

Utility Fair RAT Selection in Multi-Homed LTE/802.11 Networks

Bahar Partov, Douglas J. Leith
Trinity College Dublin, Ireland

Abstract—We consider proportional fair rate allocation in a heterogeneous network with a mix of LTE and 802.11 Random Access Technologies (RATs) which supports multipath and multi-homed operation (simultaneous connection of a user device to multiple LTE BSs and one 802.11 AP). We show that the utility fair optimisation problem is non-convex but that a global optimum can be found by solving a sequence of convex optimisations in a distributed fashion. The result is a principled approach to offload from LTE to 802.11 and for exploiting LTE/802.11 path diversity to meet user traffic demands.

I. INTRODUCTION

In this paper we consider rate allocation in a heterogeneous network with a mix of LTE and 802.11 cells. Integrated design of LTE and 802.11 is topical in view of the continuing increases in data traffic, the fact that many cellular network operators also operate a large network of 802.11 hotspots and that user handsets are now typically equipped with both LTE and 802.11 interfaces. Rather than an either/or handover-like setting where the question is which network to use, our interest is instead in settings where user devices jointly use the LTE and 802.11 networks and may send data across both simultaneously. Further, we consider situations where user devices may connect to multiple LTE Base Stations (BSs) and one 802.11 Access Points (AP) simultaneously. This allows us to encompass the Coordinated Multi Point (CoMP) features in release 11 of LTE [1] which allow coordinated transmission and reception across multiple BSs. Simultaneous transmission across multiple interfaces might be implemented by routing each connection across one or other network (in a manner akin to load balancing) or by striping connections across both networks (e.g. via use of transport layer protocols such as multi-path TCP [3]).

In view of the existing work, the authors in [4] provide a survey on network selection in HetNets. Where network selection is the task of assigning RATs to users so that a network utility is maximised. In this context, two modes of UE operation are typically considered: multi-modal and multi-homed. Although both multi-mode and multi-home operations support multiple radio interfaces, only the latter supports multiple TCP flows across disjoint paths. To solve the network selection problem, the authors in [5] and [6] consider combinatorial optimisation, [5] considers a linear utility function in a multi-mode UMTS/GSM network and [6] a linear utility function which is the sum of logarithmic

utilities of individual BS-UE pairs. In [7] a network-user association problem in a WLAN/UMTS hybrid network for multi-mode operations is solved using Markov Decision Processes (MDPs), although the complexity scales poorly with network size [8]. In [9] a non-cooperative game is formed among users of two different classes corresponding to 3G/LTE and 802.11.

Our focus in this paper is on how to allocate the available LTE and 802.11 bandwidth amongst user devices in a heterogeneous network, and in particular how to determine a proportional fair rate allocation. Our main contributions are as follows. Firstly, we develop a throughput model for heterogeneous networks that include both LTE and 802.11 links and which encompasses multipath and multi-homed operation (simultaneous connection of a user device to multiple LTE BSs and 802.11 APs). We show that although the rate region is log-convex, and is also not log-convex, the utility fair optimisation problem is non-convex. Our second contribution is therefore a sequential convex optimisation approach, based on determining a sequence of convex subsets, that we show converges to a global optimum. This optimisation approach is suited to distributed implementation. Thirdly, we present a number of application examples demonstrating how this framework can be used to develop principled approaches to offload from LTE to 802.11 and for exploiting LTE/802.11 path diversity to meet user traffic demands.

A. Motivating Example

Consider the simple example in Fig 1, where the network consists of one cellular BS and one 802.11 AP. UE u_2 is located close to the 802.11 AP and so uses a physical rate of 54Mbps. UE u_1 is further from the AP and so uses a lower physical rate of 1Mbps. Both UEs use the same physical rate of 10Mbps for their LTE connection. The physical rates available on the 802.11 and LTE links are summarised in Table I. For simplicity, we assume both UEs are saturated *i.e.* always have a packet to send.

When the UEs are contending for the available resources, the achievable throughput is determined by the channel access method. For 802.11 the random access approach leads to collisions and reduced throughput when more than one station contends for access. In the case of LTE, time-frequency slots are allocated to UEs by the cell BS in a **Round-Robin manner**. The achievable data rates using individual random access technologies by both UEs are summarised in Table I.

For this example we later show that if u_1 sends all traffic via the BS, and u_2 sends all traffic via the AP, rates of 10 Mbps

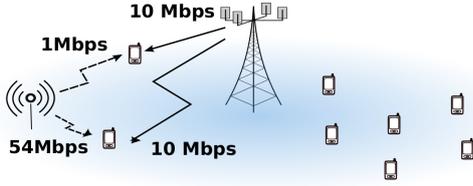


Fig. 1: Example of a mixed LTE/WiFi network, with one 802.11 AP, one LTE BS and two UEs u_1 and u_2 each equipped with both an 802.11 and an LTE interface.

and 47.5 Mbps can be achieved by the UEs and that this is the proportional fair rate allocation. This is summarised in Table I. Note that simply splitting the traffic for each UE equally over the 802.11 and LTE networks would yield rates of 5.46Mbps and 29.9Mbps *i.e.* worse for *both* users than the optimised allocation.

TABLE I: Motivating Example: Achievable data rates

	PHY Rates [Mbps]		Technology Rates [Mbps]		
	BS	AP	LTE Only	802.11 Only	Optimised Multi-RAT
u_1	10	1	5	0.46	10
u_2	10	54	5	24.9	47.5

II. NETWORK MODEL

A. Network topology

We consider a multi-RAT setup where UEs can potentially connect to multiple LTE base stations and one 802.11 AP simultaneously.

Let a identify the 802.11 access point and $\mathcal{U} := \{1, 2, \dots, U\}$ the set of users. Associated with access point a is the set of users $\mathcal{U}_a \subset \mathcal{U}$ which can associate to it (so capturing geographic and access control constraints). So a user is either associated to the AP a or not associated to it. We let \mathcal{B} denote the set of LTE base stations, \mathcal{U}_b the set of users located in the cell operated by $b \in \mathcal{B}$ and \mathcal{B}_u the set of base stations with which user u can connect.

