Exploiting Channel Memory for Multi-User Wireless Scheduling without Channel Measurement: Capacity Regions and Algorithms

Chih-ping Li, *Student Member, IEEE* and Michael J. Neely, *Senior Member, IEEE*

***Abstract*—We study the fundamental network capacity of a multi-user wireless downlink under two assumptions: (1) Channels are not explicitly measured and thus instantaneous states are unknown, (2) Channels are modeled as ON/OFF Markov chains. This is an important network model to explore because channel probing may be costly or infeasible in some contexts. In this case, we can use channel memory with ACK/NACK feedback from previous transmissions to improve network throughput. Computing in closed form the capacity region of this network is difficult because it involves solving a high dimension partially observed Markov decision problem. Instead, in this paper we construct an inner and outer bound on the capacity region, showing that the bound is tight when the number of users is large and the traffic is symmetric. For the case of heterogeneous traffic and any number of users, we propose a simple queue-dependent policy that can stabilize the network with any data rates strictly within the inner capacity bound. The stability analysis uses a novel frame-based Lyapunov drift argument. The outer-bound analysis uses stochastic coupling and state aggregation to bound the performance of a restless bandit problem using a related multi-armed bandit system. Our results are useful in cognitive radio networks, opportunistic scheduling with delayed/uncertain channel state information, and restless bandit problems.**

***Index Terms*—stochastic network optimization, Markovian channels, delayed channel state information (CSI), partially observable Markov decision process (POMDP), cognitive radio, restless bandit, opportunistic spectrum access, queueing theory, Lyapunov analysis.**

I. Introduction

Due to the increasing demand of cellular network services, in the past fifteen years efficient communication over a single-hop wireless downlink has been extensively studied. In this paper we study the fundamental network capacity of a time-slotted wireless downlink under the following assumptions: (1) Channels are never explicitly probed, and thus their instantaneous states are never known, (2) Channels are modeled as two-state ON/OFF Markov chains. This network model is important because, due to the energy and timing overhead, learning instantaneous channel states by probing may be costly or infeasible. Even if this is feasible (when channel coherence time is relatively large), the time consumed by channel probing cannot be re-used for data transmission, and transmitting data without probing may achieve higher throughput [2].[1] In addition, it has been shown that wireless channels can be adequately modeled as Markov chains [3], [4], especially in high-speed transmission regimes. Since each time slot comprises a short period of time, channel states are likely correlated across slots. In this case we shall exploit channel memory to improve network throughput.

Specifically, we consider a time-slotted wireless downlink where a base station serves N users through N (possibly different) *positively correlated* Markov ON/OFF channels. Channels are never probed so that their instantaneous states are unknown. In every slot, the base station selects at most one user to which it transmits a packet. We assume every packet transmission takes exactly one slot. Whether the transmission succeeds depends on the unknown state of the channel. At the end of a slot, an ACK/NACK is fed back from the served user to the base station. Since channels are either ON or OFF, this feedback reveals the channel state of the served user in the last slot and provides partial information of future states. Our goal is to characterize all achievable throughput vectors in this network, and to design simple throughput-achieving algorithms.

We define the *network capacity region* Λ as the closure of the set of all achievable throughput vectors. Computing the capacity region Λ in closed form is complicated. The Markov ON/OFF channels do not seem to have enough structure to characterize the capacity region Λ exactly. Using a brute-force approach, we may in principle compute Λ by locating all boundary points. Each boundary point can be solved by formulating an N-dimensional Markov decision process (MDP) [5] with system states defined as the probabilities, conditioning on the channel observation history, that channels are ON in a slot. These MDPs, however, are very complex to solve. One reason is the curse of dimensionality; the state space of the MDPs is countably infinite (shown later in Section II) and grows exponentially with N. This also

Chih-ping Li (web: http://www-scf.usc.edu/∼chihpinl) and Michael J. Neely (web: http://www-rcf.usc.edu/∼mjneely) are with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA.

This material is supported in part by one or more of the following: the DARPA IT-MANET program grant W911NF-07-0028, the NSF Career grant CCF-0747525, the Network Science Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory W911NF-09-2-0053.

This paper appears in part in [1].

[1]One quick example is to consider a time-slotted channel with state space $\{\mathsf{B}, \mathsf{G}\}$. Suppose channel states are i.i.d. over slots with stationary probabilities $\Pr[\mathsf{B}] = 0.2$ and $\Pr[\mathsf{G}] = 0.8$. At state B and G, at most 1 and 2 packets can be successfully delivered in a slot, respectively. Packet transmissions beyond the capacity will all fail and need retransmissions. Channel probing can be done on each slot, which consumes 0.2 fraction of a slot. Then the policy that always probes the channel yields throughput $0.8(2 \cdot 0.8 + 1 \cdot 0.2) = 1.44$, while the policy that never probes the channel and always sends packets at rate 2 packets/slot yields throughput $2 \cdot 0.8 = 1.6 > 1.44$.

hinders the use of linear programming to solve the optimal steady states for the MDPs (assuming they exist). To further illustrate the computational difficulty, let us consider a very simple case where all channels are independent and have the same transition probability matrix, and we seek to locate the boundary point in the direction $(1, 1, \dots, 1)$. Equivalently, the goal is to maximize the sum throughput over i.i.d. channels. It is shown in [7] that the optimal policy is to serve all channels in a round robin fashion $1 \to 2 \to \cdots N \to 1 \to \cdots$ where, on each channel, packets are continuously transmitted until a NACK is received. The resulting sum throughput μ_{sum} is easily written as

$$\mu_{\text{sum}} \triangleq \lim_{K\to\infty} \frac{\sum_{k=1}^{K}\sum_{n=1}^{N}(L_{kn}-1)}{\sum_{k=1}^{K}\sum_{n=1}^{N}L_{kn}},$$

where L_{kn} denotes the time interval the base station stays with channel n in the kth round (c.f. Section III-B and [6]). In this simple case, the value of μ_{sum} is still complex to compute when $N > 2$. The reason is that the interval process $\{L_{11}, L_{12}, \dots, L_{1N}, L_{21}, L_{22}, \dots\}$ forms a high-order Markov chain, since the distribution of each L_{kn} depends on the previous $(N-1)$ intervals $\{L_{(k-1)(n+1)}, \dots, L_{k(n-1)}\}$. Following the above discussions, it seems infeasible to characterize Λ in closed form.

