congested access point (AP) in WLAN.

Another important approach for load balancing that has attracted lot of interest in the literature is user association using cell range expansion (CRE) [3]–[13]. According to CRE technique the users associate with a BS that provides the highest biased received power. Bias value for small BSs is greater than or equal to one and for macro BSs it is equal to one. This results into increase in the coverage regions of the small BSs and thereby associating more users to them.

However, there are several challenges for using CRE technique. One of the challenges is to avoid high co-channel inter-cell interference in downlink from the macro BSs experienced by the small BSs' cell edge users. To address this problem advanced interference management techniques have been studied in the literature [3], [8], [11]. Resource partitioning is one of the techniques, in which the macro and the small BSs transmit in orthogonal time/frequency slots. Thereby, reducing the co-channel interference. Particularly, small BSs transmissions are scheduled in the almost blank subframe (ABS), in which macro BS transmit with lower power or do not transmit [11]. The advantage of using advanced interference cancellation receiver in conjunction with CRE technique is shown in [8].

Another challenge, which is the focus of our work, is to determine the optimal CRE bias values for a desired optimal performance of the network. For example, the same bias values are not optimal for both rate-optimal and delay-optimal performance of the network. A set of papers determines the optimal range of bias values by simulations and experimentations for a specific performance requirement of the network [4], [12], [19]. In [4], the authors show that bias range of $5 - 10$ dB is rate-optimal when both small and macro BSs use the same frequency band. When ABS is used in conjunction with CRE

the optimal bias range increases to $15-20$ dB. In [12], authors show that CRE bias values above 6 dB have a negative impact on the handover failure rate. In [19], authors implemented a testbed to show that CRE increases uplink bit rates at the price of little reduction of the throughput on the downlink.

Another set of papers determines optimal bias by applying heuristic, learning algorithms, and optimisation [20]–[24]. In [20], the bias is set according to the feedback from the network performance, which is maximisation of resource utilisation and improvement in quality of service for cell edge users. In [21], the authors propose an adaptive CRE scheme in which users choose among two biases depending on their signal-to-interference-plus-noise ratio (SINR). They showed through simulations that cell edge user throughput is improved while maintaining the average user throughput performance. In [25], a Q-learning based algorithm is used by all users to determine the optimal bias value by minimising the cost, which is the number of users in outage, broadcasted by the small BSs. In [22], a dynamic programming and a greedy approach are proposed to associate users to small and macro BSs with the goal to achieve a global proportional fairness. In [24], the load balancing problem is formulated as an integer optimisation problem aiming at maximising the Jain's fairness index. In [23], a simple heuristic method is proposed to adapt the bias values. According to the relative utilities of the macro and the pico-cell the bias value is stepwise increased or decreased.

In all the above papers the optimal bias values are determined for a specific optimal performance requirement of the network. However, there is no general framework for determining the optimal biases for different optimal performance requirements of the network. This is the focus of our work. We address this problem by considering an $\alpha-$fairness objective function that captures various aspects of the network performance and fairness for different α values. For $\alpha = 0$ it gives the rate-optimal policy, for $\alpha = 1$ gives the proportional fair policy, for $\alpha = 2$ gives the delay-optimal policy, and as $\alpha \to \infty$ it gives the min-max load policy. We present a novel approach, according to which the BSs learn their optimal bias values given α. In contrast to the related load balancing literature our approach is unique. Although an $\alpha-$fairness function is also considered for load balancing in [29], our approach is entirely different in terms of the user association rule, game framework, and distributed learning algorithms. We solve an optimisation problem considering a generalised objective function that is not convex unlike in [14], [29]. It is network centric, which gives full control to the network operator, as opposed to the user centric and hybrid approaches using MDP [15], [16]. It is guaranteed to converge to the optimal Nash equilibrium (NE) rather than a sub-optimal NE when compared to [17], [18]. Compared to simulations/testbeds our approach is generic, i.e. not specific to a single scenario. Its distributive nature favours ease of implementation, scalability of network, and robustness to node failures.

Another distinctive feature of our approach is the use of distributed learning algorithms for solving a non-convex opti-misation problem of load balancing. Our approach is general and can be used to distributively solve other optimisation problems. The idea is to enforce the potential game structure so that the objective function is exactly the potential function.

A. Contributions

We summarise our main contributions below.

- We present a novel approach for load balancing in heterogeneous networks that uses CRE for user association. Our approach is to distributively minimise an $\alpha-$fairness objective function that captures various performance and fairness criteria for different α.
- We prove that for $\alpha = 0$ the objective function captures rate-optimal policy of the network. For $\alpha \to \infty$, we prove that it results in the min-max load policy. We extend the classical result derived in [32] by considering a non-convex $\alpha-$fairness function.
- We solve the load balancing problem using distributed learning algorithms, where BSs learn the optimal CRE bias values. First, interactions between BSs are modelled as a potential game using wonderful life utility (WLU) structure. However, the WLU structures works only when the neighbourhood of the BSs is static. We propose a technique that allows to use the WLU structure with a time-varying neighbourhood. Next, the optimal NE of the game is determined by using log-linear learning algorithms. We consider two different settings: complete and partial information. In the former setting, we use classical log-linear learning algorithm (LLLA), whereas in the latter setting, we use binary log-linear learning algorithm (BLLLA). To the best of our knowledge, this is a novel approach in the load balancing literature.
- By running extensive simulations, we show that the proposed algorithms converge within a few tens of iterations to the optimal NE. The convergence speed of the BLLLA is comparable to that of the LLLA, meaning that partial information is sufficient in practical implementations. We also show that the classical best response (BR) algorithm does not necessarily converge to the optimal NE, although it necessitates complete information.

