Max Weight Scheduling with Base Station Running and Switching Costs

Haritha K, Chandramani Singh
Department of Electronic Systems Engineering
Indian Institute of Science
Bangalore 560012, India
Email: {haritha, chandra}iisc.ac.in

***Abstract*—We consider a cellular downlink network of N base stations(BSs) serving M users through time-varying channels. The service rate provided by the channel is observed only when the corresponding BS is switched on. In this set-up, we consider two types of costs: $(1)$the cost of maintaining BS in on state, $(2)$the cost of switching the state of BS(from on to off or from off to on). In this paper, we propose algorithms that achieve cost arbitrarily close to the optimum cost under system stability. Since traditional drift-plus-penalty approach does not suffice to prove this, we propose our proof in two steps. In the first step, we show the BS running cost arbitrarily close to the optimum under queue stability. In the second step, we show the average number of BS switch-overs is of the order $O(1/U)$. We illustrate the performance of the proposed algorithm via simulations.**

I. Introduction

The exponential growth of networking devices, such as smart phones, tablets, laptops has resulted in dramatic increase in network data traffic. According to Ericsson Mobilty report [1], it is expected that the number of mobile users will be 7.2 billion by the year 2024.

Network densification is one of the promising solution to meet data traffic [2], [3]. Network densification means adding more cell sites to increase the amount of available capacity. Though densification increases spectral efficiency, as the number of deployed base stations (BSs) increases, the their power consumption also increases. According to GeSi's SMARTer2030 report [4] the amount of carbon footprint emitted by mobile sector is more than 50% of total emissions from Information and Communication Technology (ICT) infrastructure. BSs contribute to more than half the energy consumption of cellular networks [5], [6].

A prudent approach to ensure high QoS in dense cellular networks while keeping the energy profile low is to opportunistically operate BSs and allocate rates to users [7], [8]. If we keep all the BSs ON all the time, irrespective of the traffic, we can offer maximum QoS to users, e.g., users will experience least latency. But, this would incur excessive energy cost which includes energy consumed on air-conditioning, transceivers and signal processing. We may thus want to turn OFF subsets of BSs whenever they are underutilized, i.e.., their loads are small, and turn them ON when they accumulate enough packets. However, turning ON/OFF a BS causes signaling overhead (control signaling to users, signaling over backhaul to the BS controller etc.) [9]. Further, the associated state-migration processing may also incur certain latency. Finally, the BSs that are ON should allocate rates to the associated users keeping into account the instantaneous queue length and channel states (or, rates), e.g., they should prefer users with larger backlogs and higher rates. We thus need BS activation and user scheduling algorithms that balance the BS running and switching costs and users latency.

The first author is supported by Visvesvaraya PhD Scheme and the second author by INSPIRE Faculty Research Grant (DSTO-1363).

We address a BS switching and user scheduling problem for downlink in a multi-cell cellular network. These controls are executed on a slot basis. We assume that while a BS is ON it consumes unit power irrespective of the number of users it serves. We further assume that when we switch ON/OFF a BS we incur unit cost; this may be the wear and tear cost associated with BS operations or may represent the latency associated with BS switching. We propose an algorithm to determine which BSs to switch ON/OFF and which users to serve in each slot to get a desirable tradeoff between BS running and switching costs and users' latency.

The rest of this paper is organized as follows. We present the related work and summary of our contribution in Sections I-A and I-B respectively. We describe system model in Section II. In Section III, we propose and analyze drift plus penalty based throughput optimal scheduling algorithms that achieve BS running and switching costs arbitrarily close to the optimal cost. Further, in the absence channel distribution information the algorithm learns using explore-exploit technique. In IV we propose algorithms that learn channel distribution in all time slots. We present the numerical results in Section V. We conclude with a mention of future directions in Section VI.

A. Related Work

Resource allocation algorithms to maintain QoS while economizing on operating costs and overheads have garnered significant attention in the context of cellular and other networks (e.g., data centers). Prasanna et al. and Li et al. [9], [10] consider cellular downlink with costs associated with acquiring channel sate information and with actual transmissions. They propose algorithms to determine when channel state information needs to be acquired before scheduling transmissions. Neely et al. [12] propose a generic framework to allow two-stage decisions at the decision epochs. The first stage of the decision is executed in an unknown environment. This lets the environment being revealed partially. This information is used when executing the second stage. For instance, we may have to activate a subset of BSs without knowing the channel states. Once a subset of BSs is ON, we can observe the corresponding channels to judiciously allocate rates to the associated users. The costs here depend on the instantaneous network states. All these algorithms are based on Max-weight or Drift-plus-penalty frameworks.

Subhashini et al. [11] consider cellular networks with unknown packet arrival and channel statistics and both BS running and switching costs. They propose algorithms that minimize these costs while ensuring stability. In particular, they propose learning the arrival and channel statistics to determine optimal BS activation and using Max-weight for rate allocation. Our problem does not fit in Neely et al. [12]'s framework as the switching costs depend also on the previous BS configuration. Our approach also does not require explicitly learning the packet arrival statistics as in [11].

Several works in recent past have considered BS-user switchover delays rather than switching costs. These works assume that BSs are ON all the time (i.e., ignore BS running costs) and assume that whenever a BS switches from one user to other, a certain switchover time is needed before starting the service. Celik et al. [13] propose a Variable Frame-Based Max Weight (VFMW) algorithm that computes a schedule and also a frame duration for which this schedule should be used, both using queue lengths information. When the queue lengths are large, the frame lengths also become large, reducing the switching overhead. The authors show throughput optimality of the proposed algorithm. Hsieh et al. [14] considered the same problem and proposed a delay optimal queue length biased max weight scheduling algorithm named Q-BMW. Vineeth [15] also proposed a throughput optimal one step lookahead policy based on an approximate solution of a Markov decision process formulation of the problem.

As we mention above, these works do not consider BS running costs as we do. Moreover, since user switchings also incur latencies in these works, their capacity regions are smaller than those achieved in absence of switching delays.

Wang et al. [16] studied scheduling in a data center via dynamic configuration of optical switches. They considered switch reconfiguration delays wherein data center service stalls for a certain duration every time switches are reconfigured. They derived a class of adaptive max-weight based algorithms that are provably throughput optimal. Wang et al. [17] extended this study to scheduling in a network of geographically distributed data centers. They considered switch running and reconfiguration costs and also reconfiguration delays. Here also, the authors proposed a class of throughput optimal algorithms that achieve optimum cost.

Notice that the data center links do not exhibit the uncertainty that is inherent to cellular channels. Consequently, all the decisions there are made with complete state information.