B. LTE throughput

Let \mathcal{I} denote the set of available LTE sub-channels and let $\zeta_{b,u}^i$ denote fraction of time sub-channel i of BS b is used by user u . Let \mathcal{E}^i be the set of user-BS pairs for which transmissions interfere on sub-channel i (this defines a conflict graph). We consider joint sub channel allocation and time sharing so that at a given time interfering BSs do not transmit on the same sub-channel [10]. Hence, the total allocation for each sub-channel must satisfy the following:

$$\sum_{(u,b) \in \mathcal{E}^i} \zeta_{b,u}^i \leq 1, \quad \forall i \in \mathcal{I} \quad (1)$$

When there is no frequency reuse, this constraint simplifies to $\sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \zeta_{b,u}^i \leq 1, \quad \forall i \in \mathcal{I}$. Letting ω^i denote the sub-channel bandwidth, then the achievable rate of user u is given

by:

$$r_u = \sum_{b \in \mathcal{B}} \sum_{i \in \mathcal{I}} \zeta_{b,u}^i \beta_1 \omega^i \log \left(1 + \frac{\gamma_{b,u}^i}{\beta_2} \right) \quad \forall u \in \mathcal{U} \quad (2)$$

with $\gamma_{b,u}^i = \frac{p_b^i e_{b,u}^i}{\sigma_n^2}$ denoting the SNR on sub-channel i of BS b where $p_b^i, e_{b,u}^i, \sigma_n^2$ denote the BS power on channel i , channel gain, and noise power at the receiver respectively. Parameters β_1 and β_2 reflect the LTE bandwidth and SINR implementation efficiencies [11]. Hence the rate region of the LTE network is the following

$$\mathcal{R}_{lte} = \left\{ \mathbf{r} : r_u = \sum_{b \in \mathcal{B}} \sum_{i \in \mathcal{I}} \zeta_{b,u}^i \beta_1 \omega^i \log \left(1 + \frac{\gamma_{b,u}^i}{\beta_2} \right), \right. \\ \left. r \leq r_u \leq \bar{r}, 0 \leq \zeta_{b,u}^i \leq 1, \right. \\ \left. \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \zeta_{b,u}^i \leq 1, \forall i \in \mathcal{I} \right\} \quad (3)$$

where $\mathbf{r} \in \mathbb{R}^U$ is the vector formed by stacking the user throughputs $r_u, u \in \mathcal{U}$ and we have also added the constraint that each user has maximum and minimum rates \bar{r} and r respectively. Note that r may be 0 and \bar{r} may be ∞ . It can be seen that the LTE rate region is convex (a polytope in fact, since the constraints are linear) in the time-frequency sharing factors $\zeta_{b,u}^i$.

We note that this model can be extended to include dynamic power allocation per sub-channel, for example using a similar approach to [10] and will still be convex.

C. 802.11 Operation

We consider a setup with one 802.11 AP. UEs that connect to the IEEE 802.11 AP contend for the medium based on their Contention Window (CW) size. To transmit a packet, a UE will initialise a backoff counter with a random number uniformly distributed in the $[0, CW - 1]$ interval. The UE decrements its value every time slot, and transmits when the counter reaches zero. On the first attempt, the contention window is initialised with the CW_{min} . Upon each failure, size of the contention window is doubled. The maximum value of CW however is kept at CW_{max} . After a successful transmission of a frame, the CW size is reset to CW_{min} . Although the standard defines a set of recommended medium access parameters, it also allows the AP to periodically adjust and distribute these values to the associated UEs [12].

D. 802.11 throughput

We consider a $|\mathcal{U}_a|$ number of UEs contending for the 802.11 resources. We further assume that all UEs are in saturation conditions *i.e.* always have packets to send. Let $L_u, T_{succ,u}, T_{coll}$ and σ denote, respectively, the payload in bits of a successful transmission by UE u , the mean duration of a successful transmission by UE u , the mean duration of a collision and the idle slot duration. To simplify notation we confine consideration to client uplink transmissions. Extension to include AP downlink transmissions is straightforward, simply requiring appropriate book-keeping of which fraction of

AP transmissions is directed to each client. For simplicity we also assume that $T_{succ,u}$ is the same for all users transmitting in the WLAN and let $T_{succ} = T_{succ,u}$.

We let \mathcal{X}_u be a random variable which equals to 1 if UE u attempts to transmit **at slot time** t . Hence $\tau_u := \text{Prob}(\mathcal{X}_u = 1)$ defines the probability of UE u accessing the 802.11 medium at **randomly chosen slot time**. This can be achieved in practice by setting $CW_{min,u} = CW_{max,u}$ in which case $\tau_u = 2/(CW_{min,u} + 1)$ [13]. We let \mathcal{Y}_u be a random variable that equals to 1 if UE u transmits successfully, and 0 otherwise. We let \mathcal{Z} equal 0 when one or more users transmits and 1 otherwise. Note that $P_{succ,i} = E[\mathcal{Y}_u] = \tau_u \prod_{u' \in \mathcal{U}_a \setminus \{u\}} (1 - \tau_{u'})$, $P_{idle} = E[\mathcal{Z}] = \prod_{u \in \mathcal{U}_a} (1 - \tau_u)$ respectively define successful transmission probability of UE u and idle transmission probability. Defining the $P_{succ} = \sum_{i=1}^{|\mathcal{U}_a|} P_{succ,i}$ as the probability of a successful transmission, we have that $P_{coll} = 1 - P_{idle} - P_{succ}$.

The throughput of user u therefore is defined as the ratio of the expected number of successfully transmitted bits (**over a slot time**) to the mean duration of a slot time.

$$s_u = \frac{P_{succ,u} L_u}{P_{idle} \sigma + \sum_{i=1}^{|\mathcal{U}_a|} P_{succ,i} T_{succ,i} + P_{coll} T_{coll}} \quad (4)$$

It will prove useful to work in terms of the quantity $\phi = \frac{\tau}{1-\tau}$ rather than τ . With this transformation we have that

$$s_u = \frac{\phi_u L_u / T_{coll}}{\frac{\sigma}{T_{succ}} - 1 + (\frac{T_{succ}}{T_{coll}} - 1) \sum_{i \in \mathcal{U}_a} \phi_i + \prod_{k \in \mathcal{U}_a} (1 + \phi_k)} \quad (5)$$

We can re-write the throughput expression by

$$s_u = \frac{\phi_u L_u / T_{coll}}{\Phi(\phi)} \quad (6)$$

with

$$\Phi(\phi) = \frac{\sigma}{T_{succ}} - 1 + (\frac{T_{succ}}{T_{coll}} - 1) \sum_{i \in \mathcal{U}_a} \phi_i + \prod_{k \in \mathcal{U}_a} (1 + \phi_k) \quad (7)$$

Also we will use the normalised throughput $\rho_u = \frac{s_u}{c_u}$, where $c_u = L_u / T_{succ}$.