This paper is organised as follows. In Section II, the description of system model and problem formation is given. In Section III, a potential game framework solution is presented. The various distributed algorithms that are considered in this paper are described in Section IV. In Section V, our approach is validated using extensive simulations. Finally, the conclusions and future work are given in Section VI.

II. SYSTEM MODEL

A. Network Model

We consider a cellular network (typically a LTE-Advanced network) consisting of $\mathcal{B}_e$ eNodes-B (eNB) or BSs and $\mathcal{B}_s$ small BSs in a two dimensional region $\mathcal{L}$. The set of all stations is denoted $\mathcal{S} \triangleq \mathcal{B}_e \cup \mathcal{B}_s$. Every small BS maintains and can vary a parameter called CRE bias, denoted as c_j for $j \in \mathcal{B}_s$. The CRE vector is $\bar{c} = \begin{bmatrix} c_1, & c_2, & \ldots, & c_{|\mathcal{S}|} \end{bmatrix}$, where c_i is BS i's

CRE bias, which for practical purposes takes discrete values from 1 to $c_{\max}$. The CRE biases for macro eNBs are fixed to unity, i.e., $c_k = 1, \forall k \in \mathcal{B}_e$. This leads to no bias in the received power from a macro eNB.

1) Channel Model: The received power at location x from BS i is $P_i g_i(x)$, where P_i is the transmit power and $g_i(x)$ is the channel gain, which captures the effect of path-loss. The effect of small-scale fading is not considered because the time for user association procedure is assumed to be much larger than the channel coherence time [29]. Inclusion of shadow fading increases the complexity of model and is thus left to future work. Formally, these assumptions related to channel model are summarised below.

Assumption 1: [Deterministic propagation model] The channel gain $g_i : \mathcal{R}^2 \to \mathcal{R}$ is a deterministic function of the distance between BS i and a user.

The SINR $\gamma_i(x)$ at location x provided by the BS i is a random variable defined as:

$$\gamma_i(x) = \frac{P_i g_i(x)}{\sum_{j \in \mathcal{S}} \delta_j P_j g_j(x) + N_0}, \quad (1)$$

where N_0 is thermal noise power and δ_j is a random variable which is one when station j is active (i.e. transmitting) and 0 otherwise. To get more insights without much complexity due to random variable δ, we make use of the following assumption.

Assumption 2: [Worst case interference] The variables δ_j are deterministic and equal to 1 for all j. This provides an upper bound on the interference, which is the worst case interference. The consequence of assumptions 1 and 2 is that $\gamma_i(x)$ is deterministic.

2) CRE User Association Rule: A user association rule using CRE is commonly used in the heterogeneous networks [3]–[13]. According to this rule, a user located at x is served by the BS that provides the highest biased received power and the SINR provided by it is greater than the minimum required SINR ($\gamma_{\min}$). The region $\mathcal{D}_i(\bar{c})$ served by BS i is defined as:

$$\mathcal{D}_i(\bar{c}) = \{x | \forall j \in \mathcal{S}, P_i g_i(x) c_i \ge P_j g_j(x) c_j, \gamma_i(x) \ge \gamma_{\min}\}. \quad (2)$$

Note that $\mathcal{D}_i(\bar{c})$ is a bounded region because of $\gamma_{\min}$. According to assumptions 1 and 2, the association rule should be understood for a given realisation of the shadowing mask and averaged over fast fading variations.

3) Physical Data Rate: The physical data rate received by a user served by BS i with SINR γ_i is denoted as $\nu_i(\gamma_i)$, which is a non-negative and a non-decreasing function of SINR.

4) Traffic Model: Users are assumed to arrive in the system according to a spatial random process, download a file of random size and leave the system when the download is over. This is referred to as elastic traffic. At location x, the arrival rate is denoted $\lambda(x)$ [arrivals/s/m^2] and the average file size is $1/\mu(x)$ [bits]. If x is associated to BS i, the load generated by x on BS i is $\varrho_i(x) \triangleq \frac{\lambda(x)}{\mu(x)\nu_i(\gamma_i)}$. Following [29], [33], we model every BS i as a M/G/1/PS queue of load:

$$\rho_i(\bar{c}) = \int_{x \in \mathcal{D}_i(\bar{c})} \varrho_i(x) dx. \quad (3)$$

BS i is stable if and only if $0 \le \rho_i < 1$. In this work, only stable network is considered. The *flow throughput* of users is defined as the ratio of the mean file size to the mean file download duration [33].

Assumption 3: [Time-scale separation] The process of updating the CRE parameters is supposed to be long with respect to the traffic variations. The M/G/1/PS queues describing the BSs traffic are thus supposed to have reached their stationary regime before any new change of the CRE parameters.

B. Problem Formulation and Objective Function

Following [29], we intend to minimize an $\alpha-$fairness function $\phi_\alpha(\bar{c})$ over the feasible set $\mathcal{F}$, which are given below.

$$\phi_\alpha(\bar{c}) = \begin{cases} \sum_{i \in \mathcal{S}} \frac{(1-\rho_i(\bar{c}))^{1-\alpha}}{\alpha-1}, & \alpha \ge 0, \alpha \ne 1, \\ -\sum_{i \in \mathcal{S}} \log(1-\rho_i(\bar{c})), & \alpha = 1, \end{cases} \quad (4)$$

$$\mathcal{F} = \{\rho | 0 \le \rho_i(\bar{c}) < 1, c_i \in [1, c_{\max}], \forall i \in \mathcal{S}\}. \quad (5)$$

The function $\phi_\alpha(\bar{c})$ is in general non-convex and even if it is convex the set $\mathcal{F}$ is non-convex because $\bar{c}$ takes discrete values. The function $\phi_\alpha(\bar{c})$ captures various aspects of fairness and so of load balancing of the network for different values of α, which are described below.

($\alpha = 0$) **Rate-optimal policy:** Minimising $\phi_0(\bar{c})$ gives a rate-optimal policy. See Appendix A for the proof.