B. our contribution

- We propose a drift-plus-penalty based joint BS switching and user scheduling algorithm. It is a two stage algorithm wherein in the first stage we switch ON/OFF BSs based on channel statistics and queue lengths and learn the channel states. In the second stage we schedule users also accounting for the channel state. Unlike earlier works (e.g., [11]), our approach does not rely on explicitly learning the packet arrival statistics at various BSs. In particular, we do not solve an optimization problem to get a static BS configuration policy to be used in each slot as in [11].
- Our algorithm yields BS switching cost of the order $O(\frac{1}{U})$, BS running cost within $O(\frac{1}{\gamma})$ of the optimum BS running cost and queue lengths of the order $O(\gamma U)$ where γ and U are parameters of the algorithm. We can thus drive the costs to their optimal values at the cost of increased queue lengths.
- When the channel statistics is unknown we progressively learn it using exploration slots at which all the BSs are turned ON. We propose an alternate where we update channel distribution estimates in each slot. Following this approach, we can potentially lean the statistics quickly. This approach gives comparable performance to the original approach for the scenarios that we have simulated. We thus see that we do not suffer from *Inspection paradox* mentioned in [12].

II. System Model

We consider a cellular network with N BSs serving M users through time varying channels. We consider downlink traffic. Each user m is associated with a subset of BSs N_m. Each BS receives downlink traffic for each of its associated users. BSs can occasionally be turned off to save energy. We refer to a set of active BSs as a BS configuration. Let $\mathcal{A}$ denote set of all possible BS configurations in the network. We assume that the network operates using a slotted time structure wherein BS configurations can change only at the slot boundaries.

Data Arrival Model: We assume that data packets for users arrive at the associated BSs according to a discrete time process, i.i.d across slots. Let $A_{mn}(t)$ be the number of packets for user m arriving at an associated BS n in slot t. We assume that $A(t) = (A_{mn}(t))_{m\in[M],n\in N_m}$, is an i.i.d. discrete time process. We further assume that $A_{mn}(t) \le \bar{A}$ for all $m \in [M], n \in N_m$. Let $A_{mn}(t)$ be the number of packets to the user m at time t. Let $\mathbb{E}[A_{mn}(t)] = \lambda_{mn}, m \in [M], n \in N_m$. As soon as arrivals arrive at the BS they are placed in corresponding queues. Let $Q_{mn}(t)$ represent the queue length of user m at time t.

Channel Model: Let $h(t)$ be the network-wide channel state at time t and $\mathcal{H}$ be the ensemble of all possible channel states.[1] We assume that $h(t), t > 0$ are i.i.d, with a probability distribution μ. Any BS configuration-channel state pair may result in several rate vectors, $r \in \mathcal{R}^{M\times N}$ depending on which BSs serve which users and also on communication constraints. Let $\mathcal{R}(a,h) \subset \mathcal{R}^{M\times N}$ be the set of possible rate vectors corresponding to BS configuration-channel state pair (a,h). For instance, $r_{mn} = 0$ for all $r \in \mathcal{R}(a,h)$ if BS $n \notin a$ or if user $m \notin N_m$. Let $r_{mn} \le \bar{R}$ for $r \in \mathcal{R}(a,h), a \in \mathcal{A}, h \in \mathcal{H}, m \in [M], n \in [N]$. Let $r(t) \in \mathcal{R}(a(t),h(t))$ be the rate

[1] A network channel state prescribes states of channels between all BS-user pairs.

vector chosen at time t. Then packet queues evolve as

$$Q_{mn}(t+1) = (Q_{mn}(t) - r_{mn}(t))^+ + A_{mn}(t). \quad (1)$$

A sequence of BS configurations and rate vectors $a(t) \in \mathcal{A}, \nabla(\sqcup) \in \mathcal{R}(\dashv(\sqcup), \langle(\sqcup), \sqcup \geq \infty$, is referred to as a schedule. A scheduling is sequence of mappings, one for each slot. At each time t, it maps the network's current and past states, $\{Q(l), a(l), h(l), r(l), l < t\}$ and arrivals and channel statistics to a couple $a(t), r(t)$.

Stability: The network is said to be stable if there exists a scheduling policy yielding a sequence $\{(a(t), h(t)), t \geq 1\}$, such that

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{P}\left[\sum_{mn} Q_{mn}(t) \leq \bar{Q} | a(0), Q(1)\right] > \rho \quad (2)$$

for some constants $\bar{Q}, \rho > 0$, for any initial condition $a(0), Q(1)$.

Capacity Region:: The capacity region region $\Lambda(\mu)$ is defined as the set of all arrival rate vectors for which there exists a scheduling algorithm that renders the system stable. An arrival rate vector $\lambda \in \Lambda(\mu)$ if and only if there exists distributions $\alpha(1, h) \in \mathcal{P}(\mathcal{R}(1, h)) \forall h \in \mathcal{H}$ such that, see [**?**], [11]

$$\lambda < \sum_{h\in\mathcal{H}} \mu_h \sum_{r\in R(1,h)} \alpha_r(1, h) r.$$

Throughput Optimal Algorithm: A scheduling algorithm is called throughput optimal if it stabilizes the network for any arrival rate vector λ such that $(1 + \epsilon)\lambda \in \Lambda(\mu)$ for some $\epsilon > 0$.[2]

BS Running Cost: We assume that when a BS is ON, it incurs unit cost irrespective of number of users it is serving. Formally, the BS running cost at time t is $\|a(t)\|_1$.

BS Switching Cost: We incur switching cost every time we change BS configuration, i.e., we switch ON or OFF a few BSs. The cost depends on the number of BSs that are switched ON or OFF. Formally, the BS switching cost at time t is $\|a(t) - a(t-1)\|_1$.

We define the net cost at time t to be a weighted sum of BS running and switching costs. We use constants $U, V > 0$ to specify relative weights of the two costs. Specifically, we define the net cost at time t as

$$C(t) = V\|a(t)\|_1 + U\|a(t-1) - a(t)\|_1.$$

A. Optimal Scheduling Problem

We aim to minimize the time averaged expected net cost while ensuring network stability. Mathematically, the optimal scheduling problem can be put as

$$\min \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[C(t)]$$
subject to (2).