1) *802.11 Rate Region*: With the above change of variables the rate region of the 802.11 network is,

$$\mathcal{R}_{wifi} = \left\{ \mathbf{s} : s_u = \rho_u c_u, \quad \rho_u \leq \frac{\phi}{\Phi(\phi)} \frac{T_{succ}}{T_{coll}}, \quad \forall u \in \mathcal{U} \right\} \quad (8)$$

where $\mathbf{s} \in \mathbb{R}^U$ is the vector obtained by stacking the user throughputs s_u , $u \in \mathcal{U}$. Letting $\tilde{s}_u = \log s_u$, $\tilde{\rho}_u = \log \rho_u$, $\tilde{c}_u = \log c_u$, and $\tilde{\phi}_u = \log \phi_u$, the log rate region of the 802.11 network can be written as

$$\tilde{\mathcal{R}}_{wifi} = \left\{ \tilde{\mathbf{s}} : \tilde{s}_u = \tilde{\rho}_u + \tilde{c}_u, \quad \tilde{\rho}_u \leq \tilde{\phi}_u + \log \frac{T_{succ}}{T_{coll}} - \log \left(\frac{\sigma}{T_{succ}} - 1 + (\frac{T_{succ}}{T_{coll}} - 1) \sum_{i \in \mathcal{U}_a} e^{\tilde{\phi}_i} + \prod_{k \in \mathcal{U}_a} (1 + e^{\tilde{\phi}_k}) \right), \quad \forall u \in \mathcal{U} \right\} \quad (9)$$

E. Convexity Properties of Network Rate Region

The network rate region is the joint LTE/802.11 rate region $\mathcal{R}_{lte} \times \mathcal{R}_{wifi}$. The LTE rate region \mathcal{R}_{lte} is convex and so the convexity/non-convexity of the network rate region depends on the convexity/non-convexity of the 802.11 component of the network rate region \mathcal{R}_{wifi} . When we allow the UEs to adjust

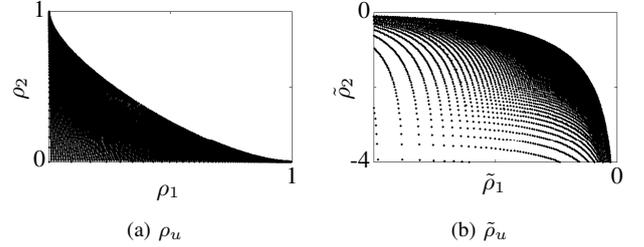


Fig. 2: Example of 802.11 normalised rate region where probability of accessing the channel (τ) varies where $T_{succ} = T_{coll} = 10\mu\text{sec}$, $\sigma = 1\mu\text{sec}$

the fraction of time (τ) that they access the 802.11 channel, the rate region becomes non-convex, as illustrated in Figure 2(a). **However we can show that by taking logs, the 802.11 rate region becomes log-convex.**

Lemma 1: $\log \Phi(e^{\tilde{\phi}})$ is convex in $\tilde{\phi}$.

Proof: Re-writing $\log \Phi(e^{\tilde{\phi}})$ by

$$\Phi(e^{\tilde{\phi}}) = \log \left(a + \sum_{i \in \mathcal{U}_a} e^{\tilde{\phi}_i + b} + \sum_{k=1}^{|\mathcal{U}_a|} e^{\sum_{j \in \mathcal{P}_k(\mathcal{U}_a)} \tilde{\phi}_j} \right)$$

with $a = \frac{\sigma}{T_{succ}} - 1$, $b = \frac{T_{succ}}{T_{coll}}$, and $\mathcal{P}_k(\mathcal{U}_a)$ denoting subsets of \mathcal{U}_a with cardinality k . It follows that the log sum of exponential is convex [14, p. 72], and Lemma 1 holds. ■

Lemma 2: 802.11 log Rate region defined by $\tilde{\mathcal{R}}_{wifi}$, is convex.

Proof: By Lemma 1 we have that the constraint governing the rate region is log-convex. By Eq (9) it follows that the rate region is also log-convex. ■

III. PROPORTIONAL FAIR RATE ALLOCATION

A. Utility Fair Optimisation

We are now in a position to consider the main focus of this work, namely finding a proportional fair rate allocation for the joint LTE/802.11 network. This is the solution to the following utility optimisation P ,

$$\max_{\tilde{\mathbf{s}}, \mathbf{r}} \sum_{u \in \mathcal{U}} \log(e^{\tilde{s}_u} + r_u) \quad (10)$$

$$\text{s.t. } \tilde{\mathbf{s}} \in \tilde{\mathcal{R}}_{wifi}, \quad \mathbf{r} \in \mathcal{R}_{lte} \quad (11)$$

It would prove useful to re-write the initial optimisation problem P in the following (epigraph) form

$$\max_{\tilde{\mathbf{s}}, \mathbf{r}} \sum_{u \in \mathcal{U}} t_u \quad (12)$$

$$\text{s.t. } \tilde{\mathbf{s}} \in \tilde{\mathcal{R}}_{wifi}, \quad \mathbf{r} \in \mathcal{R}_{lte} \quad (13)$$

$$t_u \leq \log(e^{\tilde{s}_u} + r_u), \quad \forall u \in \mathcal{U} \quad (14)$$

with $\mathbf{t} \in \mathbb{R}^{|\mathcal{U}|}$ denoting a vector with elements $t_u, \forall u \in \mathcal{U}$. It can easily be seen that the solution to the optimisation problem above is optimal if and only if it is also optimal for its initial form, P .

Observe that the LHS of the epigraph constraint (14) is a convex function. Hence despite the convexity of the LTE and 802.11 rate regions, the utility fair optimisation problem P is non-convex. However, as we will see the optimisation possesses sufficient convex structure to allow near-optimal solutions to be found in an efficient manner. We proceed by first considering methods for approximating a non-convex set by a convex subset.