The first contribution of this paper is that, instead of computing Λ exactly, we construct an outer and an inner bound on Λ. The outer bound comes from analyzing a fictitious channel model in which every scheduling policy yields higher throughput than it does in the real network. The inner bound is the achievable rate region of a special class of *randomized round robin policies* (introduced in Section IV-A). These policies are simple and take advantage of channel memory. In the case of symmetric channels (that is, channels are i.i.d.) and when the network serves a large number of users, we show that as data rates are more *balanced*, or in a geometric sense as the direction of the data rate vector in the Euclidean space is closer to the 45-degree angle, the inner bound converges geometrically fast to the outer bound, and the bounds are tight. Round robin policies are first shown in [6], [7] to achieve optimal sum throughput over Markov ON/OFF channels in some special cases. Here we capitalize these results to construct our inner capacity bound.

The inner capacity bound is indeed useful. First, the structure of the bound itself shows how channel memory improves throughput when there are multiple users in the network.[2] An intuition is that, as compared to treating channels as i.i.d. over slots, incorporating channel memory enlarges the control space of the network and thus the optimal performance can only be improved. The throughput gain due to channel memory will be made precise later in Lemma 4 (Section III-B); an example given in Section IV-D. Second, we show analytically that a large class of intuitively good heuristic policies achieve throughput that is at least as good as this bound, and hence the bound acts as a (non-trivial) performance guarantee. Finally, supporting throughput outside this bound may inevitably involve solving a much more complicated POMDP. Thus, for simplicity and practicality, we may regard the inner bound as an *operational* network capacity region.

In this paper we also derive a simple queue-dependent dynamic round robin policy that stabilizes the network whenever the arrival rate vector is interior to our inner bound. This policy has polynomial time complexity and is derived by a novel *variable-length frame-based Lyapunov analysis*, first used in [10] in a different context. This analysis is important because the inner bound is based on a mixture of many round robin policies acting on different subsets of channels, and an offline computation of the proper time average mixtures needed to achieve a given point in this complex inner bound would require solving $\Theta(2^N)$ unknowns in a linear system, which is impractical when N is large. The Lyapunov analysis overcomes this complexity difficulty with online queue-dependent decisions.

This paper applies to the emerging area of opportunistic spectrum access in cognitive radio networks (see [11] and references therein), where the channel occupancy of a primary user acts as a Markov ON/OFF channel to the secondary users. In previous work, [6]–[8] optimize the sum throughput of the network using dynamic programming and coupling methods, and show greedy round robin policies are optimal in some special cases; both positively and negatively correlated channels are studied. A Whittle's index [17] policy is constructed in [9] and shown to achieve near-optimal sum throughput by simulations. In this paper, we focus on analytically characterizing *the set of all achievable throughput vectors* in the network. One of the motivations is to study this partially observable wireless network using a mathematical programming approach, of which the first step is to characterize the performance region. An easy-to-use characterization of the performance region prepares us for studying more complex control problems using stochastic optimization theory. For example, in recent work [21] we have considered maximizing a concave utility function of throughput vectors (as limiting time averages) over the inner bound of the network capacity region we build in this paper. Equivalently, [21] considers optimizing *a general functional objective* over a restless bandit problem, which seems difficult using dynamic programming or Whittle's index theory that are used to optimize the often considered linear objectives.

This paper is also a study on efficient scheduling over wireless networks with delayed/uncertain channel state information (CSI) (see [12]–[14] and references therein). The work on delayed CSI that is most closely related to ours is [13], [14], where the authors study the capacity region and throughput-optimal policies of different wireless networks, assuming that channel states are *persistently* probed but fed back *with delay*. We note that our paper is significantly different. Here channels are never probed, and new (delayed) CSI of a channel is only acquired when the channel is served. Implicitly, acquiring the delayed CSI of any channel is part of the control decisions in this paper. This paper also applies to an important scenario in partial channel probing (see [2], [15] and references therein) where at most one channel is probed in every slot, and

[2]When there is one user, channel memory does not improve throughput. Serving the user every slot is throughput-optimal, and there is no need to consider channel memory.

data can only be served over the probed channel but not on unknown ones (as far as throughput is concerned and that we neglect probing overhead, this scenario is equivalent to blindly transmitting over a channel in every slot). Different from previous works which usually assume channels are i.i.d. over slots, here we show how channel memory improves throughput under a limited probing regime.

This paper is organized as follows. The network model is given in Section II, inner and outer bounds are constructed in Sections III and IV, and compared in Section V in the case of symmetric channels. Section VI gives the queue-dependent policy to achieve the inner bound.

II. Network Model

Consider a base station transmitting data to N users through N Markov ON/OFF channels. Suppose time is slotted with normalized slots $t \in \mathbb{Z}^+$. Each channel $n \in \{1, \dots, N\}$ is modeled as a two-state ON/OFF Markov chain (see Fig. 1). Let $s_n(t) \in \{\mathsf{OFF}, \mathsf{ON}\}$ denote the state of channel n in slot t.

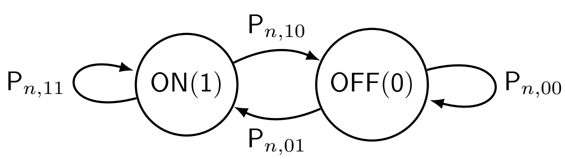

Fig. 1. A two-state Markov ON/OFF chain for channel $n \in \{1, 2, \dots, N\}$.

The state $s_n(t)$ of channel n evolves according to the transition probability matrix

$$\mathbf{P}_n = \begin{bmatrix} \mathsf{P}_{n,00} & \mathsf{P}_{n,01} \\ \mathsf{P}_{n,10} & \mathsf{P}_{n,11} \end{bmatrix},$$

where state ON is represented by 1 and OFF by 0, and $\mathsf{P}_{n,ij}$ denotes the transition probability from state i to j. We assume $\mathsf{P}_{n,11} < 1$ for all n so that no channel is constantly ON. Incorporating constantly ON channels like wired links is easy and thus omitted in this paper. We suppose channel states are fixed in every slot and may only change at slot boundaries. We assume all channels are positively correlated, which, in terms of transition probabilities, is equivalent to assuming $\mathsf{P}_{n,11} > \mathsf{P}_{n,01}$ or $\mathsf{P}_{n,01} + \mathsf{P}_{n,10} < 1$ for all n.[3] We suppose the base station keeps N queues of infinite capacity to store exogenous packet arrivals destined for the N users. At the beginning of every slot, the base station attempts to transmit a packet (if there is any) to a selected user. We suppose the base station has no channel probing capability and must select users oblivious of the current channel states. If a user is selected and its current channel state is ON, one packet is successfully delivered to that user. Otherwise, the transmission fails and zero packets are served. At the end of a slot in which the base station serves a user, an ACK/NACK message is fed back from the selected user to the base station through an independent error-free control channel, according to whether the transmission succeeds. Failing to receive an ACK is regarded as a NACK. Since channel states are either ON or OFF, such feedback reveals the channel state of the selected user in the last slot.