Let $C(\lambda, \mu)$ denote the optimal values of the above optimization problem. Now we characterize above optimization problem. Markov static-split algorithm is characterized by

1) an irreducible, aperiodic stochastic matrix P over $\mathcal{A}$.
2) distributions $\alpha(a, h)$ over $\mathcal{R}(a, h)$, for all $a \in \mathcal{A}, h \in \mathcal{H}$

At any time t, given the previous BS configuration $a(t-1)$, the algorithm selects the new BS configuration according to P. Further, it selects the rate vector according to $\alpha(a, h)$. For any $\lambda \in \Lambda(\mu)$, the optimal cost across all the stabilizing Markov-static-split algorithms can be obtained via solving the following problem. [3]

$$C(\lambda, \mu) = \inf_{P,\alpha} U \sum_{a',a} \sigma_{a'}(P) P_{a'a} \|a' - a\|_1 + V \sum_{a\in\mathcal{A}} \sigma_a \|a\|_1 \quad (3)$$

$$\text{such that, } \alpha(a, h) \in \mathcal{P}_{\mathcal{R}(a,h)}$$
$$\lambda < \sum_{a\in\mathcal{A}} \sigma_a(P) \sum_{h\in\mathcal{H}} \mu_h \sum_{r\in R(a,h)} \alpha_r(a, h) r, \ \forall a \in \mathcal{A}, h \in \mathcal{H}.$$

The authors in [11] show that for any $\lambda \in \Lambda(\mu)$, $C(\lambda, \mu)$ equals the optimal value of this problem. They further show that the joint Markov-static-split algorithms that yield

1) BS switching cost arbitrarily close to zero.
2) BS running cost arbitrarily close to the optimal value of the following problem.

$$\bar{C}(\lambda, \mu) = V \inf_{\sigma,\beta} \sum_{a\in\mathcal{A}} \sigma_a \|a\|_1, \text{ such that} \quad (4)$$
$$\sigma \in \mathcal{P}_{|\mathcal{A}|},$$
$$\beta_r(a, h) \geq 0, \forall a \in \mathcal{A}, \forall h \in \mathcal{H} \forall r \in \mathcal{R}(a, h)$$
$$\sigma_a = \sum_{r\in R(a,h)} \beta_r(a, h), \forall a \in \mathcal{A}, \forall h \in \mathcal{H}$$
$$\lambda \leq \sum_{a\in\mathcal{A}, h\in\mathcal{H}, r\in\mathcal{R}(a,h)} \beta_r(a, h) \mu_h r$$

It is known that the class of Markov-static-split scheduling algorithms that yield cost arbitrarily close to $C(\lambda, \mu)$, see [11]. Let $\bar{C}(\lambda, \mu)$ denote the optimal value of this problem.

Remarks 2.1: 1) For any λ in the interior of $\Lambda(\mu)$, there exists a $\epsilon_g > 0$ such that $\lambda + \epsilon_g \in \Lambda$. Let σ, β be the solutions of the above optimization problem corresponding to $\lambda + \epsilon_g$ and μ. Then we can define distributions $\alpha(a, h)$ and an irreducible aperiodic stochastic matrix P for $\epsilon_s \in (0, 1)$ as follows.

$$\alpha_r(a, h) = \frac{\beta_r(a, h)}{\sigma_a}$$
$$P(\sigma, \epsilon_s) := \epsilon_s 1_{|\mathcal{A}|} \sigma + (1 - \epsilon_s) I_{|\mathcal{A}|}$$

These P and α constitute a Markov-static-split policy. By choosing small enough ϵ_g and ϵ_s, the cost of this can be made arbitrarily close to $\bar{C}(\lambda, \mu)$.
2) The constraints λ in the original optimization problem (3) can be written as

$$\lambda \leq \sum_a \sigma_a \sum_h \mu_h \sum_{r\in\mathcal{R}(a,h)} \alpha_r(a, h) r.$$

So this optimization problem is relaxed version of (3). Consequently $\bar{C}(\lambda, \mu) \leq C(\lambda, \mu)$.

[2] With slight abuse of notation, we write $\lambda + 1\epsilon = \lambda + \epsilon$.

[3] $\sigma(P)$ denotes unique stationary distribution associated with P.

III. Optimal Cost Scheduling Algorithm

We propose drift plus penalty based scheduling algorithms that are throughput optimal and also yield BS running and switching costs arbitrarily close to the optimal cost. More specifically, for any μ and $\lambda \in \Lambda(\mu)$, the proposed algorithms yield

1) BS running cost arbitrarily close to $\bar{C}(\lambda, \mu)$
2) BS switching cost arbitrarily close to 0

Unlike [11], the proposed algorithms need not know or learn arrival rate vector. However, these algorithms still need channel statistics, μ. In the following, we first consider scenario where we know μ a priori. We then move to scenario where we gradually learn μ by observing channel states over the course of the algorithm.

A. Known Channel Statistics

The proposed algorithm is presented in Algorithm 1. At each time t, scheduling consists of two steps

1) BS switching: We determine which BSs have to be turned ON/OFF in accordance with step (2). We do not see the complete channel state $h(t)$ but we see it partially depending on $a(t)$. However, this partial knowledge reveals $\mathcal{R}(a(t), h(t))$.
2) Rate allocation: We let the BSs serve users so as to obtain a rate vector $r(t)$ as in step (4).

Algorithm 1 Known Channel Statistics

1. **for** all $t > 0$ **do**
2. Given $Q(t), a(t-1)$, select BS configuration
$$a(t) \in \underset{a \in \mathcal{A}}{\operatorname{argmax}} \sum_h \mu_h \max_{r \in \mathcal{R}(a,h)} \left(\sum_{m,n} Q_{mn}(t) r_{mn} - V\|a\|_1 - U\|a - a(t-1)\|_1 \right)$$
3. Observe $h(t)|_{a(t)}$
4. Select rate vector
$$r(t) \in \underset{r \in \mathcal{R}(a(t),h(t))}{\operatorname{argmax}} \sum_{m,n} Q_{mn}(t) r_{mn}$$
5. **end for**

The following theorem formalizes optimality of Algorithm 1.

Theorem 3.1: For any arrival rate vector λ, such that $\lambda + \epsilon_g \in \Lambda(\mu)$ for some $\epsilon_g > 0$, Algorithm 1 is throughput optimal and achieves cost arbitrarily close to the optimal value as shown below.

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{m,n} \mathbb{E}[Q_{mn}(t)] \leq \frac{B + V\bar{C}(\lambda + \epsilon_g, \mu) + NU}{\epsilon_g},$$
$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[\|a(t)\|_1] \leq \frac{B}{V} + \frac{UN}{V} + \bar{C}(\lambda + \epsilon_g, \mu),$$
$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[\|a(t) - a(t-1)\|_1] \leq \frac{\bar{R}^2}{U} + \frac{MN\bar{A}\bar{R}}{U},$$

where $B = MN\frac{\bar{A}^2 + \bar{R}^2}{2}$.

Proof: This follows from the proof of Theorem 3.3 that asserts optimality of a similar Algorithm 2 for unknown channel statistics. By choosing V, U values appropriately Algorithm 1 achieves cost arbitrarily close to optimal cost. ■

Remarks 3.1: 1)The special cases where only BS running cost is considered and where only BS switching cost is considered, can be handled $U = 0$ and $V = 0$, respectively, in Theorem 3.1.
2) Let us set $V = \gamma U$. We can bring the number of BS switchings close to zero by increasing U. Similarly we can drive BS running cost close to optimal by increasing γ.

a) Discussion:: We now explain the intuition behind Theorem 3.3 (3). We determine BS switching decisions before observing the channel states. In particular we switch a BS ON, when the corresponding queue lengths become large enough. Similarly we switch a BS OFF when queue lengths become sufficiently small. Since the queue lengths vary gradually, BS ON/OFF events can be made apart choosing appropriate U.