($\alpha = 1$) **Proportional fair policy:** In this case, $\phi_1(\bar{c})$ captures the proportional fairness of the network [32].

($\alpha = 2$) **Delay-optimal policy:** It can be shown that minimising $\phi_2(\bar{c})$ corresponds to minimising the average number of flows of the network. *Little's law* says that minimising the average number of flows is equivalent to minimising the average delay experienced by a typical flow. Therefore, minimising $\phi_2(\bar{c})$ is equivalent to minimising the average delay of the network. For more detailed discussion refer to [29].

($\alpha \to \infty$) **Minmax policy:** As $\alpha \to \infty$ the minimiser of $\phi_\alpha(\bar{c})$ tends to the min-max load vector. It is a standard result with convex objective functions [29], [32], [34]. We prove this result for our non-convex objective function in Theorem 1.

Definition 1: [Min-max load vector [34]] Let all the vectors in $\mathcal{F}$ be sorted in increasing order. A vector $\rho \in \mathcal{F}$ is min-max if ρ is lexicographically not greater than any vector in $\mathcal{F}$. The vector ρ is *lexicographically lower* than y, denoted $\rho \prec y$, if the first non-zero component of $\rho - y$ is negative. We say that ρ is not greater than y, denoted by $\rho \preceq y$, if $\rho \prec y$ or $\rho = y$. Let $r_i(\bar{c}) = 1 - \rho_i(\bar{c}), \forall i \in \mathcal{S}$. Let $\mathcal{X} = \{r \in \mathcal{R}^{|\mathcal{S}|} | \exists \bar{c} : \rho(\bar{c}) \in \mathcal{F}, r(\bar{c}) = r\}$. Load vector ρ^* is a min-max if and only if r^* is max-min vector.

Theorem 1: Let $r_\alpha \in \underset{r \in \mathcal{X}}{\operatorname{argmax}} \sum_{i \in \mathcal{S}} \frac{r_i^{1-\alpha}}{1-\alpha}$. Then, any accumulation vector of the trajectory $\{r_\alpha\}_{\alpha > 1}$ is max-min in $\mathcal{X}$.

Proof The proof is given in Appendix B.

In Fig. 1, we show an example of set $\mathcal{F}$ obtained with 2 BSs having different transmit powers located on a two-dimensional region. It is clear from the figure that even if the CRE set were continuous, $\mathcal{F}$ would not be convex. We also show the
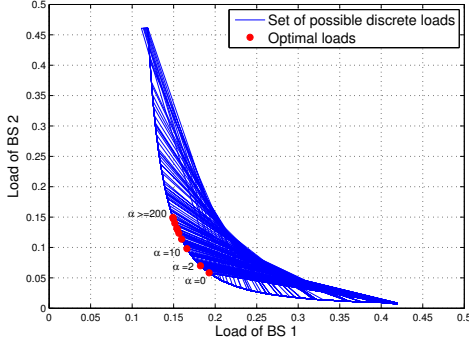
Fig. 1: Feasible set $\mathcal{F}$ for 2 BSs.

optimal loads obtained for different α values. All the optimal load points are located on the Pareto frontier. The point for $\alpha \geq 200$ in Fig. 1 is the min-max load point because a point of equal coordinates on the Pareto frontier is the min-max point.

III. POTENTIAL GAME FRAMEWORK

In this section, we present an approach using potential game framework for distributed optimisation of the objective function. We do not aim to describe and analyse selfish nature of the BSs that aim to minimise their costs. Rather, our goal is to achieve the global objective of load balancing by prescribing a cost function to the BSs. For this context, potential games provide a good framework because players of such a game distributively optimise the potential function.

We model the problem as a game, where the BSs are players and allowed CRE bias values are their strategies. Formally, a *CRE game* is defined by the tuple $\Gamma = \{\mathcal{S}, \{X_i\}_{i \in \mathcal{S}}, \{U_i\}_{i \in \mathcal{S}}\}$, where $\mathcal{S}$ is the set of BSs, $X = X_1 \times X_2 \times \ldots \times X_{|\mathcal{S}|}$ is a strategy set or an action set, and $U_i : X \rightarrow \mathcal{R}$ is a cost function. X_i is a discrete set of CRE bias values ranging from 1 to $c_{\max}$. The BSs play the CRE game with the objective of minimising their costs. A pure NE (PNE) of the game is reached when no player can benefit by changing its strategy unilaterally.

Definition 2: [Pure Nash equilibrium] A PNE is a vector $\bar{c}^*$ of bias values in strategy set X. Given the other BSs' equilibrium strategies c^*_{-i} the PNE strategy of BS i is c^*_i if and only if $c^*_i \in \arg\min_{c_i \in X_i} U_i(c_i, c^*_{-i}), \forall i \in \mathcal{S}$.

Definition 3: [Exact potential game [35]] If there is a function called potential function $P : X \rightarrow \mathcal{R}$ such that $\forall i \in \mathcal{S}, \forall c_i, c'_i \in X_i$ and $\forall c_{-i} \in X_{-i}$,

$$U_i(c_i, c_{-i}) - U_i(c'_i, c_{-i}) = P(c_i, c_{-i}) - P(c'_i, c_{-i}), \quad (6)$$

then the game is an exact potential game.

An exact potential game has at least one PNE and local optimisers of the potential function are PNEs [35]. Furthermore, there exist several distributed learning algorithms that converge to a PNE when the game has a potential (see next Section). In our problem, we intend to turn our objective function (4) into a potential function by designing cost functions of the BSs.

Two ways of achieving it are by using identical interest utility (IIU) and WLU [36].

Identical interest utility: In this structure, the cost function $U_i(\bar{c})$ is completely aligned with the potential function, i.e., $U_i(\bar{c}) = \phi_\alpha(\bar{c}), \quad \forall i \in \mathcal{S}$. However, it requires the BS i to know the loads of all the other BSs in order to compute $U_i(\bar{c})$.