We define the N-dimensional *information state vector* $\boldsymbol{\omega}(t) = (\omega_n(t))_{n=1}^N$ where $\omega_n(t)$ is the probability that channel n is ON in slot t conditioning on the past observation history. In other words,

$$\omega_n(t) \triangleq \Pr\left[s_n(t) = \mathsf{ON} \mid \text{all past observations of channel } n\right].$$

We assume initially $\omega_n(0) = \pi_{n,\mathsf{ON}}$ for all n, where $\pi_{n,\mathsf{ON}}$ denotes the stationary probability that channel n is ON. As discussed in [5, Chapter 5.4], vector $\boldsymbol{\omega}(t)$ is a *sufficient statistic*. That is, instead of tracking the whole system history, the base station can act optimally only based on $\boldsymbol{\omega}(t)$. The base station shall keep track of the $\{\boldsymbol{\omega}(t)\}$ process.

Throughout the paper, we assume the transition probability matrices $\mathbf{P}_n$ of all channels are known to the base station. In practice, the matrix $\mathbf{P}_n$ for channel n may be learned in an initial training period, in which the base station continuously transmits packets over channel n in every slot. In this period we compute a sample average $\overline{Y_n}$ of the durations $(Y_{n,1}, Y_{n,2}, Y_{n,3}, \dots)$ that channel n is continuously ON. It is easy to see that $Y_{n,k}$ are i.i.d. over k with $\mathbb{E}[Y_{n.k}] = 1/\mathsf{P}_{n,10}$. As a result, we may use $1/\overline{Y_n}$ as an estimate of $\mathsf{P}_{n,10}$. The transition probability $\mathsf{P}_{n,01}$ can be estimated similarly. This estimation method works when channels are stationary.

Next, let $n(t) \in \{1, 2, \dots, N\}$ denote the user served in slot t. Based on the ACK/NACK feedback, the information state vector $\boldsymbol{\omega}(t)$ is updated as follows. For $1 \le n \le N$,

$$\omega_n(t+1) = \begin{cases} \mathsf{P}_{n,01}, & \text{if } n = n(t),\ s_n(t) = \mathsf{OFF} \\ \mathsf{P}_{n,11}, & \text{if } n = n(t),\ s_n(t) = \mathsf{ON} \\ \omega_n(t)\mathsf{P}_{n,11} + (1 - \omega_n(t))\mathsf{P}_{n,01}, & \text{if } n \neq n(t). \end{cases} \tag{1}$$

If in the most recent use of channel n, we observed (through feedback) its state was $i \in \{0, 1\}$ in slot $(t-k)$ for some $k \le t$, then $\omega_n(t)$ is equal to the k-step transition probability $\mathsf{P}_{n,i1}^{(k)}$. In general, for any fixed n, probabilities $\omega_n(t)$ take values in the countably infinite set $\mathcal{W}_n = \{\mathsf{P}_{n,01}^{(k)}, \mathsf{P}_{n,11}^{(k)} : k \in \mathbb{N}\} \cup \{\pi_{n,\mathsf{ON}}\}$. By eigenvalue decomposition on $\mathbf{P}_n$ [16, Chapter 4], we can show the k-step transition probability matrix $\mathbf{P}_n^{(k)}$ is

$$\begin{aligned} \mathbf{P}_n^{(k)} &\triangleq \begin{bmatrix} \mathsf{P}_{n,00}^{(k)} & \mathsf{P}_{n,01}^{(k)} \\ \mathsf{P}_{n,10}^{(k)} & \mathsf{P}_{n,11}^{(k)} \end{bmatrix} = (\mathbf{P}_n)^k \\ &= \frac{1}{x_n} \begin{bmatrix} \mathsf{P}_{n,10} + \mathsf{P}_{n,01}(1 - x_n)^k & \mathsf{P}_{n,01}\left(1 - (1 - x_n)^k\right) \\ \mathsf{P}_{n,10}(1 - (1 - x_n)^k) & \mathsf{P}_{n,01} + \mathsf{P}_{n,10}(1 - x_n)^k \end{bmatrix}, \end{aligned} \tag{2}$$

where we have defined $x_n \triangleq \mathsf{P}_{n,01} + \mathsf{P}_{n,10}$. Assuming that channels are positively correlated, i.e., $x_n < 1$, by (2) we have the following lemma.

Lemma 1. *For a positively correlated* ($\mathsf{P}_{n,11} > \mathsf{P}_{n,01}$) *Markov* ON/OFF *channel with transition probability matrix* $\mathbf{P}_n$*, we have*

[3] Assumption $\mathsf{P}_{n,11} > \mathsf{P}_{n,01}$ yields that the state $s_n(t)$ of channel n has auto-covariance $\mathbb{E}\left[(s_n(t) - \mathbb{E}s_n(t))(s_n(t+1) - \mathbb{E}s_n(t+1))\right] > 0$. In addition, we note that the case $\mathsf{P}_{n,11} = \mathsf{P}_{n,01}$ corresponds to a channel having i.i.d. states over slots. Although we can naturally incorporate i.i.d. channels into our model and all our results still hold, we exclude them in this paper because we shall show how throughput can be improved by channel memory, which i.i.d. channels do not have. The degenerate case where all channels are i.i.d. over slots is fully solved in [2].

1) *The stationary probability* $\pi_{n,\mathsf{ON}} = \mathsf{P}_{n,01}/x_n$.
2) *The* k*-step transition probability* $\mathsf{P}_{n,01}^{(k)}$ *is nondecreasing in* k *and* $\mathsf{P}_{n,11}^{(k)}$ *nonincreasing in* k*. Both* $\mathsf{P}_{n,01}^{(k)}$ *and* $\mathsf{P}_{n,11}^{(k)}$ *converge to* $\pi_{n,\mathsf{ON}}$ *as* $k \to \infty$.