Often BS switch incurs switching or reconfiguration delay see [10], [13]. So the network resources are wasted whenever switching occurs. Following similar reasoning above, we can make this wastage arbitrarily small by appropriately setting U. Further, we can continue the ideas here, with those in [11] to ensure network stability and to simultaneously optimize BS running cost for any $\lambda \in \Lambda(\mu)$.

In the following we illustrate two scenarios where the above reasoning does not hold. Hence these scenarios to be dealt differently.

1) Assume that at each time t, the channel state $h(t)$ is known before determining the BS configuration $a(t)$. Then the optimum $a(t)$ could also take into account $h(t)$. Since the channel state could arbitrarily change from one slot to next, the optimum BS configuration could also change. In this case it is challenging to derive algorithms that discourage frequent BS switching.
2) Given a channel state h and BS configuration a, we can obtain different states in $\mathcal{R}(a, h)$ depending on BSs' service discipline. For instance let the BSs use TDMA to serve the associated users. Then the rate vector in a slot will depend on which BS serves which user. Let us further assume that a switching cost is incurred whenever a BS moves from one associated user to other. Since optimal service would depend on instantaneous rates to various associated users, which can again be widely different across slots. It is challenging to obtain optimal switching cost, queue length trade-off in this scenario. Hence, BSs can switch across the associated users without accounting for their instantaneous rates. In this case, we can get a similar cost, queue length trade-off as presented in this paper. However, since we may not using the best available rates the capacity region will shrink.

B. Unknown Channel Statistics

Here, we do not know the channel statistics μ a priori. We rather learn it by observing the complete channel states a number of times over the course of the algorithm. The relative frequencies of various channel states until time t yield an estimate of μ at that time. Let $\mu(t)$ denote the estimate at time t. The complete algorithm (Algorithm 2) is presented in Unknown Channel Statistics.

To elaborate, we categorize the slots as "exploration" and "exploitation" slots based on whether $E(t) = 1$ or 0 (see steps $(4, 9)$) in Algorithm 2 and $\epsilon(t) = \mathbb{E}[E(t)] = \frac{2\log t}{t}$. Let $K(t)$ denote the total number of explorations until time t; $K(t) = \sum_{s=1}^{t} E(s)$. In an exploration slot t, we switch all the BSs ON (i.e., $a(t) = 1$) and observe the complete channel state $h(t)$. We then appropriately update the empirical distribution $\mu(t)$. In an exploitation slot t we use the current estimate $\mu(t)$ to perform BS switching. As in the known channel statistics case, we see the channel state $h(t)$ partially which is sufficient to reveal $\mathcal{R}(a(t), h(t))$. In all the slots the algorithm we use rate allocation exactly as in the case of known channel statistics.

Let $\hat{\mu}_k$ denote the empirical channel distribution after k explorations. Clearly, $\mu(t) = \hat{\mu}_{K(t)}$. Let us define the following events for every $t > 0, k > 0, \delta > 0$.

$$
\begin{aligned}
\varepsilon(t) &:= \Big\{K(t) \geq \frac{1}{2}\log^2 t\Big\},\\
\varepsilon_k(\delta) &:= \Big\{\|\hat{\mu}_k - \mu\| \leq \delta\Big\},\\
\varepsilon(t,\delta) &:= \Big\{\|\mu(l) - \mu\| \leq \delta, \forall l \geq t\Big\}.
\end{aligned}
$$

We can see that, $\varepsilon(t,\delta) \supseteq \varepsilon(t) \bigcap_{k \geq 1/2\log^2 t} \varepsilon_k(\delta)$. The following result from [18] ensures that the time to reach $\delta-$close to the true channel distribution has bounded second moment.

Lemma 3.1: For any $\delta > 0$, let $\Gamma(\delta) := \min\{t : \varepsilon(t,\delta) \text{ is true}\}$. Then $\mathbb{E}[\Gamma(\delta)^2] < \infty$.

Proof: Please see [18, Lemma C.4]. ■

The following theorem states that, $\forall t \geq \Gamma(\delta)^4$ the mean conditional Lyapanov drift is negative for large enough queue lengths. It also gives an upper bound on average number of active BSs $\forall t \geq \Gamma$.

Theorem 3.2: For any $\delta_1 > 0, \lambda$, such that $\lambda + \epsilon_g \in \Lambda(\mu)$ for some $\epsilon_g > 0$ there exists constant B_1, B_2 and a random time Γ such that $\mathbb{E}[\Gamma^2|Q(1)] < \infty$, and for any $t \geq \Gamma$ Algorithm 2 achieves,

$$
\mathbb{E}\left[\Delta(t)|Q(t),\Gamma,N(t-1)\right] \leq B_2 - \frac{\epsilon_g}{2}\sum_{m=1}^{M} Q_m(t).
$$

$$
\begin{aligned}
\mathbb{E}\Big[\|a(t)\|_1|Z(t)\Big] &\leq \frac{B_1}{V} + N\left(\frac{4\delta_1}{\mu_{\min} - \delta_1}\right)\\
&+ \frac{NU}{V}\left(1 + \frac{3\delta_1}{\mu_{\min} - \delta_1}\right) + \bar{C}(\lambda + \epsilon_g, \mu)
\end{aligned}
$$

where $B_1 = \frac{1}{2}MN(\bar{A}^2 + \bar{R}^2)$, $B_2 = B_1 + V\bar{C}(\lambda + \epsilon_g, \mu) + (4V + 2U)N\frac{\delta_1}{\mu_{\min} - \delta_1} + UN$.

Proof: Please see our technical report [19]. ■

Theorem 3.3: For any $\delta_1 > 0$ and λ, such that $\lambda + \epsilon_g \in \Lambda(\mu)$, under the Algorithm 2,

1) the system is stable,

[4] From now onwards we use Γ for $\Gamma(\delta)$.

Algorithm 2 Unknown Channel Statistics

1. Initialize $\mu(0) = \text{uniform}(\mathcal{H}), K(0) = 0$
2. **for** all $t > 0$ **do**
3. Generate $E(t) \sim \text{Bernolulli}(\epsilon(t))$
4. **if** $E(t) = 1$ **then**
5. Set $a(t) = 1$
6. Observe $h(t)$
7. Set $K(t) = K(t-1) + 1$
8. Set $\mu(t) = \dfrac{\mu(t-1)K(t-1) + \mathbb{1}_{\{h(t)=h\}}}{K(t)}$
9. **else**
10. Select
$$
a(t) \in \underset{a\in\mathcal{A}}{\operatorname{argmax}} \sum_h \mu_h(t) \max_{r\in\mathcal{R}(a,h)} \Bigg(\sum_{m,n} Q_{mn}(t) r_{mn} - V\|a\|_1 - U\|a - a(t-1)\|_1\Bigg)
$$
11. Observe $h(t)|_{a(t)}$
12. **end if**
13. Select
$$
r(t) \in \underset{r\in\mathcal{R}(a(t),h(t))}{\operatorname{argmax}} \sum_{m,n} Q_{mn}(t) r_{mn}
$$
14. **end for**

---

2) the time average BS running cost

$$
\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\|a(t)\|_1] \leq \frac{B_1}{V} + \frac{NU}{V} + \bar{C}(\lambda + \epsilon_g, \mu),
$$

3) The average number of migrations is $O\Big(\frac{1}{U}\Big)$

$$
\begin{aligned}
\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}&\mathbb{E}\left[\|a(t) - a(t-1)\|_1\right]\\
&\leq \frac{1}{U}[MN\bar{A}\bar{R} + \bar{R}^2],
\end{aligned}
$$

where $B_1 = \frac{1}{2}MN(\bar{A}^2 + \bar{R}^2)$.