B. Approximate Optimisation Problem

Let $\mathbf{x} := [\zeta^T \mathbf{r}^T \mathbf{s}^T \tilde{\rho}^T \tilde{\phi}^T \mathbf{t}^T]^T \in \mathbb{R}^n$ be the vector obtained by stacking the optimisation variables with $\zeta \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{B}| \times U}$, $\tilde{\rho}, \tilde{\phi} \in \mathbb{R}^{|\mathcal{U}|}$ denoting, respectively, the vectors with elements $\zeta_{b,u}^i, \tilde{\rho}_u, \tilde{\phi}_u$, for $b \in \mathcal{B}, i \in \mathcal{I}, u \in \mathcal{U}$. **Re-writing the rate region constraints in the standard form and by transforming the maximisation to minimisation, optimisation problem P is cast as follows**

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad (15)$$

$$s.t. \quad h^{(i)}(\mathbf{x}) - g^{(i)}(\mathbf{x}) \leq 0, \quad i = 1, 2, \dots, l \quad (16)$$

with $f: \mathbb{R}^n \rightarrow \mathbb{R}, h^{(i)}: \mathbb{R}^n \rightarrow \mathbb{R}, g^{(i)}: \mathbb{R}^n \rightarrow \mathbb{R}$ being convex continuously differentiable functions, expressions of which are given in the Appendix. **We define a convex approximation of the concave functions $-g^{(i)}(\cdot)$ by $-\hat{g}^{(i)}(\mathbf{x}; \bar{\mathbf{x}})$.**

$$-\hat{g}^{(i)}(\mathbf{x}; \bar{\mathbf{x}}) = -g^{(i)}(\bar{\mathbf{x}}) - \partial g_{\mathbf{x}}^{(i)}(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) \quad (17)$$

Then the approximate optimisation problem $P_{\bar{\mathbf{x}}}$ is given by

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad (18)$$

$$s.t. \quad h^{(i)}(\mathbf{x}) - \hat{g}^{(i)}(\mathbf{x}; \bar{\mathbf{x}}) \leq 0, \quad i = 1, 2, \dots, l \quad (19)$$

By forming convex subsets as described above, we obtain a convex optimisation the solution of which is feasible for non-convex optimisation P but will, in general, be sub-optimal. The Slater condition is satisfied and so strong duality holds. We can therefore solve this convex optimisation in a distributed manner using a primal-dual subgradient approach. Letting $\Lambda := [\lambda^{(1)}, \dots, \lambda^{(l)}]^T$ denote the set of multipliers associated with the rate region constraints 1 to l , the Lagrangian for optimisation problem $P_{\bar{\mathbf{x}}}$ is

$$L(\mathbf{x}, \Lambda; \bar{\mathbf{x}}) = f(\mathbf{x}) + \sum_{i=1}^l \lambda^{(i)} \left(h^{(i)}(\mathbf{x}) - \hat{g}^{(i)}(\mathbf{x}; \bar{\mathbf{x}}) \right)$$

The standard primal-dual subgradient approach in Algorithm 1 can then be used, for example, to find a solution to optimisation $P_{\bar{\mathbf{x}}}$.

To implement the updates in Algorithm 1, only local communication between an AP/a BS and its users is required. Furthermore, the multiplier associated with the sub-channel constraint need to be exchanged between neighbour BSs. The multiplier associated with the sub-channel constraint can also be exchanged by the UEs

Algorithm 1 Distributed primal-dual algorithm

Initialise: $t = 0, \mathbf{x}(0), \Lambda(0)$, step size $\alpha > 0$

do

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{x}(t) - \alpha \partial_{\mathbf{x}} L(\mathbf{x}(t), \Lambda(t); \bar{\mathbf{x}}) \\ \Lambda(t+1) &= [\Lambda(t) + \alpha \partial_{\Lambda} L(\mathbf{x}(t), \Lambda(t); \bar{\mathbf{x}})]^+ \\ t &\leftarrow t+1 \end{aligned}$$

loop

beaconing their used channels to the neighbour UEs. $\tilde{\phi}_{a,u}$ can be updated at the APs, by exchanging it between the AP and its UEs. The choice of convex subset $\bar{\mathbf{x}}$ can also be updated by exchanging $\tilde{\rho}_u$ between the AP and its users.

C. Adaptation of Convex Subsets

To find the best approximation to the optimisation problem P , we can iteratively update the choice of the convex subsets based on the current solution to the convex optimisation problem ($P_{\bar{\mathbf{x}}}$). Having obtained the solution (\mathbf{x}) to the convex optimisation for the current choice of approximation points ($\bar{\mathbf{x}}$), one natural approach is to use the component $\tilde{\rho}_u$ of this solution as the values for the parameter $\tilde{\rho}_u$.

1) Convergence: In more detail, let $C := \{\mathbf{x} \in B : h^{(i)}(\mathbf{x}) - g^{(i)}(\mathbf{x}) \leq 0, i = 1, \dots, l\}$ with $B \subset \mathbb{R}^n$ non-empty, convex and compact (closed and bounded). As we will see, set B is needed for technical reasons, to ensure that C is compact, but B can otherwise be chosen arbitrarily and can be viewed as augmenting the set of convex constraints $h^{(i)}$. Of course, we assume that set C is non-empty. Now consider the iterative update

$$\mathbf{x}_{k+1} \in D_{\mathbf{x}_k} f(\mathbf{x}), \quad k = 1, \dots \quad (20)$$

with $\mathbf{x}_1 \in C$,

$$D_{\mathbf{x}_k} f(\mathbf{x}) = \begin{cases} \arg \min_{\mathbf{x} \in C_{\mathbf{x}_k}} f(\mathbf{x}) & \text{if } f(\mathbf{x}_k) = f_{\mathbf{x}_k}^* \\ \{\mathbf{x} \in C_{\mathbf{x}_k} : f(\mathbf{x}) < f(\mathbf{x}_k)\} & \text{otherwise} \end{cases} \quad (21)$$

and $f_{\mathbf{x}_k}^* := \min_{\mathbf{x} \in C_{\mathbf{x}_k}} f(\mathbf{x}), C_{\mathbf{x}_k} := \{\mathbf{x} \in B : h^{(i)}(\mathbf{x}) - \hat{g}^{(i)}(\mathbf{x}; \mathbf{x}_k) \leq 0\}$.