Wonderful life utility: With the WLU, a BS needs only to know the loads of its neighbour BSs. The WLU structure of the individual BSs is defined as

$$U_i(c_i, c_{-i}) = \sum_{j \in N_i} \frac{(1 - \rho_j(c_i, c_{-i}))^{1-\alpha}}{\alpha - 1}, \quad (7)$$

where N_i is the neighbour set of BS i,

$$N_i = \bigcup_{\bar{c}} \{j \in \mathcal{S} | \exists x \in \mathcal{D}_i(\bar{c}), P_i g_i(x) c_i = P_j g_j(x) c_j\}. \quad (8)$$

A neighbour set N_i is all possible BSs that share boundary with BS i for at least one possible bias value. The motivation behind WLU is that the action of the BS i only affects its neighbour BSs.

Lemma *1:* The WLU cost function (7) leads to an exact potential game.

Proof: Consider the strategy profiles $a = (c_i, c_{-i})$ and $b = (c'_i, c_{-i})$. From the objective function (4) we have

$$\phi_\alpha(a) - \phi_\alpha(b) = \sum_{j \in \mathcal{S}} \frac{(1 - \rho_j(a))^{1-\alpha}}{\alpha - 1} - \sum_{j \in \mathcal{S}} \frac{(1 - \rho_j(b))^{1-\alpha}}{\alpha - 1}, \quad (9)$$

$$= \sum_{j \in N_i} \frac{(1 - \rho_j(a))^{1-\alpha}}{\alpha - 1} - \sum_{j \in N_i} \frac{(1 - \rho_j(b))^{1-\alpha}}{\alpha - 1}. \quad (10)$$

The second equation above is true because the CRE bias of the BS i will only affect its neighbour BSs. Hence, the utility of the BS i is aligned with the potential function and the resultant game is an exact potential game. ∎

IV. DISTRIBUTED LEARNING ALGORITHMS

Recall that the potential function property enables finding a PNE through distributed learning algorithms. In this section, we introduce distributed learning algorithms that are used to find the PNE of the CRE game. First, we present the BR algorithm and the LLLA for the complete information setting. Next, the BLLLA for the partial information setting is described.

A. Best Response Algorithm

Best response algorithm is an asynchronous algorithm where at any given time only a single BS updates its strategy. Assume a time-varying random process with which a BS is chosen to revise its strategy[1]. At any time step t the selected player i chooses a strategy c_i that minimises his cost, given the strategies c_{-i} of other players. In other words, player i chooses a strategy from his best response set B_i,

$$B_i(c_{-i}) = \arg\min_{c_i} U_i(c_i, c_{-i}). \quad (11)$$

---

[1]Uniform probability or stationarity of the process is not required, it is only required that the probability of selecting any player is positive.

Note that the BR algorithm requires complete information, in which the effects of choosing all the other strategies are known, and it is not guaranteed to converge to the optimal NE [35].

B. Log-linear Learning Algorithm

LLLA is a generalisation of BR. It is summarised in Algorithm 1. At each step the LLLA deviates from BR with a probability that tends to zero as the parameter τ goes to zero. However, as τ tends to infinity the LLLA selects actions randomly with uniform probability. It guarantees the convergence to the optimal NE with probability that tends to one as τ goes to zero [37]. However, for this algorithm the BSs require again the complete information. For example, given the strategies of others, the BS has to know the cost function value for all its strategies. With this information, it selects a strategy to play according to a probability distribution. In general, acquiring this amount of information is not feasible. To overcome this difficulty in the next subsection we propose to use BLLLA.

Algorithm 1 Log-linear Learning Algorithm [37]

1: **Initialisation:** Arbitrary set CRE bias $c_i \ \forall i \in \mathcal{S}$.
2: Set parameter τ.
3: **While** $t \geq 1$ **do**
4: Randomly select a BS i.
5: Select its CRE $c_i(t)$ from X_i with probability $p_i^{c_i}(t)$,

$$p_i^{c_i}(t) = \frac{\exp\left(-\frac{1}{\tau}U_i\left(c_i, c_{-i}(t-1)\right)\right)}{\sum_{c_i' \in X_i} \exp\left(-\frac{1}{\tau}U_i\left(c_i', c_{-i}(t-1)\right)\right)}. \quad (12)$$

6: All the other BSs must repeat their previous actions, i.e., $c_{-i}(t) = c_{-i}(t-1)$.

C. Binary Log-linear Learning Algorithm

The BLLLA works even if only partial information about the game is available to the players. Partial information is the information that a player has about its current strategy. Unlike complete information the effect of choosing any other strategy is not known to the player. As LLLA the BLLLA is also an asynchronous algorithm. In this algorithm, whenever the BS updates its strategy it does it in two steps. In the first step, the BS tries a strategy from its strategy set to obtain its payoff. In the second step, the BS randomly chooses among the two strategies (present strategy and trial strategy) as summarised in Algorithm 2.