As a corollary of Lemma 1, it follows that

$$\mathsf{P}_{n,11} \geq \mathsf{P}_{n,11}^{(k_1)} \geq \mathsf{P}_{n,11}^{(k_2)} \geq \pi_{n,\mathsf{ON}} \geq \mathsf{P}_{n,01}^{(k_3)} \geq \mathsf{P}_{n,01}^{(k_4)} \geq \mathsf{P}_{n,01} \tag{3}$$

for any integers $k_1 \leq k_2$ and $k_3 \geq k_4$ (see Fig. 2). To maximize network throughput, (3) has some fundamental implications. We note that $\omega_n(t)$ represents the transmission success probability over channel n in slot t. Thus we shall keep serving a channel whenever its information state is $\mathsf{P}_{n,11}$, for it is the best state possible. Second, given that a channel was OFF in its last use, its information state improves as long as the channel remains idle. Thus we shall wait as long as possible before reusing such a channel. Actually, when channels are symmetric ($\mathbf{P}_n = \mathbf{P}$ for all n), it is shown that a myopic policy with this structure maximizes the sum throughput of the network [7].

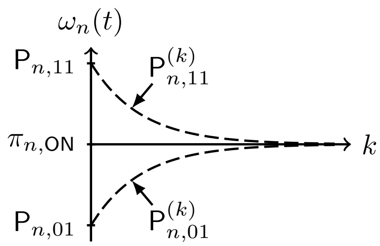

Fig. 2. Diagram of the k-step transition probabilities $\mathsf{P}_{n,01}^{(k)}$ and $\mathsf{P}_{n,11}^{(k)}$ of a positively correlated Markov ON/OFF channel.

III. A Round Robin Policy

For any integer $M \in \{1, 2, \ldots, N\}$, we present a special round robin policy $\mathsf{RR}(M)$ serving the first M users $\{1, 2, \ldots, M\}$ in the network. The M users are served in the circular order $1 \to 2 \to \cdots \to M \to 1 \to \cdots$. In general, we can use this policy to serve any subset of users.

A. The Policy

Round Robin Policy $\mathsf{RR}(M)$:

1) At time 0, the base station starts with channel 1. Suppose initially $\omega_n(0) = \pi_{n,\mathsf{ON}}$ for all n.
2) Suppose at time t, the base station switches to channel n. Transmit a *data* packet to user n with probability $\mathsf{P}_{n,01}^{(M)}/\omega_n(t)$ and a *dummy* packet otherwise. In both cases, we receive ACK/NACK information at the end of the slot.
3) At time $(t+1)$, if a dummy packet is sent at time t, switch to channel $(n \mod M) + 1$ and go to Step 2. Otherwise, keep transmitting data packets over channel n until we receive a NACK. Then switch to channel $(n \mod M) + 1$ and go to Step 2. We note that dummy packets are only sent on the first slot every time the base station switches to a new channel.
4) Update $\boldsymbol{\omega}(t)$ according to (1) in every slot.

Step 2 of $\mathsf{RR}(M)$ only makes sense if $\omega_n(t) \geq \mathsf{P}_{n,01}^{(M)}$, which we prove in the next lemma.

Lemma 2. *Under* $\mathsf{RR}(M)$*, whenever the base station switches to channel* $n \in \{1, 2, \ldots, M\}$ *for another round of transmission, its current information state satisfies* $\omega_n(t) \geq \mathsf{P}_{n,01}^{(M)}$.

Proof of Lemma 2: See Appendix A. ■

We note that policy $\mathsf{RR}(M)$ is very conservative and not throughput-optimal. For example, we can improve the throughput by always sending data packets but no dummy ones. Also, it does not follow the guidelines we provide at the end of Section II for maximum throughput. Yet, we will see that, in the case of symmetric channels, throughput under $\mathsf{RR}(M)$ is close to optimal when M is large. Moreover, the underlying analysis of $\mathsf{RR}(M)$ is tractable so that we can mix such round robin policies over different subsets of users to form a non-trivial inner capacity bound. The tractability of $\mathsf{RR}(M)$ is because it is equivalent to the following *fictitious* round robin policy (which can be proved as a corollary of Lemma 3 provided later).

Equivalent Fictitious Round Robin:

1) At time 0, start with channel 1.
2) When the base station switches to channel n, *set its current information state to* $\mathsf{P}_{n,01}^{(M)}$.[4] Keep transmitting data packets over channel n until we receive a NACK. Then switch to channel $(n \mod M) + 1$ and repeat Step 2.

For any round robin policy that serves channels in the circular order $1 \to 2 \to \cdots \to M \to 1 \to \cdots$, the technique of resetting the information state to $\mathsf{P}_{n,01}^{(M)}$ creates a system with an information state that is *worse* than the information state under the actual system. To see this, since in the actual system channels are served in the circular order, after we switch away from serving a particular channel n, we serve the other $(M-1)$ channels for at least one slot each, and so we return to channel n after at least M slots. Thus, its starting information state is always at least $\mathsf{P}_{n,01}^{(M)}$ (the proof is similar to that of Lemma 2). Intuitively, since information states represent the packet transmission success probabilities, resetting them to lower values degrades throughput. This is the reason why our inner capacity bound constructed later using $\mathsf{RR}(M)$ provides a throughput lower bound for a large class of policies.

B. Network Throughput under $\mathsf{RR}(M)$

Next we analyze the throughput vector achieved by $\mathsf{RR}(M)$.

1) General Case: Under $\mathsf{RR}(M)$, let L_{kn} denote the duration of the kth time the base station stays with channel n. A sample path of the $\{L_{kn}\}$ process is

$$(\underbrace{L_{11}, L_{12}, \ldots, L_{1M}}_{\text{round } k=1}, \underbrace{L_{21}, L_{22}, \ldots, L_{2M}}_{\text{round } k=2}, L_{31}, \ldots). \tag{4}$$

[4]In reality we cannot *set* the information state of a channel, and therefore the policy is fictitious.

The next lemma presents useful properties of L_{kn}, which serve as the foundation of the throughput analysis in the rest of the paper.