Observe that the solution to convex optimisation $P_{\mathbf{x}_k}$ lies in set $D_{\mathbf{x}_k} f(\mathbf{x})$ and so can be used in update (20). However, set $D_{\mathbf{x}_k} f(\mathbf{x})$ defines a wider class of updates, including suitable approximate solutions to $P_{\mathbf{x}_k}$ that may be easier/faster to find, and generalises the concave-convex approach of [15].

We have the following convergence result.

Theorem 1 (Local Convergence): Iterative update (20) converges to a stationary point of non-convex optimisation P .

Proof: **This can be proved using the facts that (i) update (20) is a descent update, and so will converge (provided the variables stay bounded) and (ii) the Fritz John conditions satisfied at the optimum of the convex problem, ensure that the optimum also satisfies the Fritz John conditions for the non-convex problem and so update (20) converges to a stationary point of optimisation problem P .** ■

By adding a stochastic search component to update (20) we can strengthen this to obtain a global convergence result.

Theorem 2 (Global Optimisation): Consider iterative update

$$\mathbf{x}_{k+1} \in \begin{cases} \mathbf{X}_{k+1} & \text{w.p. } \epsilon \\ D_{\mathbf{x}_k} f(\mathbf{x}) & \text{otherwise} \end{cases} \quad (22)$$

$$y_{k+1} = \arg \min \{f(y_k), f(\mathbf{x}_{k+1})\}$$

with $\mathbf{x}_1 \in C$ and where $\{\mathbf{X}_k\}$ is a sequence of random variables taking values in C satisfying $Prob(\mathbf{X}_{k+1} \in \Sigma | \mathbf{X}_k \in \Upsilon) \geq \eta \mu(\Upsilon \cap \Sigma)$ for any $\Upsilon, \Sigma \subset C$ where $\mu(\Upsilon \cap \Sigma)$ denotes the volume of their intersection and parameter $\eta > 0$. Then, y_k converges to the optimum of optimisation P with probability one.

Proof: By applying the covering and sub-covering sets of the set C we can prove that for any initial condition and every $\xi > 0$, for every optimum x^* of optimisation P a ball $B_\xi(x^*)$ containing x^* is visited with positive probability and the stated result now follows. ■

Update (22) is no longer purely a descent update, but rather with probability ϵ , an update is made which may lead to the objective f increasing, allowing escape from unfavourable stationary points (saddle points *etc*) and from local minima. The requirement on random variable \mathbf{X}_{k+1} can be satisfied by, for example, selecting \mathbf{X}_{k+1} uniformly at random within a ball about the current point \mathbf{x}_k .

Although the proof of Theorem 2 could be used to upper bound the convergence, this bound would be very loose and so while Theorem 2 provides some comfort as to the ability of update (22) to find a global optimum, evaluation of the convergence rate really needs to be carried out via numerical experiments.

2) *Descent Step:* Descent step $D_{\mathbf{x}_k} f(\mathbf{x})$ can be realised in many ways. One is by use of a truncated version of the primal-dual subgradient approach:

Algorithm 2 $D_{\mathbf{x}_k} f(\mathbf{x})$: Truncated Primal-Dual Updates

```

do
     $\mathbf{z}(t) = \mathbf{x}(t)$ 
     $\mathbf{z}(t+1) = \mathbf{z}(t+1) - \alpha \partial_{\mathbf{z}} L(\mathbf{z}(t), \Lambda(t); \mathbf{x}_k)$ 
     $\Lambda^{(i)}(t+1) = \left[ \Lambda^{(i)}(t) + \alpha \partial_{\Lambda} L(\mathbf{z}(t), \Lambda(t); \mathbf{x}_k) \right]^+$ 
     $t = t + 1$ 
until  $f(\mathbf{z}(t+1)) \leq f(\mathbf{x}_k)$  and  $\mathbf{z}(t+1) \in C_{\mathbf{x}_k}$ 
 $\mathbf{x}_{k+1} = \mathbf{z}(t+1)$ 

```

While this primal-dual approach might not lead to a feasible point at every step, the requirement that \mathbf{x}_{k+1} be feasible can be relaxed to the requirement that (i) \mathbf{x}_{k+1} is bounded and (ii) \mathbf{x}_{k+1} is feasible for sufficiently large k . Theorem 1 then continues to hold, with only minor changes to the proof.

IV. NUMERICAL RESULTS

A. Motivating Example: Revisited

Armed with our proportional fair problem formulation and non-convex optimisation tools we now revisit the example in

Fig 1 in more detail. The network consists of one LTE base-station, one 802.11 AP and two UEs.

We assume that the AP uses 802.11n [16] with the settings detailed in Table II. **Note that we consider frame aggregation via Aggregated-Mac Packet Data Units (A-MPDUs) frame structures in 802.11n.** For simplicity we assume that only two LTE sub-channels are available, denoted sub-band one and sub-band two. The LTE system parameters are detailed in Table II. **Figure 3 illustrates the convergence**

TABLE II: Network parameters.

802.11 Parameters	Physical Rate, $(c_{a,1}, c_{a,2})$ Physical Rate, $(c_{b,1}, c_{b,2})$ Basic Rate (R_{basic}) Slot Time (σ) PLCP Preamble PLCP Header MAC Header Payload size (L) Number of MPDU subframes for u_1 (K_1) Number of MPDU subframes for u_2 (K_2) FSC Delimiter Padding ACK RTS CTS	(1, 54) Mbps (5, 5) Mbps 1 Mbps 9 μ s 16 μ s 48 bits @ $c_{a,u}$ 192 bits @ $c_{a,u}$ 1500 bytes @ $c_{a,u}$ 1 54 32 bits @ $c_{a,u}$ 32 bits @ $c_{a,u}$ 32 bits @ $c_{a,u}$ 112 bits @ R_{basic} 160 bits @ R_{basic} 112 bits @ R_{basic}
LTE Parameters	Duplex Mode System Bandwidth Carrier Frequency Number of Antennas at BS Number of Antennas at UE Guard Band Overhead DMRS Overhead Random Access Overhead Cyclic Prefix Overhead Bandwidth Efficiency (β_1) SNR Efficiency (β_2) SNR $(\gamma_{1,1}^{(i)}, \gamma_{1,2}^{(i)})$, $i = 1, 2$	FDD 2×10 MHz 2 GHz 1 1 10% 14.3% 0.625% 6.66% 71% 1 (4.9, 4.9) dB
Optimisation:	Convex Subset 1 ($\bar{\mathbf{x}}_1$) with $\{(\bar{\rho}_1, \bar{\rho}_2)\}$ Convex Subset 2 ($\bar{\mathbf{x}}_2$) with $\{(\bar{\rho}_1, \bar{\rho}_2)\}$ Convex Subset 3 ($\bar{\mathbf{x}}_3$) with $\{(\bar{\rho}_1, \bar{\rho}_2)\}$ α (step size) ϵ	$\{(-1, 0)\}$ $\{(-2, -2)\}$ $\{(0, -1)\}$ 0.1 0.2