Proof: Please see Appendix A. ■

IV. More Scheduling Algorithms

The algorithm Algorithm 2 proposed in previous section updates the empirical distribution on complete channel states $\mathcal{H}$, in exploration slots only. One can observe channel state partially $\mathcal{H}|_{a(t)}$ in exploitation slots also. In this section we propose two algorithms Algorithm 3, Algorithm 4 that make use this information in updating the channel state distribution.

Let $K_a(t)$ and $\nu_{a,h}(t)$ denote the total number of times BS configuration a is scheduled and empirical distribution on $\mathcal{H}|_a$ till time $t, \forall a \in \mathcal{H}$. Observe that Algorithm 3 updates K_a and $\nu_{a,h}$ in exploitation slots also. In algorithm Algorithm 4 there are no exploration slots at all. Since there are no explorations, with $K_a(0) = 0, \forall a \in \mathcal{A}$ one cannot guarantee convergence of distribution $\nu_{a,h}, \forall a \in \mathcal{A}$. So we are initializing the counters to one, i.e., $K_a(0) = 1, \forall a \in \mathcal{A}$.

Through simulations in Section V we demonstrate the cost optimality above algorithms.

Algorithm 3 Unknown Channel Statistics

Initialize $\nu_{a,h} = \text{uniform}(\mathcal{H}|_a), K_a(0) = 0, \forall a \in \mathcal{A}$
for all $t > 0$ **do**
    Generate $E(t) \sim$ Bernolulli$(\epsilon(t))$
    if $E(t) = 1$ **then**
        Set $a(t) = 1$
    else
        Select

$$a(t) \in \underset{a \in \mathcal{A}}{\text{argmax}} \sum_{h \in \mathcal{H}|_a} \nu_{a,h}(t) \max_{r \in \mathcal{R}(a,h)} \Big[\sum_m Q_{mn}(t) r_{mn} - V\|a\|_1 - U\|a - a(t-1)\|_1 \Big]$$

    end if
    Observe $h(t) \in \mathcal{H}|_{a(t)}$
    Set $K_{a(t)}(t) = K_{a(t)}(t-1) + 1$
    Set $\nu_a(t) = \dfrac{K_{a(t)}(t-1)\nu_a(t-1) + \mathbb{1}_{h=h(t)}}{K_{a(t)}(t)}$
    Select

$$r(t) \in \underset{r \in \mathcal{R}(a(t),h(t))}{\text{argmax}} \sum_{m,n} Q_{mn}(t) r_{mn}$$

end for

Algorithm 4 Unknown Channel Statistics

Initialize $\nu_{a,h} = \text{uniform}(\mathcal{H}|_a), K_a(0) = 1, \forall a \in \mathcal{A}$
for all $t > 0$ **do**
    Select

$$a(t) \in \underset{a \in \mathcal{A}}{\text{argmax}} \sum_{h \in \mathcal{H}|_a} \nu_{a,h}(t) \max_{r \in \mathcal{R}(a,h)} \Big[\sum_m Q_{mn}(t) r_{mn} - V\|a\|_1 - U\|a - a(t-1)\|_1 \Big]$$

    Observe $h(t) \in \mathcal{H}|_{a(t)}$
    Set $K_{a(t)}(t) = K_{a(t)}(t-1) + 1$
    Set $\nu_a(t) = \dfrac{K_{a(t)}(t-1)\nu_a(t-1) + \mathbb{1}_{h=h(t)}}{K_{a(t)}(t)}$
    Select

$$r(t) \in \underset{r \in \mathcal{R}(a(t),h(t))}{\text{argmax}} \sum_{m} Q_{mn}(t) r_{mn}$$

end for

V. Simulations

In this section, we provide simulation results for the optimization algorithms we proposed. We consider the following setup of a wireless network with two BSs with four users. Each BS is associated with two users, i.e., BS1 is associated with users $1, 2$ and BS2 is associated with users $3, 4$. The Cellular downlink is ON-OFF fading channel. The channel states are $\mathcal{H} = \{(0,0,0,0), (1,0,1,0), (0,1,0,1), (1,1,1,1)\}$

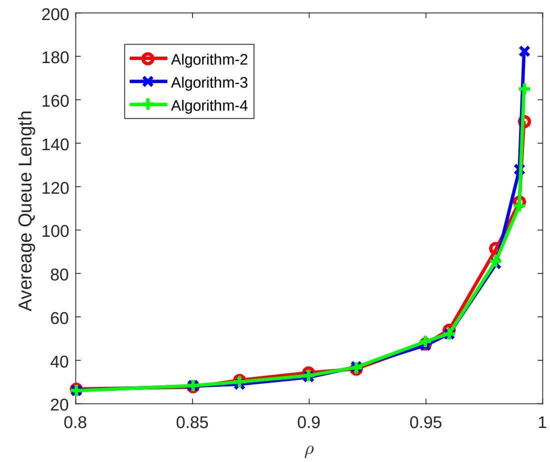

Fig. 1. Average queue length vs load intensity.

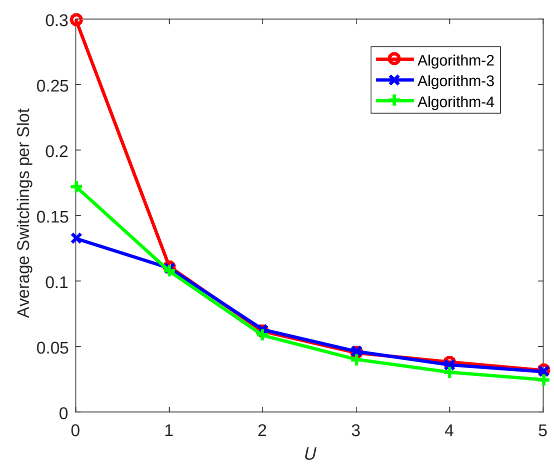

Fig. 2. Average BS switchings per slot vs U.

and are equiprobable. We assume Bernoulli arrivals on each queue. The capacity region of this network include,

$$\Big\{ \lambda : (\lambda_m n) = \rho(0.25, 0.5, 0.25, 0.5), 0 \le \rho < 1 \Big\},$$

where ρ is referred as load intensity.