Lemma 3. *For any integer k and $n \in \{1, 2, \ldots, M\}$,*

1) *The probability mass function of L_{kn} is independent of k, and is*
$$L_{kn} = \begin{cases} 1 & \text{with prob. } 1 - \mathsf{P}_{n,01}^{(M)} \\ j \geq 2 & \text{with prob. } \mathsf{P}_{n,01}^{(M)} \left(\mathsf{P}_{n,11}\right)^{(j-2)} \mathsf{P}_{n,10}. \end{cases}$$
As a result, for all $k \in \mathbb{N}$ we have
$$\mathbb{E}\left[L_{kn}\right] = 1 + \frac{\mathsf{P}_{n,01}^{(M)}}{\mathsf{P}_{n,10}} = 1 + \frac{\mathsf{P}_{n,01}(1 - (1 - x_n)^M)}{x_n \mathsf{P}_{n,10}}.$$
2) *The number of data packets served in L_{kn} is $(L_{kn} - 1)$.*
3) *For every fixed channel n, time durations L_{kn} are* i.i.d. *random variables over all k.*

Proof of Lemma 3:

1) Note that $L_{kn} = 1$ if, on the first slot of serving channel n, either a dummy packet is transmitted or a data packet is transmitted but the channel is OFF. This event occurs with probability
$$\left(1 - \frac{\mathsf{P}_{n,01}^{(M)}}{\omega_n(t)}\right) + \frac{\mathsf{P}_{n,01}^{(M)}}{\omega_n(t)} \left(1 - \omega_n(t)\right) = 1 - \mathsf{P}_{n,01}^{(M)}.$$
Next, $L_{kn} = j \geq 2$ if in the first slot a data packet is successfully served, and this is followed by $(j-2)$ consecutive ON slots and one OFF slot. This happens with probability $\mathsf{P}_{n,01}^{(M)} \left(\mathsf{P}_{n,11}\right)^{(j-2)} \mathsf{P}_{n,10}$. The expectation of L_{kn} can be directly computed from the probability mass function.
2) We can observe that one data packet is served in every slot of L_{kn} except for the last one (when a dummy packet is sent over channel n, we have $L_{kn} = 1$ and zero data packets are served).
3) At the beginning of every L_{kn}, we observe from the equivalent fictitious round robin policy that $\mathsf{RR}(M)$ effectively fixes $\mathsf{P}_{n,01}^{(M)}$ as the current information state, regardless of the true current state $\omega_n(t)$. Neglecting $\omega_n(t)$ is to discard all system history, including all past $L_{k'n}$ for all $k' < k$. Thus L_{kn} are i.i.d.. Specifically, for any $k' < k$ and integers $l_{k'}$ and l_k we have
$$\Pr\left[L_{kn} = l_k \mid L_{k'n} = l_{k'}\right] = \Pr\left[L_{kn} = l_k\right].$$

∎

Now we can derive the throughput vector supported by $\mathsf{RR}(M)$. Fix an integer $K > 0$. By Lemma 3, the time average throughput over channel n after all channels finish their Kth rounds, which we denote by $\mu_n(K)$, is
$$\mu_n(K) \triangleq \frac{\sum_{k=1}^{K} (L_{kn} - 1)}{\sum_{k=1}^{K} \sum_{n=1}^{M} L_{kn}}.$$

Passing $K \to \infty$, we get
$$\begin{aligned} &\lim_{K\to\infty} \mu_n(K) \\ &= \lim_{K\to\infty} \frac{\sum_{k=1}^{K} (L_{kn} - 1)}{\sum_{k=1}^{K} \sum_{n=1}^{M} L_{kn}} \\ &= \lim_{K\to\infty} \frac{(1/K) \sum_{k=1}^{K} (L_{kn} - 1)}{\sum_{n=1}^{M} (1/K) \sum_{k=1}^{K} L_{kn}} \\ &\overset{(a)}{=} \frac{\mathbb{E}\left[L_{1n}\right] - 1}{\sum_{n=1}^{M} \mathbb{E}\left[L_{1n}\right]} \\ &\overset{(b)}{=} \frac{\mathsf{P}_{n,01}(1 - (1 - x_n)^M)/(x_n \mathsf{P}_{n,10})}{M + \sum_{n=1}^{M} \mathsf{P}_{n,01}(1 - (1 - x_n)^M)/(x_n \mathsf{P}_{n,10})}, \end{aligned} \tag{5}$$
where (a) is by the Law of Large Numbers (noting by Lemma 3 that L_{kn} are i.i.d. over k), and (b) is by Lemma 3.

2) Symmetric Case: We are particularly interested in the sum throughput under $\mathsf{RR}(M)$ when channels are symmetric, that is, all channels have the same statistics $\mathbf{P}_n = \mathbf{P}$ for all n. In this case, by channel symmetry every channel has the same throughput. From (5), we can show the sum throughput is
$$\sum_{n=1}^{M} \lim_{K\to\infty} \mu_n(K) = \frac{\mathsf{P}_{01}(1 - (1 - x)^M)}{x \, \mathsf{P}_{10} + \mathsf{P}_{01}(1 - (1 - x)^M)},$$
where in the last term the subscript n is dropped due to channel symmetry. It is handy to define a function $c_{(\cdot)} : \mathbb{N} \to \mathbb{R}$ as
$$c_M \triangleq \frac{\mathsf{P}_{01}(1 - (1 - x)^M)}{x \, \mathsf{P}_{10} + \mathsf{P}_{01}(1 - (1 - x)^M)}, \quad x \triangleq \mathsf{P}_{01} + \mathsf{P}_{10}, \tag{6}$$
and define $c_\infty \triangleq \lim_{M\to\infty} c_M = \mathsf{P}_{01}/(x\mathsf{P}_{10} + \mathsf{P}_{01})$ (note that $x < 1$ because every channel is positively correlated over time slots). The function $c_{(\cdot)}$ will be used extensively in this paper. We summarize the above derivation in the next lemma.