as well as the solutions found, using update (20) for various choices of initial convex subsets, detailed in Table II. For comparison, when u_1 and u_2 use only the 802.11 WLAN their data rates are, respectively 0.46Mbps and 24.9Mbps while when the UEs use only the LTE the BS allocates a rate of 5Mbps to each UE. Splitting the traffic for each UE equally over the 802.11 and LTE networks would yield rates of 5.46Mbps and 29.9Mbps.

For comparison, Figure 4 shows the convergence of both updates (20) and (22). Observe that in this example the extra complexity of update (22) does not yield a better optimum, and indeed we also observed this in the other examples presented below.

B. 802.11 Offload

Our problem formulation can be readily extended to provide a principled approach for offload of data traffic from LTE to 802.11. In an offload setting we would like to preferentially use the 802.11 network where possible. We can capture this requirement by augmenting the proportional fair utility function $\sum_{u \in \mathcal{U}} \log(s_u + r_u)$ with a cost related to LTE usage.

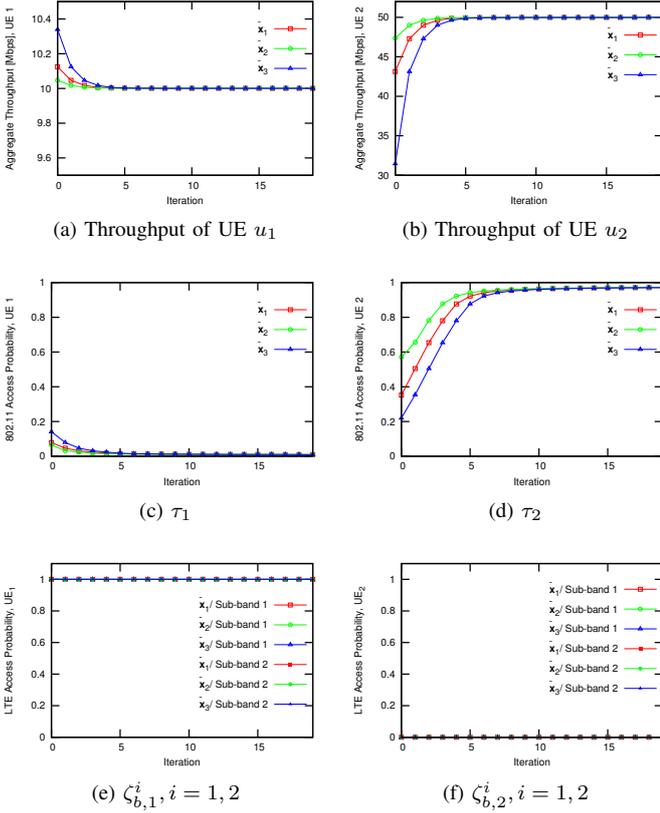


Fig. 3: Convergence of update (20) for various choices of initial convex subset, detailed in Table II.

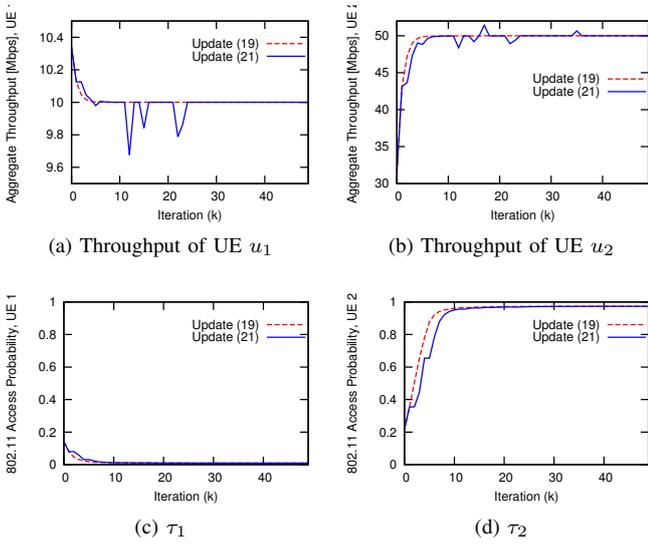


Fig. 4: Convergence of updates (20) and (22) when \mathbf{X}_{k+1} is drawn uniformly random within the ball of radius 0.025.

For example, the cost function

$$H(\zeta_{b,u}^i) = \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \sum_{i \in \mathcal{I}} \frac{\zeta_{b,u}^i}{\beta_1 \omega^i \log(1 + \frac{\gamma_{b,u}^i}{\beta_2})} \quad (23)$$

TABLE III: 802.11 offload example data rates.

	PHY Rates [Mbps]		Technology Rates [Mbps]		
	BS	AP1	LTE only	802.11 only	Optimised Multi-RAT
u_1	4	54	2	16.5	16.5
u_2	4	54	2	16.5	16.5

associates a cost with the LTE airtime used by each UE (the airtime being inversely proportional to the data rate).

Consider a simple network setup with one BS, one AP and two UEs. Suppose the 802.11 physical rate is 54Mbps for both users and the LTE rates are 4Mbps (corresponding to an SNR of 0.75dB), see Table III. Other 802.11 LTE parameters are detailed in Table II.