In Figure.1 we demonstrate throughput optimality of the algorithms. For this we consider the above network with parameter values $U = 1, V = 5$ and plot queue length vs load intensity.

For remaining simulation results we fix arrival rate as $\lambda = (0.25, 0.25, 0.25, 0.25)$. In Figure.2 we demonstrate how average number of BS switchings vary with U varies for fixed V. We observe that as U is increases average number of BS switchings per slot decreases.

In Figure. 3 we plot total network cost for all the algorithms Algorithm 2, Algorithm 3 and Algorithm 4. In [12] Neely et al. claim that using the information from exploitation slots may degrade the performance (*Inspection Paradox*), we observe the algorithms Algorithm 3, Algorithm 4 that use information from exploitation slots achieve same performance as algorithm

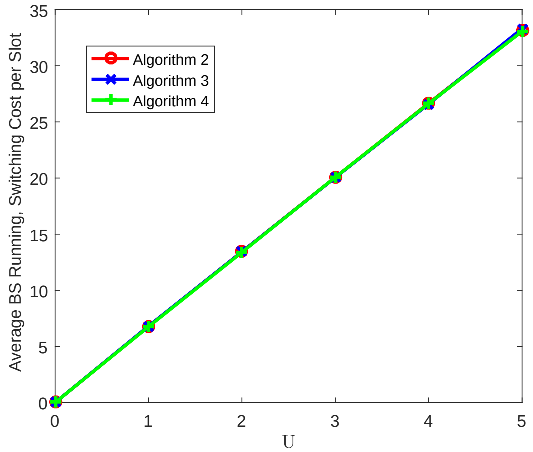

Fig. 3. Average BS running and swithching Cost per slot vs U.

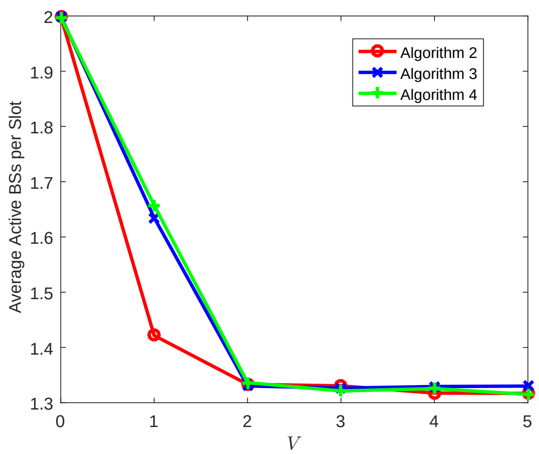

Fig. 4. Average active BSs per slot vs V.

Algorithm 2. In Figure.4 we demonstrate how the BS running cost varies with V for a fixed U. We observe that as V value increases we see that average number of active BSs is close to the optimal value.

VI. Conclusion

We proposed drift-plus-penalty based joint BS switching and user scheduling algorithms. We first considered a scenario with known channel statistics. We showed that the proposed algorithm (1) can yield close to optimal BS running and switching cost while ensuring network stability (3.1). We then moved to unknown channel statistics. We proposed an algorithm (2) that gradually learns channel distribution in the so called exploration slots, and uses this statistics for resource management in all the slots. This algorithm delivers performance similar to 1 (3.1). Finally, we presented an algorithm (3) that updates channel distribution estimates in each slot.

Our future work entails developing similar algorithms for scenarios where reconfiguration costs or delays are incurred also when BSs switch users. We would also like to consider networks with non-persistent users.

Appendix A
Proof of Theorem 3.3

Part 1: This follows from Theorem 3.2 and [18, Lemma C.3].

Part 2: Given Γ , the average BS running cost till time Γ can be upper bounded by

$$\mathbb{E}\Big[\sum_{t=1}^{\Gamma-1}\|a(t)\|_1|Q(t),\Gamma,a(t-1)\Big] \leq \Gamma N. \tag{5}$$

Using (5) and BS running cost from Theorem 3.2 we can write,

$$\mathbb{E}\left[\sum_{t=0}^{\Gamma-1}\|a(t)\|_1+\sum_{t=\Gamma}^{T}\|a(t)\|_1|Q(t),\Gamma,a(t-1)\right]\leq \Gamma N \\ +(T-\Gamma)\Bigg(\frac{B_1}{V}+N\left(\frac{4\delta_1}{\mu_{\min}-\delta_1}\right) \\ +\frac{NU}{V}\left(1+\frac{2\delta_1}{\mu_{\min}-\delta_1}\right)+\bar{C}(\lambda+\epsilon_g,\mu)\Bigg)$$

Now taking expectation with respect to $Q(t),\Gamma,a(t-1)$, and taking time average,

$$\frac{1}{T}\mathbb{E}\left[\sum_{t=0}^{T}\|a(t)\|_1\right]\leq\frac{1}{T}\mathbb{E}[\Gamma]N \\ +\frac{1}{T}(T-\mathbb{E}[\Gamma])\Bigg(\frac{B_1}{V}+N\left(\frac{4\delta_1}{\mu_{\min}-\delta_1}\right) \\ +\frac{NU}{V}\left(1+\frac{2\delta_1}{\mu_{\min}-\delta_1}\right)+\bar{C}(\lambda+\epsilon_g,\mu)\Bigg).$$

$$\lim_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{t=0}^{T}\|a(t)\|_1\right]\leq\frac{B_1}{V}+\bar{C}(\lambda+\epsilon_g,\mu) \\ +N\left(\frac{4\delta_1}{\mu_{\min}-\delta_1}\right)+\frac{NU}{V}\left(1+\frac{2\delta_1}{\mu_{\min}-\delta_1}\right).$$

Now by appropriately choosing U,V and δ_1 values average BS running cost is arbitrarily close to optimal cost $\bar{C}(\lambda+\epsilon_g)$.

Part 3:

Now consider another configuration at every t, $\tilde{a}(t)=a(t-1)$, [5], which do not cause BS switchings at all.

$$\sum_h \mu_h(t)\max_{r\in\mathcal{R}(a(t-1),h)}\sum_{m,n}Q_{mn}(t)r_{mn}(t)-V\|a(t-1)\|_1$$

Let

$$\tilde{r}_h(t)\in\operatorname*{argmax}_{r\in\mathcal{R}(a(t-1),h(t))}Q_{mn}(t)r_{mn}.$$

$$\sum_h\mu_h(t)\sum_{m,n}Q_{mn}(t)\bar{r}_{h,mn}(t)-V\|a(t)\|_1-U\|a(t)-a(t-1)\|_1$$