Lemma 4. *Policy $\mathsf{RR}(M)$ serves channel $n \in \{1, 2, \ldots, M\}$ with throughput*
$$\frac{\mathsf{P}_{n,01}(1 - (1 - x_n)^M)/(x_n \mathsf{P}_{n,10})}{M + \sum_{n=1}^{M} \mathsf{P}_{n,01}(1 - (1 - x_n)^M)/(x_n \mathsf{P}_{n,10})}.$$
In particular, in symmetric channels the sum throughput under $\mathsf{RR}(M)$ is c_M defined as
$$c_M = \frac{\mathsf{P}_{01}(1 - (1 - x)^M)}{x \, \mathsf{P}_{10} + \mathsf{P}_{01}(1 - (1 - x)^M)}, \quad x = \mathsf{P}_{01} + \mathsf{P}_{10},$$
and every channel has throughput c_M/M.

We remark that the sum throughput c_M of $\mathsf{RR}(M)$ in the symmetric case is nondecreasing in M, and thus can be improved by serving more channels. Also, we have shown in [2] that the maximum sum throughput when channel memory is neglected is equal to $\pi_{\mathsf{ON}} = c_1$, which is strictly less than the memory-assisted throughput c_M whenever $M \geq 2$ and $x < 1$. Interestingly, here we see that the sum throughput is improved by having *multiuser diversity* and channel memory in the network, even though instantaneous channel states are never known.

C. *How Good is* $\mathsf{RR}(M)$*?*

Next, in symmetric channels, we quantify how close the sum throughput c_M is to optimal. The following lemma presents a useful upper bound on the maximum sum throughput.

Lemma 5 ([6], [7])**.** *In symmetric channels, any scheduling policy that confines to our model has sum throughput less than or equal to* c_∞.[5]

By Lemma 4 and 5, the loss of the sum throughput of $\mathsf{RR}(M)$ is no larger than $c_\infty - c_M$. Define $\widetilde{c}_M$ as

$$\widetilde{c}_M \triangleq \frac{\mathsf{P}_{01}(1-(1-x)^M)}{x\mathsf{P}_{10}+\mathsf{P}_{01}} = c_\infty(1-(1-x)^M)$$

and note that $\widetilde{c}_M \leq c_M \leq c_\infty$. It follows

$$c_\infty - c_M \leq c_\infty - \widetilde{c}_M = c_\infty(1-x)^M. \tag{7}$$

The last term of (7) decreases to zero geometrically fast as M increases. This indicates that $\mathsf{RR}(M)$ yields near-optimal sum throughput even when it only serves a moderately large number of channels.

IV. Randomized Round Robin Policy, Inner and Outer Capacity Bound

A. *Randomized Round Robin Policy*

Lemma 4 specifies the throughput vector achieved by implementing $\mathsf{RR}(M)$ over a particular collection of M channels. Here we are interested in the set of throughput vectors achievable by randomly mixing $\mathsf{RR}(M)$-like policies over different channel subsets and allowing a different round-robin ordering on each subset. To generalize the $\mathsf{RR}(M)$ policy, first let Φ denote the set of all N-dimensional binary vectors excluding the all-zero vector $(0,0,\dots,0)$. For any binary vector $\boldsymbol{\phi} = (\phi_1,\phi_2,\dots,\phi_N)$ in Φ, we say channel n is *active* in $\boldsymbol{\phi}$ if $\phi_n = 1$. Each vector $\boldsymbol{\phi} \in \Phi$ represents a different subset of active channels. We denote by $M(\boldsymbol{\phi})$ the number of active channels in $\boldsymbol{\phi}$.

For each $\boldsymbol{\phi} \in \Phi$, consider the following round robin policy $\mathsf{RR}(\boldsymbol{\phi})$ that serves active channels in $\boldsymbol{\phi}$ in every round.

Dynamic Round Robin Policy $\mathsf{RR}(\boldsymbol{\phi})$:

1) *Deciding the service order in each round*:
   At the beginning of each round, we denote by τ_n the time duration between the last use of channel n and the beginning of the current round. Active channels in $\boldsymbol{\phi}$ are served in the decreasing order of τ_n in this round (in other words, the active channel that is *least recently used* is served first).
2) *On each active channel in a round*:
   a) Suppose at time t the base station switches to channel n. Transmit a *data* packet to user n with probability $\mathsf{P}_{n,01}^{(M(\boldsymbol{\phi}))}/\omega_n(t)$ and a *dummy* packet otherwise. In both cases, we receive ACK/NACK information at the end of the slot.
   b) At time $(t+1)$, if a dummy packet is sent at time t, switch to the next active channel following the order given in Step 1. Otherwise, keep transmitting data packets over channel n until we receive a NACK. Then switch to the next active channel and go to Step 2a. We note that dummy packets are only sent on the first slot every time the base station switches to a new channel.
3) Update $\boldsymbol{\omega}(t)$ according to (1) in every slot.

Using $\mathsf{RR}(\boldsymbol{\phi})$ as building blocks, we consider the following class of *randomized round robin policies*.

Randomized Round Robin Policy RandRR:

1) Pick $\boldsymbol{\phi} \in \Phi$ with probability $\alpha_{\boldsymbol{\phi}}$, where $\sum_{\boldsymbol{\phi}\in\Phi} \alpha_{\boldsymbol{\phi}} = 1$.
2) Run policy $\mathsf{RR}(\boldsymbol{\phi})$ for one round. Then go to Step 1.

Note that active channels may be served in different order in different rounds, according to the least-recently-used service order. This allows more time for OFF channels to return to better information states (note that $\mathsf{P}_{n,01}^{(k)}$ is nondecreasing in k) and thus improves throughput. The next lemma guarantees the feasibility of executing any $\mathsf{RR}(\boldsymbol{\phi})$ policy in RandRR (similar to Lemma 2, whenever the base station switches to a new channel n, we need $\omega_n(t) \geq \mathsf{P}_{n,01}^{(M(\boldsymbol{\phi}))}$ in Step 2a of $\mathsf{RR}(\boldsymbol{\phi})$).

Lemma 6. *When* $\mathsf{RR}(\boldsymbol{\phi})$ *is chosen by* RandRR *for a new round of transmission, every active channel* n *in* $\boldsymbol{\phi}$ *starts with information state no worse than* $\mathsf{P}_{n,01}^{(M(\boldsymbol{\phi}))}$.