Using update (20) to solve the optimisation problem, Table ?? summarises the solution found. The UEs share the 802.11 AP and do not send traffic via the LTE network (the LTE rates $r_{u_1} = 0 = r_{u_2}$). The throughput of each UE is 16.5Mbps, which is higher than the data rate of 2Mbps provided by the LTE network. Figure 5(a) plots how the fraction $\zeta_{1,b,u}^i$ of LTE airtime and 802.11 access probabilities, (τ_u) used by each UE change as the LTE data rate is increased, while Figure 5(b) shows the aggregate throughput (LTE plus 802.11) of each UE. It can be seen that for LTE data rates less than 4Mbps the LTE network is not used and data is fully offloaded to the 802.11 network. However, as the LTE data rate increases the LTE network is increasingly used to enhance the user throughputs.

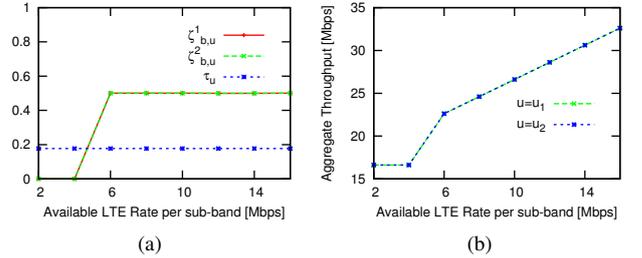


Fig. 5: Illustrating 802.11 offload as LTE data rate is varied, (a): LTE airtime, (b): Aggregate LTE plus 802.11 UE throughput.

We can readily extend consideration to include situations where UEs have a specified traffic load by adding additional constraints to the optimisation problem. For example, suppose that both UEs have a traffic load of 12Mbps video, the LTE available rate for each UE is 12Mbps (i.e. the total LTE capacity is 24 Mbps, enough to support the traffic load) and the 802.11 physical rate for UE u_1 is fixed at 6Mbps while that of UE u_2 is varied. Using Eq (23) to account for the cost of using the LTE connection we expect the traffic to be offloaded to the 802.11 network as long as its capacity is sufficient to meet the load. **This is illustrated in Fig 6.** Observe also that the 802.11 usage by UE u_1 falls to zero when the 802.11 rate is 12Mbps. This allows UE u_2 to make full use of the 802.11 network to meet its traffic demand, without incurring the overhead of collisions. Once the 802.11 rate of the UE u_2 increases further, the extra capacity is then used by UE u_1 .

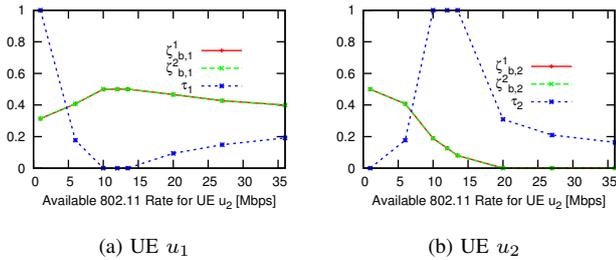


Fig. 6: Illustrating 802.11 offload as 802.11 rate is varied.

C. LTE Multihoming

We next consider an example with two BSs (b_1 and b_2), one AP, and 4 UEs (u_1 , u_2 , u_3 , and u_4). The physical rates available on the 802.11 and LTE links are summarised in Table IV. We also consider the conventional user assignments to LTE BSs *i.e.* each UE associates with the LTE BS that provides the UE with the maximum received power. The achievable LTE rates when using this maximum received power assignment are also summarised in Table IV. Table V shows the proportional

TABLE IV: LTE multihoming example: data rates.

	PHY Rates [Mbps]			Technology Rates [Mbps]				
	BS b_1	BS b_2	AP	LTE b_1 only	LTE b_2 only	LTE (Maximum Rx Power)	802.11 only	Optimised Multi-RAT
u_1	26	25	54	6.5	6.25	8.27	7.51	14.6
u_2	10	25	27	2.5	6.25	2.4	3.75	11.94
u_3	5	29	54	1.25	7.25	3.56	7.51	14.02
u_4	11	10	13.5	2.75	2.5	2.29	1.88	5.25

fair LTE sub-band allocations and 802.11 access probabilities obtained using update (20). From Table IV it can be seen that a relatively large physical rate of 54 Mbps is available to UEs u_1 and u_3 and so it can be seen from Table V that in the proportional fair allocation these UEs make use of the 802.11 network but not the LTE network, so freeing up capacity in the LTE network for UEs u_2 and u_4 . UEs u_2 and u_4 make use of the LTE rather than the 802.11 network (so reducing collisions and increasing the 802.11 capacity available to UEs u_1 and u_3), and share the LTE network evenly between them. It can be seen from the right-hand column of Table IV that all users benefit from this use of multi-homing.

V. SUMMARY AND CONCLUSION

We consider proportional fair rate allocation in a heterogeneous network with LTE and 802.11 RATs which supports multipath and multi-homed operation (simultaneous connection of a user device to multiple LTE BSs and one 802.11

TABLE V: Rate allocation for LTE multi-homing example.

Parameter	u_1	u_2	u_3	u_4
$\zeta_{1,u}^1$	0	0	0	0.47
$\zeta_{1,u}^2$	0	0	0	0.47
$\zeta_{2,u}^1$	0	0.47	0.05	0
$\zeta_{2,u}^2$	0	0.47	0.05	0
r_u	0	11.82	1.65	5.17
ρ_u	0.46	0.01	0.39	0.01
τ_u	0.5	0.02	0.46	0.02
s_u	14.6	0.12	12.37	0.08

APs). We show that the utility fair optimisation problem is non-convex but that a global optimum can be found by solving a sequence of convex optimisations in a distributed fashion. The result is a principled approach to offload from LTE to 802.11 and for exploiting LTE/802.11 path diversity to meet user traffic demands.