[5] $\tilde{a}(t)$ is the different from the one which we used in the proof of Theorem 3.2

$$\geq \sum_h \mu_h(t) \sum_{m,n} Q_{mn}(t)\tilde{r}_{h,mn}(t) - V\|a(t-1)\|_1$$

$$\begin{aligned}
&U\|a(t)-a(t-1)\|_1 \\
&\quad \leq \sum_h \mu_h(t) \sum_{m,n} Q_{mn}(t)(\bar{r}_{h,mn}(t) - \tilde{r}_{h,mn}(t)) \\
&\qquad + \sum_h \mu_h(t) V(\|a(t-1)\|_1 - \|a(t)\|_1)
\end{aligned}$$

Telescopic sum $t = \{1, \ldots T\}$.

$$\begin{aligned}
&\sum_{t=1}^{T} \|a(t)-a(t-1)\|_1 \\
&\leq \sum_{t=1}^{T} \sum_h \mu_h(t) \sum_{m,n} Q_{mn}(t)(\bar{r}_{h,mn}(t) - \tilde{r}_{h,mn}(t)) \\
&\quad + \sum_{t=1}^{T} \sum_h \mu_h(t) V(\|a(t-1)\|_1 - \|a(t)\|_1) \\
&\leq \sum_h \sum_{m,n} \Big(\mu_h(T) Q_{mn}(T)\bar{r}_{h,mn}(T) - \mu_h(1) Q_{mn}(1)\tilde{r}_{h,mn}(1) \Big) \\
&\quad + \sum_{t=1}^{T-1} \mu_h(t) Q_{mn}(t)\bar{r}_{h,mn}(t) - \sum_{t=2}^{T} \mu_h(t) Q_{mn}(t)\tilde{r}_{h,mn}(t) \\
&\quad + V \sum_{t=1}^{T} \sum_h \left[\mu_h(t-1) - \mu_h(t)\right] \|a(t-1)\|_1 \\
&\leq \sum_h \sum_{m,n} \Big(\mu_h(T) Q_{mn}(T) r_{mn}(T) \\
&\quad + \sum_{t=1}^{T-1} \mu_h(t) Q_{mn}(t)\bar{r}_{h,mn}(t) - \sum_{t=2}^{T} \mu_h(t) Q_{mn}(t)\tilde{r}_{h,mn}(t) \Big)
\end{aligned}$$

The last inequality is due to

$$\sum_h \Big(\mu_h(t-1) - \mu_h(t) \Big) = 0$$

and at $t = 0$ we assume that all BSs are switched off.

$$\begin{aligned}
&\sum_{t=1}^{T} U\|a(t)-a(t-1)\|_1 \\
&\leq \sum_h \mu_h(T) \sum_{m,n} Q_{mn}(T)\bar{r}_{h,mn}(t) \\
&\quad + \sum_{t=1}^{T-1} \mu_h(t) Q_{mn}(t)\bar{r}_{h,mn}(t) - \sum_{t=2}^{T} \mu_h(t) Q_{mn}(t)\tilde{r}_{h,mn}(t) \\
&\leq TMN\bar{A}\bar{R} + \sum_{t=1}^{T-1} \sum_h \sum_{m,n} (\mu_h(t) \Big(Q_{mn}(t)\bar{r}_{h,mn}(t) \\
&\quad - \mu_h(t+1) Q_{mn}(t+1)\tilde{r}_{h,mn}(t+1) \Big) \\
&\leq TMN\bar{A}\bar{R} + \sum_{t=1}^{T-1} \sum_h \sum_{m,n} \max_{r \in \mathcal{R}(a(t),h)} r_{h,mn} \Big(\mu(t) Q_{mn}(t) \\
&\quad - \mu_h(t+1) Q_{mn}(t+1) \Big) \\
&\leq TMN\bar{A}\bar{R} + \sum_{t=1}^{T-1} \sum_h \sum_{m,n} \hat{r}_{h,mn}(t) \Big(\mu(t) Q_{mn}(t) \\
&\quad - \mu_h(t+1) Q_{mn}(t+1) \Big) \\
&\leq TMN\bar{A}\bar{C} + \sum_{t=1}^{T-1} \sum_h \sum_{m,n} \hat{r}_{h,mn}(t) (\mu(t) Q_{mn}(t) \\
&\quad - \mu_h(t+1)(Q_{mn}(t) - \bar{R})) \\
&\leq TMN\bar{A}\bar{R} + T\bar{R}^2 \\
&\quad + \sum_{t=1}^{T-1} \sum_h \sum_{m,n} Q_{mn}(t)\hat{r}_{h,mn}(t)(\mu_h(t) - \mu_h(t+1)) \\
&\leq TMN\bar{A}\bar{R} + T\bar{R}^2 + \bar{C}\bar{A} \sum_{t=1}^{T-1} \sum_h \sum_{m=1}^{M} t(\mu_h(t) - \mu_h(t+1))
\end{aligned} \tag{6}$$

Now consider $\mu_h(t+1)$, channel statistics are upgraded only during exploration times. If $t+1$ is not an exploration slot then $\mu_h(t) = \mu_h(t+1)$, last term in the above inequality is zero. Let $K_h(t)$ be the number of times channel state h is observed till time t and $K(t)$ be the total explorations observed till t. We assume $K(t) > 0$ for this whole discussion. If $E(t+1) = 1$ then $t+1$ is an exploration sample. Using these we can write,

$$\begin{aligned}
\mu_h(t+1) &= \frac{K_h(t) + \mathbb{1}_{h(t)=h}}{K(t) + E(t+1)} \\
&= \frac{\mu_h(t)K(t)}{K(t) + E(t+1)} + \frac{\mathbb{1}_{h(t)=h}}{K(t) + E(t+1)}
\end{aligned}$$

$$\begin{aligned}
&\mu_h(t) - \mu_h(t+1) \\
&\quad = \mu_h(t) - \frac{\mu_h(t)K(t)}{K(t) + E(t+1)} - \frac{\mathbb{1}_{h(t)=h}}{K(t) + E(t+1)} \\
&\quad \leq \mu_h(t) - \frac{\mu_h(t)K(t)}{K(t) + 1} - \frac{\mathbb{1}_{h(t)=h}}{K(t) + 1} \\
&\quad \leq \mu_h(t) - \frac{\mu_h(t)K(t)}{K(t) + 1} - \frac{\mathbb{1}_{h(t)=h}}{K(t) + 1} \\
&\quad = \frac{\mu_h(t)}{K(t) + 1} - \frac{\mathbb{1}_{h(t)=h}}{K(t) + 1} \leq \frac{\mu_h(t)}{K(t)}
\end{aligned}$$

Using this in (6)

$$\begin{aligned}
&\sum_{t=1}^{T} U\|a(t)-a(t-1)\|_1 \\
&\quad \leq MNT\bar{A}\bar{R} + T\bar{R}^2 + \bar{R}\bar{A} \sum_{t:E(t+1)=1} \sum_h \sum_{m,n} t \frac{\mu_h(t)}{K(t)} \\
&\quad \leq MNT\bar{A}\bar{R} + T\bar{R}^2 + \bar{R}\bar{A}MN \sum_{t:E(t+1)=1} t \frac{1}{K(t)} \\
&\quad \leq MNT\bar{A}\bar{R} + T\bar{R}^2 + \bar{R}\bar{A}MN \sum_{t=1}^{T-1} t \frac{E(t+1)}{\sum_{\tau=1}^{t} E(\tau)}
\end{aligned} \tag{7}$$