Proof of Lemma 6: See Appendix B. ∎

Although RandRR randomly selects subsets of users and serves them in an order that depends on previous choices, we can surprisingly analyze its throughput. This is done by using the throughput analysis of $\mathsf{RR}(M)$, as shown in the following corollary to Lemma 3:

Corollary 1. *For each policy* $\mathsf{RR}(\boldsymbol{\phi})$*,* $\boldsymbol{\phi} \in \Phi$*, within time periods in which* $\mathsf{RR}(\boldsymbol{\phi})$ *is executed by* RandRR*, denote by* $L_{kn}^{\boldsymbol{\phi}}$ *the duration of the kth time the base station stays with active channel* n*. Then:*

1) *The probability mass function of* $L_{kn}^{\boldsymbol{\phi}}$ *is independent of* k*, and is*
$$L_{kn}^{\boldsymbol{\phi}} = \begin{cases} 1 & \text{with prob. } 1-\mathsf{P}_{n,01}^{(M(\boldsymbol{\phi}))} \\ j \geq 2 & \text{with prob. } \mathsf{P}_{n,01}^{(M(\boldsymbol{\phi}))}\,(\mathsf{P}_{n,11})^{(j-2)}\,\mathsf{P}_{n,10}. \end{cases}$$
   As a result, for all $k \in \mathbb{N}$ *we have*
$$\mathbb{E}\left[L_{kn}^{\boldsymbol{\phi}}\right] = 1 + \frac{\mathsf{P}_{n,01}^{(M(\boldsymbol{\phi}))}}{\mathsf{P}_{n,10}}. \tag{8}$$
2) *The number of data packets served in* $L_{kn}^{\boldsymbol{\phi}}$ *is* $(L_{kn}^{\boldsymbol{\phi}}-1)$.
3) *For every fixed* $\boldsymbol{\phi}$ *and every fixed active channel* n *in* $\boldsymbol{\phi}$*, the time durations* $L_{kn}^{\boldsymbol{\phi}}$ *are i.i.d. random variables over all* k.

B. *Achievable Network Capacity — An Inner Capacity Bound*

Using Corollary 1, next we present the achievable rate region of the class of RandRR policies. For each $\mathsf{RR}(\boldsymbol{\phi})$ policy,

[5] We note that the throughput analysis in [6] makes a minor assumption on the existence of some limiting time average. Using similar ideas of [6], in Theorem 2 of Section IV-C we will construct an upper bound on the maximum sum throughput for general positively correlated Markov ON/OFF channels. When restricted to the symmetric case, we get the same upper bound without any assumption.

define an N-dimensional vector $\boldsymbol{\eta}^{\boldsymbol{\phi}} = (\eta_1^{\boldsymbol{\phi}}, \eta_2^{\boldsymbol{\phi}}, \dots, \eta_N^{\boldsymbol{\phi}})$ where

$$\eta_n^{\boldsymbol{\phi}} \triangleq \begin{cases} \frac{\mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]-1}{\sum_{n:\phi_n=1}\mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]} & \text{if channel } n \text{ is active in } \boldsymbol{\phi}, \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

where $\mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]$ is given in (8). Intuitively, by the analysis prior to Lemma 4, round robin policy $\mathsf{RR}(\boldsymbol{\phi})$ yields throughput $\eta_n^{\boldsymbol{\phi}}$ over channel n for each $n \in \{1, 2, \dots, N\}$. Incorporating all possible random mixtures of $\mathsf{RR}(\boldsymbol{\phi})$ policies for different $\boldsymbol{\phi}$, RandRR can support any data rate vector that is entrywise dominated by a convex combination of vectors $\{\boldsymbol{\eta}^{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$ as shown by the next theorem.

Theorem 1 (Generalized Inner Capacity Bound)**.** *The class of* RandRR *policies supports all data rate vectors* $\boldsymbol{\lambda}$ *in the set* Λ_{int} *defined as*

$$\Lambda_{\mathit{int}} \triangleq \left\{\boldsymbol{\lambda} \mid \mathbf{0} \le \boldsymbol{\lambda} \le \boldsymbol{\mu},\ \boldsymbol{\mu} \in \operatorname{conv}\left(\{\boldsymbol{\eta}^{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}\right)\right\},$$

where $\boldsymbol{\eta}^{\boldsymbol{\phi}}$ *is defined in* (9)*,* $\operatorname{conv}(A)$ *denotes the convex hull of set* A*, and* $\le$ *is taken entrywise.*

Proof of Theorem 1: See Appendix C. ■

Applying Theorem 1 to symmetric channels yields the following corollary.

Corollary 2 (Inner Capacity Bound for Symmetric Channels)**.** *In symmetric channels, the class of* RandRR *policies supports all rate vectors* $\boldsymbol{\lambda} \in \Lambda_{\mathit{int}}$ *where*

$$\Lambda_{\mathit{int}} = \left\{\boldsymbol{\lambda} \mid \mathbf{0} \le \boldsymbol{\lambda} \le \boldsymbol{\mu},\, \boldsymbol{\mu} \in \operatorname{conv}\left(\left\{\frac{c_{M(\boldsymbol{\phi})}}{M(\boldsymbol{\phi})}\boldsymbol{\phi}\right\}_{\boldsymbol{\phi}\in\Phi}\right)\right\},$$

where $c_{M(\boldsymbol{\phi})}$ *is defined in* (6)*.*

The inner capacity bound Λ_{int} in Theorem 1 comprises all rate vectors that can be written as a convex combination of the zero vector and the throughput vectors $\boldsymbol{\eta}^{\boldsymbol{\phi}}$ (see (9)) yielded by the round robin policies $\mathsf{RR}(\boldsymbol{\phi})$ serving all subsets $\boldsymbol{\phi}$ of channels. This convex hull characterization shows that the inner bound Λ_{int} contains a very large class of policies and is intuitively near optimal because it is constructed by randomizing the efficient round robin policies on all subsets of channels. A simple example of the inner bound Λ_{int} will be provided later in Section IV-D.

We also remark that, although Λ_{int} is conceptually simple, supporting any given rate vector $\boldsymbol{\lambda}$ within Λ_{int} could be difficult because finding the right convex combination that supports $\boldsymbol{\lambda}$ is of exponential complexity. In Section VI, we provide a simple queue-dependent dynamic policy that supports all data rate vectors within the inner bound Λ_{int} *with polynomial complexity*.