D. Effect of time-varying neighbours

For all the above algorithms, the BS i needs to know its neighbours N_i to calculate its cost using the WLU structure. For a given CRE bias vector, providing to every BS the neighbours set is a standard task for network operators. This task can be performed automatically e.g., using automated neighbour relation (ANR) standardised by 3GPP [1]. The difficulty here is to deal with time-varying neighbourhood. If neighbourhood is changing then WLU doesn't lead to a

---

Algorithm 2 Binary Log-linear Learning Algorithm [37]

1: **Initialisation:** Arbitrary set CRE bias $c_i \ \forall i \in \mathcal{S}$.
2: Set parameter τ.
3: **While** $t \geq 1$ **do**
4: Randomly select a BS i.
5: Select a trial action $\hat{c}_i \in X_i$ with uniform probability. Play $\hat{c}_i$ and observe its cost.
6: Play the action $c_i(t) \in \{c_i(t-1), \hat{c}_i\}$ as given below.

$$c_i(t) = \begin{cases} c_i(t-1), & \text{w.p. } \frac{e^{-\frac{1}{\tau}U_i(\bar{c}(t-1))}}{e^{-\frac{1}{\tau}U_i(\bar{c}(t-1))}+e^{-\frac{1}{\tau}U_i(\hat{c}_i, c_{-i}(t-1))}}, \\ \hat{c}_i, & \text{w.p. } \frac{e^{-\frac{1}{\tau}U_i(\hat{c}_i, c_{-i}(t-1))}}{e^{-\frac{1}{\tau}U_i(\bar{c}(t-1))}+e^{-\frac{1}{\tau}U_i(\hat{c}_i, c_{-i}(t-1))}}. \end{cases}$$
(13)

7: All the other BSs must repeat their previous actions, i.e., $c_{-i}(t) = c_{-i}(t-1)$.

TABLE I: Simulation parameters.

Parameter	Variable	Value
Number of BSs	N_s	8
Transmit power of macro BS	P_{macro}	46 dBm
Transmit power of small BS	P_{small}	24 dBm
Average file size	$\frac{1}{\mu}$	0.5 Mbytes
Average traffic load density	$\frac{\lambda}{\mu}$	64 $bits/s/m^2$
System bandwidth	W	20 MHz
Noise power	N_0	-174+10log(W) dBm
Minimum SINR	γ_{min}	-10 dB
Path-loss exponent	η	3.5
CRE bias set	c_i	$\{1, 1.1, 1.2, \ldots, 16\}$

potential game. To address this problem, we propose the following technique.

The algorithms above can use the WLU structure without knowing any neighbour set at the start of the algorithm. In the process of learning, whenever BS i updates its strategy if it finds new neighbours then it should remember and include them in N_i. As the CRE bias set is finite, there is a time instant after which all neighbourhood sets $N_i, \forall i$, are constant. Thus, from this time instant, the game becomes a potential game and the above algorithms will converge to the optimal PNE.

V. SIMULATION RESULTS

In this section, we show simulation results considering standard parameters as adopted in 3GPP [2]. These parameters are listed in Table I. The region $\mathcal{L}$ considered is a square of side 1000 m. Among the 8 BSs located in $\mathcal{L}$, BS 1 at the center is a macro BS that transmits with P_{macro} and the rest are small BSs that transmit with P_{small}. The user traffic is fixed in time but varies with location across an average traffic density of 64 $bits/s/m^2$. There are two hotspots where the traffic is 5 times the average traffic, which can be seen in Fig. 2.

The following channel model is used in the simulations [38].

$$g_i(x) = \min\left\{1, K \left|x_i - x\right|^{-\eta}\right\}, \quad (14)$$

where K is a constant, $|x_i - x|$ is the distance between the location x_i of the BS i and the location x of the user, and η is

(a) ($\alpha = 0$) Rate-optimal policy.

(b) ($\alpha = 2$) Delay-optimal policy.
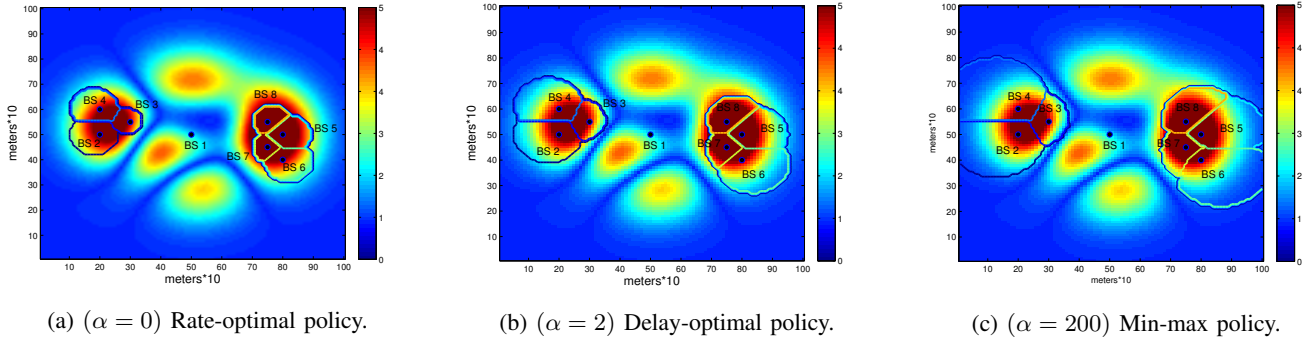
(c) ($\alpha = 200$) Min-max policy.

Fig. 2: The variations of the coverage regions of BSs obtained using the optimal CRE for different α. Varying traffic, which is normalised with the average traffic of 64 Kbits/s/m^2, is shown using varying intensity of colours.

the path-loss exponent. We use the classical Shannon formula for calculating channel capacity $\nu_i(\gamma_i)$ at any location x.

The BSs play CRE game and learn the optimal NE using the proposed learning algorithms. The optimal coverage regions obtained using the optimal CRE bias values for different α are shown in Fig. 2. The corresponding optimal bias values and loads are shown in Table II. In Fig. 2a, the rate-optimal policy, which is obtained for $\alpha = 0$, is shown. The optimal CRE bias values of all the small BSs is close to one. This is intuitive because for rate-optimal policy the bias values should be equal to one. If we take more iterations of the algorithms then the optimal bias values are guaranteed to converge to one. Therefore, this case corresponds to the classical user association without the use of CRE bias leading to a heavy load imbalance. We can observe that the load of the macro BS 1 is 92%, which is near an overload. On the other hand, the loads of all the small BSs are less than 12%, which is a heavy under-utilisation. This case serves as a benchmark for other cases of α to compare for load balancing.