Now consider the last term in above inequality and take iterated expectation with respect to $E[1:t] = \{E(1), \dots \{E(t)\}$.

$$\sum_{t=1}^{T-1} t\mathbb{E}\left[\mathbb{E}\left[\frac{E(t+1)}{\sum_{\tau=1}^{t} E(\tau)} | E[1:t]\right]\right] = \sum_{t=1}^{T-1} t\frac{2\log t}{t}\mathbb{E}\left[\frac{1}{\sum_{\tau=1}^{t} E(\tau)}\right]$$

From the proof of 3.1 for $t \geq 3$ we have ,

$$\mathbb{P}\left[\sum_{\tau=1}^{t} E(\tau) \leq \frac{\log^2 t}{2}\right] \leq \exp\left(\frac{-1}{24}\log^2 t\right)$$

from this we can write

$$\mathbb{P}\left[\frac{1}{\sum_{\tau=1}^{t} E(\tau)} \leq \frac{2}{\log^2 t}\right] \geq 1 - \exp\left(\frac{-1}{24}\log^2 t\right).$$

From this we can write

$$\mathbb{E}\left[\frac{1}{\sum_{\tau=3}^{t} E(\tau)}\right] \leq \frac{2}{\log^2 t}\left(1 - \exp\left(\frac{-1\log^2 t}{32}\right)\right) + \exp\left(\frac{-1\log^2 t}{32}.\right)$$

Now consider the last term in (7), use above inequality,

$$\begin{aligned}
&\sum_{t=3}^{T-1} t\mathbb{E}\left[\mathbb{E}\left[\frac{E(t+1)}{\sum_{\tau=3}^{t} E(\tau)} | E[1:t]\right]\right] \\
&\leq \sum_{t=1}^{T-1} 2\log(t+1)\left(\frac{2}{\log^2 t}(1 - \exp(\frac{-1}{32}\log^2 t)) + \exp\left(\frac{-1}{32}\log^2 t\right)\right) \\
&\leq \sum_{t=1}^{T-1}\left(\frac{4\log(t+1)}{\log^2 t} + 2\log(t+1)\exp\left(\frac{-1}{32}\log^2 t\right)\right)
\end{aligned}$$

The above inequality is due to $E(t)$ are i.i.d. random variables and $\mathbb{E}[E(t)] = \frac{2\log t}{t}$. In first two slots $t = 1, 2$ at most $2N$ BS switching can happen and taking expectation of (7) gives,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\Big[U\|a(t) - a(t-1)\|_1\Big] \\
&\leq \frac{1}{U}(MN\bar{A}\bar{R} + \bar{R}^2 +) + \frac{2N}{TU} \\
&\quad + MN\frac{1}{TU}\bar{R}\bar{A}\sum_{t=1}^{T-1}\left(\frac{4\log(t+1)}{\log^2 t} + 2\log(t+1)\exp\left(\frac{-1}{32}\log^2 t\right)\right) + \frac{2N}{TU}.
\end{aligned}$$

Now taking $\lim_{T\to\infty}$ on both sides,

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\Big[U\|a(t) - a(t-1)\|_1\Big] \leq \frac{1}{U}(MN\bar{A}\bar{R} + \bar{R}^2).$$

The last inequality is due to $\lim_{T\to\infty}\left(\frac{4\log(t+1)}{\log^2 t} + 2\log(t+1)\exp\left(\frac{-1}{32}\log^2 t\right) + \frac{2N}{TU}\right) = 0$.

References

[1] "Ericsson mobility report," November 2018, https://www.ericsson.com/assets/local/mobility-report/documents/2018/ericsson-mobility-report-november-2018.pdf.

[2] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5g," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, February 2014.

[3] H. Zhang, S. Huang, C. Jiang, K. Long, V. C. M. Leung, and H. V. Poor, "Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1936–1947, Sep. 2017.

[4] "Smarter2030,ict solutions for 21st century challenges," April 2015, http://smarter2030.gesi.org/downloads/Full_report.pdf.

[5] J. Wu, Y. Zhang, M. Zukerman, and E. K. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 803–826, Secondquarter 2015.

[6] M. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *2009 IEEE International Conference on Communications Workshops*, June 2009, pp. 1–5.

[7] G. Wu, C. Yang, S. Li, and G. Y. Li, "Recent advances in energy-efficient networks and their application in 5g systems," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 145–151, April 2015.

[8] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 56–61, June 2011.

[9] P. Chaporkar and A. Proutiere, "Optimal joint probing and transmission strategy for maximizing throughput in wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1546–1555, October 2008.

[10] C. Li and M. J. Neely, "Energy-optimal scheduling with dynamic channel acquisition in wireless downlinks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 4, pp. 527–539, April 2010.

[11] S. Krishnasamy, P. T. Akhil, A. Arapostathis, S. Shakkottai, and R. Sundaresan, "Augmenting max-weight with explicit learning for wireless scheduling with switching costs," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, May 2017, pp. 1–9.

[12] M. J. Neely, S. T. Rager, and T. F. La Porta, "Max weight learning algorithms for scheduling in unknown environments," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1179–1191, May 2012.

[13] G. D. Celik and E. Modiano, "Scheduling in networks with time-varying channels and reconfiguration delay," *IEEE/ACM Transactions on Networking*, vol. 23, no. 1, pp. 99–113, Feb 2015.

[14] P. Hsieh, I. Hou, and X. Liu, "Delay-optimal scheduling for queueing systems with switching overhead," *CoRR*, vol. abs/1701.03831, 2017. [Online]. Available: http://arxiv.org/abs/1701.03831

[15] B. S. Vineeth, "Throughput optimal scheduling for wireless downlinks with reconfiguration delay," *National Conference on Communications*, no. 5, Jan 2018.

[16] C. Wang, T. Javidi, and G. Porter, "End-to-end scheduling for all-optical data centers," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, April 2015, pp. 406–414.

[17] C. Wang, J. Llorca, A. M. Tulino, and T. Javidi, "Dynamic cloud network control under reconfiguration delay and cost," *CoRR*, vol. abs/1802.06581, March 2018.

[18] K. Subhashini, "Online learning and decision-making from implicit feedback," Ph.D. dissertation, The University of Texas, Austin, 2017.

[19] K. Haritha and C. Singh, "Lightweight max weight scheduling algorithms for wireless networks," March 2019, https://indianinstituteofscience-my.sharepoint.com/:b:/g/personal/haritha_iisc_ac_in/ESJmX_QNqe5KvHg0-dRsjaABK13TdVIT8P7nyCYv787RtA?e=lX6oNA.