REFERENCES

- [1] 3GPP TR 36.819 V11.1.0 (2011-12) Technical Specification Group Radio Access Network; 3rd Generation Partnership Project; Coordinated multi-point operation for LTE physical layer aspects (Release 11), 2011.
- [2] D. Giustiniano, E. G. Llauro, J. M. Pozo, A. L. Toledo, and P. Rodriguez, "Clubadsl: When your neighbors are your friends," in *Computers and Communications, 2009. ISCC 2009. IEEE Symposium on*, pp. 25–29, IEEE, 2009.
- [3] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses," Internet Draft draft-ietf-mptcp-rfc6824bis-04, IETF, Mar. 2015.
- [4] L. Wang and G.-S. Kuo, "Mathematical modeling for network selection in heterogeneous wireless networks—a tutorial," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 1, pp. 271–292, 2013.
- [5] I. Blau, G. Wunder, I. Karla, and R. Sigle, "Decentralized utility maximization in heterogeneous multicell scenarios with interference limited and orthogonal air interfaces," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, p. 2, 2009.
- [6] M. Ismail and W. Zhuang, "A distributed multi-service resource allocation algorithm in heterogeneous wireless access medium," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 2, pp. 425–432, 2012.
- [7] D. Kumar, E. Altman, and J. M. Kelif, "User-network association in a wlan-umts hybrid cell: Global & individual optimality," *CoRR*, vol. abs/cs/0608025, 2006.
- [8] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, "An overview of load balancing in hetnets: Old myths and open problems," *arXiv preprint arXiv:1307.7779*, 2013.
- [9] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "Rat selection games in hetnets," in *INFOCOM, 2013 Proceedings IEEE*, pp. 998–1006, IEEE, 2013.
- [10] J. Huang, V. Subramanian, R. Berry, and R. Agrawal, "Scheduling and resource allocation in ofdma wireless systems," *book chapter, submitted*, 2009.
- [11] P. Mogensen, W. Na, I. Z. Kovács, F. Frederiksen, A. Pokhariyal, K. I. Pedersen, T. Kolding, K. Hugl, and M. Kuusela, "Lte capacity compared to the shannon bound," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, pp. 1234–1238, IEEE, 2007.
- [12] I. L. S. Committee *et al.*, "Ieee 802.11-wireless lan medium access control (mac) and physical layer (phy) specifications," in *IEEE*, June, 2007.
- [13] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 3, pp. 535–547, 2000.
- [14] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [15] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.
- [16] "802.11n-2009 - IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput," *IEEE Std*, 2009.
- [17] P. Billingsley, *Probability and measure*. John Wiley & Sons, 2008.

APPENDIX

A. Optimisation problem

With the obvious abuse of notation we add subscripts to the functions $h^{(i)}$ and $g^{(i)}$ to streamline the presentation since there is no risk of confusion. Details of the objective and

constraints functions are as follows:

$$\begin{aligned}
f(\mathbf{x}) &= - \sum_{u \in \mathcal{U}} t_u \\
h_u^{(1)}(\mathbf{x}) &= t_u \\
h_u^{(2)}(\mathbf{x}) &= \tilde{\rho}_u - \tilde{\phi}_u - \log \frac{T_{succ}}{T_{coll}} + \log \Phi(\tilde{\phi}) \\
h_i^{(3)}(\mathbf{x}) &= \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \zeta_{b,u}^i - 1 \\
h_u^{(4)}(\mathbf{x}) &= r_u - \sum_{b \in \mathcal{B}} \sum_{i \in \mathcal{I}} \zeta_{b,u}^i \beta_1 \omega^i \log \left(1 + \frac{\gamma_{b,u}^i}{\beta_2} \right) \\
g_u^{(1)}(\mathbf{x}) &= \log (e^{\tilde{\rho}_u} c_u + r_u) \\
g_u^{(2)}(\mathbf{x}) &= 0, \quad g_i^{(3)}(\mathbf{x}) = 0, \quad g_u^{(4)}(\mathbf{x}) = 0
\end{aligned}$$

$$\begin{aligned}
L(\mathbf{x}, \Lambda; \bar{\mathbf{x}}) &= - \sum_{u \in \mathcal{U}} t_u \\
&+ \sum_{u \in \mathcal{U}} \lambda_u^1 (t_u - \log (e^{\tilde{\rho}_u} (1 + \tilde{\rho}_u - \bar{\rho}_u) c_u + r_u)) \\
&+ \sum_{u \in \mathcal{U}} \lambda_u^2 \left(\tilde{\rho}_u - \tilde{\phi}_u - \log \frac{T_{succ}}{T_{coll}} + \log \Phi(\tilde{\phi}) \right) \\
&+ \sum_{i \in \mathcal{I}} \lambda_i^3 \left(\sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} \zeta_{b,u}^i - 1 \right) \\
&+ \sum_{u \in \mathcal{U}} \lambda_u^4 \left(r_u - \sum_{b \in \mathcal{B}} \sum_{i \in \mathcal{I}} \zeta_{b,u}^i \beta_1 \omega^i \log \left(1 + \frac{\gamma_{b,u}^i}{\beta_2} \right) \right)
\end{aligned} \tag{24}$$

Having obtained the Lagrangian $L(\mathbf{x}, \Lambda; \bar{\mathbf{x}})$ to the approximate convex problem $P_{\bar{\mathbf{x}}}$, we have

$$\begin{aligned}
\partial_{t_u} L &= -1 + \lambda_u^1 \\
\partial_{\tilde{\rho}_u} L &= -\lambda_u^1 \frac{e^{\tilde{\rho}_u} c_u}{(1 + \tilde{\rho}_u - \bar{\rho}_u) e^{\tilde{\rho}_u} c_u + r_u} + \lambda_u^2 \\
\partial_{\tilde{\phi}_u} L &= -\lambda_u^2 + \sum_{u \in \mathcal{U}} \lambda_u^2 \frac{\partial_{\tilde{\phi}_u} \Phi(\tilde{\phi})}{\Phi(\tilde{\phi})} \\
\partial_{\zeta_{b,u}^i} L &= -\lambda_u^{(3)} \beta_1 \omega^i \log \left(1 + \frac{\gamma_{b,u}^i}{\beta_2} \right) + \lambda_i^{(4)}
\end{aligned} \tag{25}$$

with

$$\partial_{\tilde{\phi}_u} \Phi(\tilde{\phi}) = \left(\frac{T_{succ}}{T_{coll}} - 1 \right) e^{\tilde{\phi}_u} + e^{\tilde{\phi}_u} \prod_{v \in \mathcal{U}_a \setminus \{u\}} (1 + e^{\tilde{\phi}_v})$$

To implement the distributed primal-dual Algorithm 1, the above subgradients need to be calculated at each AP/BS.