In Fig. 2b, the coverage regions for delay-optimal policy, which is obtained for $\alpha = 2$, is shown. As α increases to 2 the coverage regions of all small BSs increase and that of the macro BS decreases. This happens due to the increase in the optimal bias values. The load of the macro BS is decreased to 62% and the utilisation of small BSs is increased.

In Fig. 2c, the coverage regions for the min-max policy, which is obtained for $\alpha = 200$, is shown. When α increases to higher value, here $\alpha = 200$, the load of the macro BS 1 is further reduced to 42%. It can be observed that the loads of all the BSs are equalised except for BS 3 and 7. The utilisation of these BSs cannot be increased further because these BSs are near to the macro BS 1 that causes heavy downlink interference to their users.
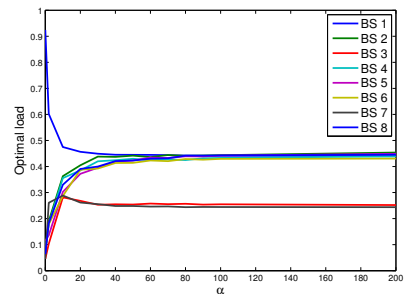
The evolution of the optimal CRE biases and optimal loads for α values ranging from 0 to 200 is shown in Fig. 3. The figure shows how the optimality changes from rate-optimal to min-max optimal as α increases from 0 to 200. Therefore, initially at $\alpha = 0$, the optimal bias of all BSs is close to one and correspondingly there is a heavy load imbalance. As α increases, the load of the macro BS decreases and that of

TABLE II: Comparison of optimal CRE, optimal loads of BSs for different α.

BS i	$\alpha = 0$		$\alpha = 2$		$\alpha \to \infty$	
	c_i^*	$\rho_i^*\%$	c_i^*	$\rho_i^*\%$	c_i^*	$\rho_i^*\%$
1	1	92	1	62	1	42
2	1.1	9	3.1	20	16	51
3	1	4	3.6	11	16	21
4	1	7	2.8	17	14.8	49
5	1.1	12	3.4	23	7.7	42
6	1.1	8	3.4	20	7.7	41
7	1.1	5	3.5	12	16	25
8	1	6	3.2	18	7.2	42
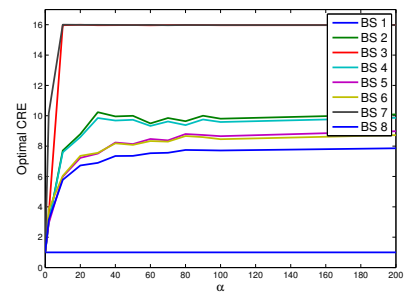
(a) Optimal CRE bias.

(b) Optimal loads.

Fig. 3: Evolution of optimal CRE and optimal load with α.

(a) ($\alpha = 0$) Rate-optimal policy. (b) ($\alpha = 2$) Delay-optimal policy. (c) ($\alpha = 200$) Min-max policy.
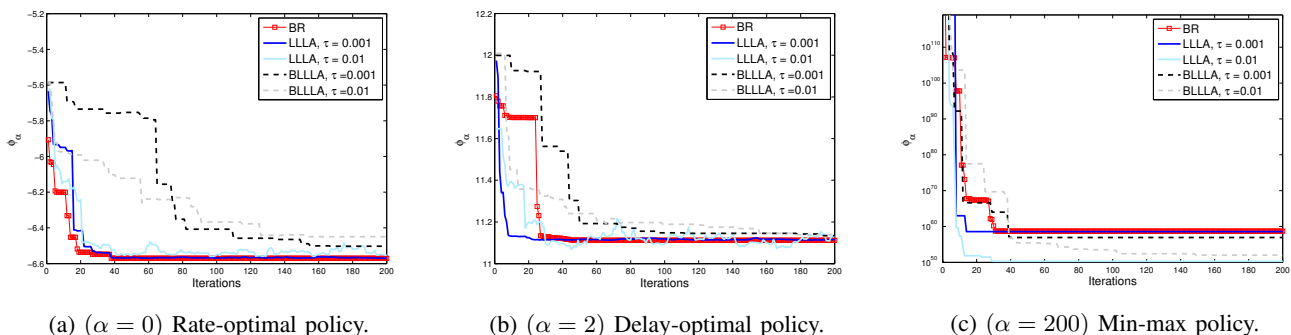
Fig. 4: Convergence of BR, LLLA, and BLLLA.

the small BSs increases. At around $\alpha = 50$ the min-max load policy is reached. It is observed that different BSs have different optimal bias values for different α.

The evolution of the objective function for different algorithms is shown in Fig. 4. In all the cases, we know that the LLLA and the BLLLA converge to the global minimum. In Fig. 4a, we observe that BR converges much faster than the LLLA and the BLLLA, whereas it requires complete information. On the other hand, the BLLLA requires only partial information, which is an advantage for practical implementations but does not loose much in terms of convergence speed. The figure also shows that smaller values of τ result in faster convergence for the LLLA and the BLLLA.

In Fig. 4b, we see that the BR does not converge to the global minimum because the LLLA for $\tau = 0.01$ sometimes achieves lower values of potential. Also, the LLLA for $\tau = 0.001$ converges to the same local minimum as BR because for smaller τ the LLLA behaviour is similar to the BR. However, the LLLA and the BLLLA for $\tau = 0.01$ converge to the global minimum. In Fig. 4c, it is clear that the BR and the LLLA for $\tau = 0.001$ converge to a local minimum, whereas both the LLLA and the BLLLA for $\tau = 0.01$ converge to the global minimum.

VI. CONCLUSIONS

In this paper, a novel approach for load balancing using CRE association technique is presented. Our approach exploits the potential game structure and distributed learning algorithms. By running extensive simulations in two settings, which are complete and partial information settings, we show that the proposed algorithms converge within a few tens of iterations to the optimal NE, which is also a minimiser of a $\alpha-$fairness function of the network. The convergence speed of the BLLLA that uses partial information is comparable to the LLLA that uses complete information, meaning that partial information is sufficient in practical implementations. We also show that the classical BR algorithm does not necessarily converge to the optimal NE, although it necessitates complete information. As future work we intend to extend our work for the following: 1) noisy estimation of the BSs utilities, 2) admission control policies to overcome the overload situation, 3) effect of using advanced interference management techniques in conjunction with CRE association, and 4) adaption of the algorithms to varying traffic.

REFERENCES

[1] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access (E-UTRAN); Overall description; Stage 2," TS 36.300, 3GPP, Sept. 2008.

[2] 3GPP, "Technical specification group radio access network; evolved universal terrestrial radio access (eutra); further advancements for e-utra physical layer aspects," TR 36.814, 3GPP, Mar. 2010.

[3] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," IEEE Wireless Commun., vol. 18, no. 3, pp. 10–21, 2011.

[4] J. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," IEEE Wireless Commun., vol. 21, pp. 18–25, April 2014.

[5] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous LTE-A cellular networks," IEEE J. Sel. Areas Commun., vol. 28, pp. 1479–1489, Dec. 2010.

[6] K. Okino, T. Nakayama, C. Yamazaki, H. Sato, and Y. Kusano, "Pico cell range expansion with interference mitigation toward lte-advanced heterogeneous networks," in ICC Workshop, pp. 1–5, June 2011.

[7] 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," in ISWCS, pp. 844–848, Nov. 2011.

[8] M. Vajapeyam, A. Damnjanovic, J. Montojo, T. Ji, Y. Wei, and D. Malladi, "Downlink FTP performance of heterogeneous networks for LTE-advanced," in ICC Workshop, pp. 1–5, June 2011.

[9] I. Guvenc, "Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination," IEEE Commun. Lett., vol. 15, pp. 1084–1087, Oct. 2011.

[10] A. Morimoto, N. Miki, H. Ishii, and D. Nishikawa, "Investigation on transmission power control in heterogeneous network employing cell range expansion for LTE-advanced uplink," in European Wireless Conf., pp. 1–6, Apr. 2012.

[11] J. Oh and Y. Han, "Cell selection for range expansion with almost blank subframe in heterogeneous networks," in Proc. PIMRC, pp. 653–657, Sept 2012.

[12] K. Kitagawa, T. Komine, T. Yamamoto, and S. Konishi, "Performance evaluation of handover in LTE-advanced systems with pico cell range expansion," in Proc. PIMRC, pp. 1071–1076, Sep. 2012.

[13] M. Eguizabal and A. Hernandez, "Interference management and cell range expansion analysis for LTE picocell deployments," in Proc. PIMRC, pp. 1592–1597, Sept 2013.

[14] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. Andrews, "User association for load balancing in heterogeneous cellular networks," IEEE Trans. Wireless Commun., vol. 12, pp. 2706–2716, June 2013.

[15] E. Stevens-Navarro, Y. Lin, and V. W. Wong, "An MDP-based vertical handoff decision algorithm for heterogeneous wireless networks," IEEE Trans. Veh. Technol., vol. 57, no. 2, pp. 1243–1254, 2008.

[16] S.-E. Elayoubi, E. Altman, M. Haddad, and Z. Altman, "A hybrid decision approach for the association problem in heterogeneous networks," in *Proc. INFOCOM*, pp. 1–5, 2010.

[17] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 2008–2017, May 2009.

[18] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "RAT selection games in HetNets," in *Proc. INFOCOM*, pp. 998–1006, 2013.

[19] P. Ökvist and A. Simonsson, "LTE HetNet trial - range expansion including micro/pico indoor coverage survey," in *Proc. VTC (Fall)*, pp. 1–5, Sep 2012.

[20] P. Tian, H. Tian, J. Zhu, L. Chen, and X. She, "An adaptive bias configuration strategy for range extension in LTE-advanced heterogeneous networks," in *ICCTA* , pp. 336–340, Oct 2011.

[21] K. Kikuchi and H. Otsuka, "Proposal of adaptive control CRE in heterogeneous networks," in *Proc. PIMRC*, pp. 910–914, Sep 2012.

[22] J. Wang, J. Liu, D. Wang, J. Pang, and G. Shen, "Optimized fairness cell selection for 3GPP LTE-A macro-pico HetNets," in *Proc. VTC (Fall)*, pp. 1–5, Sep 2011.

[23] M. Al-Rawi, "A dynamic approach for cell range expansion in interference coordinated LTE-advanced heterogeneous networks," in *Proc. ICCS*, pp. 533–537, Nov 2012.

[24] I. Siomina and D. Yuan, "Load balancing in heterogeneous LTE: Range optimization via cell offset and load-coupling characterization," in *Proc. ICC*, pp. 1357–1361, June 2012.

[25] T. Kudo and T. Ohtsuki, "Cell range expansion using distributed Q-learning in heterogeneous networks," *EURASIP J. Wireless Commun. Netw.*, pp. 1–10, Mar. 2013.

[26] O. K. Tonguz and E. Yanmaz, "The mathematical theory of dynamic load balancing in cellular networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 12, pp. 1504–1518, 2008.

[27] K. Son, S. Chong, and G. Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3566–3576, 2009.

[28] B. Rengarajan and G. De Veciana, "Architecture and abstractions for environment and traffic-aware system-level coordination of wireless networks," *IEEE Trans. Netw. ACM*, vol. 19, no. 3, pp. 721–734, 2011.

[29] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed alpha-optimal user association and cell load balancing in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 20, pp. 177–190, Feb 2012.

[30] F. Xu, C. C. Tan, Q. Li, G. Yan, and J. Wu, "Designing a practical access point association protocol," in *Proc. INFOCOM*, pp. 1–9, 2010.

[31] Y. Bejerano and S.-J. Han, "Cell breathing techniques for load balancing in wireless LANs," *IEEE Trans. Mobile Comput.*, vol. 8, pp. 735–749, June 2009.

[32] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, 2000.

[33] T. Bonald and A. Proutière, "Wireless downlink data channels: User performance and cell dimensioning," in *In Proc. ACM International Conf. Mobile Comput. Netw.*, MobiCom '03, pp. 339–352, Sept. 2003.

[34] T. Bonald and L. Massoulié, "Impact of fairness on internet performance," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 29, pp. 82–91, 2001.

[35] D. Monderer and L. S. Shapley, "Potential Games," *Games and Economic Behavior*, vol. 14, pp. 124–143, May 1996.

[36] N. Li and J. R. Marden, "Designing games for distributed optimization," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 2, pp. 230–242, 2013.

[37] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and a payoff-based implementation," *Games and Economic Behaviour*, vol. 75, pp. 788–808, July 2012.

[38] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

APPENDIX

A. Proof of Rate-optimal Policy

It can be shown that the minimiser of $\phi_\alpha(c)$ is same as the minimiser of the arithmetic mean of the BSs loads, i.e., $\sum_{i \in |\mathcal{S}|} \rho_i(\bar{c})$. We will prove that minimising $\phi_0(c)$ is indeed rate optimal policy by contradiction. Consider that at the minimum ϕ_0^* a location x_0 is associated with a BS j with rate $r_j(x_0)$, which is not the maximum. Then, there exist a

BS k that provides the maximum rate $r_k(x_0)$. Let ϕ_0^k be the value of the objective function when the UE at location x_0 is associated with BS k. The difference of the objective function due to the loads of BSs j and k is

$$\phi_0^k - \phi_0^* = \frac{\lambda(x_0)}{\mu(x_0)} \left(\frac{1}{r_k(x_0)} - \frac{1}{r_j(x_0)} \right) < 0, \quad (15)$$

which is a contradiction that ϕ_0^* is the minimum. Therefore, we conclude that at ϕ_0^* all the locations will be served by the BS that provide the highest rate.

B. Proof of Theorem 1

The proof is divided into the following two lemmas. Lemma 2 gives the property that is required for proving the Lemma 3, which concludes the proof of the Theorem 1.

Lemma 2: Let $f_\alpha(r) = \sum_{i \in \mathcal{S}} \frac{r_i^{1-\alpha}}{1-\alpha}$. If $r \succ y$ then there is $A > 0$ large enough such that for all $\alpha \geq A$, $f_\alpha(r) > f_\alpha(y)$.

Proof: Let $\alpha > 1$. Without loss of generality, assume that r and y are sorted in increasing order and that $r_1 > y_1$. Let $\delta = r_1 - y_1$. Then

$$(1-\alpha)(f_\alpha(r) - f_\alpha(y)) = \sum_{i=1\ldots n} \left(r_i^{1-\alpha} - y_i^{1-\alpha} \right) \quad (16)$$

$$\leq r_1^{1-\alpha} - y_1^{1-\alpha} + \sum_{i=2\ldots n} \left(r_i^{1-\alpha} - y_i^{1-\alpha} \right) \quad (17)$$

$$\leq r_1^{1-\alpha} - (r_1 - \delta)^{1-\alpha} + (n-1)r_1^{1-\alpha} - \sum_{i=2\ldots n} y_i^{1-\alpha} \quad (18)$$

$$\leq nr_1^{1-\alpha} - (r_1 - \delta)^{1-\alpha}. \quad (19)$$

Then we have $f_\alpha(r) > f_\alpha(y)$ if and only if

$$nr_1^{1-\alpha} - (r_1 - \delta)^{1-\alpha} \leq 0 \quad (20)$$

$$\Leftrightarrow \frac{\log(n)}{\log\left(\frac{r_1 - \delta}{r_1}\right)} \geq 1 - \alpha \quad (21)$$

$$\Leftrightarrow 1 + \frac{\log(n)}{\log(r_1) - \log(r_1 - \delta)} \leq \alpha. \quad (22)$$

∎

Lemma 3: Let $\mathcal{X}$ be a compact subset of $\mathbb{R}^{|\mathcal{S}|}$. Consider the set

$$S = \bigcap_{A>1} \overline{\bigcup_{\alpha \geq A} \operatorname{argmax}_{x \in \mathcal{X}} f_\alpha(x)}.$$

Then S is non-empty and is made of max-min vectors in $\mathcal{X}$.

Proof:

Let $S_A = \overline{\bigcup_{\alpha \geq A} \operatorname{argmax}_{x \in \mathcal{X}} f_\alpha(x)}$. It is a decreasing nested sequence of non-empty compact sets. By Cantor's intersection theorem, it is not empty and compact.

Let $x^* \in \bigcap_{A>1} S_A$. There is an increasing sequence $\alpha(n)$ and $x_{\alpha(n)} \in \operatorname{argmax}_{x \in \mathcal{X}} f_{\alpha(n)}(x)$ with $x_{\alpha(n)} \to x^*$. Assume there is $y \succ x^*$. Then, by Prop. 2, there is N such that, for every $n \geq N$, $f_{\alpha(n)}(y) > f_{\alpha(n)}(x^*)$. But, by definition of $x_{\alpha(n)}$, $f_{\alpha(n)}(x_{\alpha(n)}) \geq f_{\alpha(n)}(y)$, which is a contradiction with the fact that $f_{\alpha(n)}(x_{\alpha(n)}) \to f_{\alpha(n)}(x^*)$, which is ensured by Berge maximum theorem.

∎