C. Outer Capacity Bound

We construct an outer bound on Λ using several novel ideas. First, by state aggregation, we transform the information state process $\{\omega_n(t)\}$ for each channel n into non-stationary two-state Markov chains (in Fig. 4 provided later). Second, we create a set of bounding stationary Markov chains (in Fig. 5 provided later), which has the structure of a multi-armed bandit system. Finally, we create an outer capacity bound by relating the bounding model to the original non-stationary Markov chains using stochastic coupling. We note that since the control of the set of information state processes $\{\omega_n(t)\}$ for all n can be viewed as a restless bandit problem [17], it is interesting to see how we bound the optimal performance of a restless bandit problem by a related multi-armed bandit system.

We first *map* channel information states $\omega_n(t)$ into *modes* for each $n \in \{1, 2, \dots, N\}$. Inspired by (3), we observe that each channel n must be in one of the following two modes:

M1 The last observed state is ON, and the channel has not been seen (through feedback) to turn OFF. In this mode the information state $\omega_n(t) \in [\pi_{n,\mathsf{ON}}, \mathsf{P}_{n,11}]$.

M2 The last observed state is OFF, and the channel has not been seen to turned ON. Here $\omega_n(t) \in [\mathsf{P}_{n,01}, \pi_{n,\mathsf{ON}}]$.

On channel n, recall that $\mathcal{W}_n$ is the state space of $\omega_n(t)$, and define a map $f_n : \mathcal{W}_n \to \{\mathsf{M1}, \mathsf{M2}\}$ where

$$f_n(\omega_n(t)) = \begin{cases} \mathsf{M1} & \text{if } \omega_n(t) \in (\pi_{n,\mathsf{ON}}, \mathsf{P}_{n,11}], \\ \mathsf{M2} & \text{if } \omega_n(t) \in [\mathsf{P}_{n,01}, \pi_{n,\mathsf{ON}}]. \end{cases}$$

This mapping is illustrated in Fig. 3.

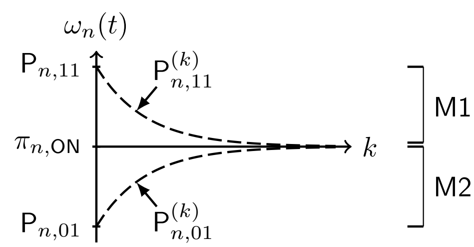

Fig. 3. The mapping f_n from information states $\omega_n(t)$ to modes $\{\mathsf{M1}, \mathsf{M2}\}$.

For any information state process $\{\omega_n(t)\}$ (controlled by some scheduling policy), the corresponding mode transition process under f_n can be represented by the Markov chains shown in Fig. 4. Specifically, when channel n is served in a slot, the associated mode transition follows the upper non-stationary chain of Fig. 4. When channel n is idled in a slot, the mode transition follows the lower stationary chain of Fig. 4. In the upper chain of Fig. 4, regardless what the current mode is, mode M1 is visited in the next slot if and only if channel n is ON in the current slot, which occurs with probability $\omega_n(t)$. In the lower chain of Fig. 4, when channel n is idled, its information state changes from a k-step transition probability to the $(k+1)$-step transition probability with the same most recent observed channel state. Therefore, the next mode stays the same as the current mode. We emphasize that, in the upper chain of Fig. 4, at mode M1 we always have $\omega_n(t) \le \mathsf{P}_{n,11}$, and at mode M2 it is $\omega_n(t) \le \pi_{n,\mathsf{ON}}$. A packet is served if and only if M1 is visited in the upper chain of Fig. 4.

To upper bound throughput, we compare Fig. 4 to the mode transition diagrams in Fig. 5 that corresponds to a fictitious model for channel n. This fictitious channel has constant information state $\omega_n(t) = \mathsf{P}_{n,11}$ whenever it is in mode M1, and $\omega_n(t) = \pi_{n,\mathsf{ON}}$ whenever it is in M2. In other words, when the fictitious channel n is in mode M1 (or M2), it sets its current information state to be the best state possible when the

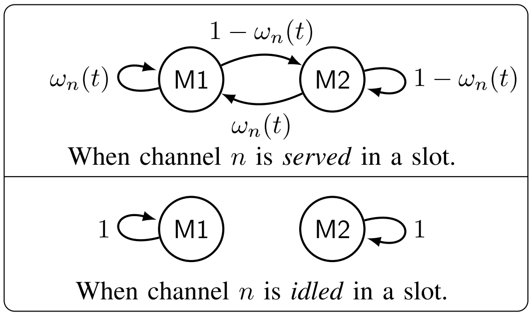

Fig. 4. Mode transition diagrams for the real channel n.

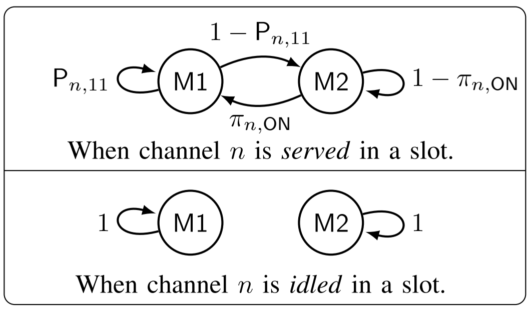

Fig. 5. Mode transition diagrams for the fictitious channel n.

corresponding real channel n is in the same mode. It follows that, when both the real and the fictitious channel n are served, the probabilities of transitions $\mathsf{M1} \to \mathsf{M1}$ and $\mathsf{M2} \to \mathsf{M1}$ in the upper chain of Fig. 5 are greater than or equal to those in Fig. 4, respectively. In other words, the upper chain of Fig. 5 is *more likely* to go to mode M1 and serve packets than that of Fig. 4. Therefore, intuitively, if we serve both the real and the fictitious channel n in the same infinite sequence of time slots, the fictitious channel n will yield higher throughput for all n. This observation is made precise by the next lemma.

Lemma 7. *Consider two discrete-time Markov chains $\{X(t)\}$ and $\{Y(t)\}$ both with state space $\{0,1\}$. Suppose $\{X(t)\}$ is stationary and ergodic with transition probability matrix*

$$\mathbf{P} = \begin{bmatrix} \mathsf{P}_{00} & \mathsf{P}_{01} \\ \mathsf{P}_{10} & \mathsf{P}_{11} \end{bmatrix},$$

and $\{Y(t)\}$ is non-stationary with

$$\mathbf{Q}(t) = \begin{bmatrix} \mathsf{Q}_{00}(t) & \mathsf{Q}_{01}(t) \\ \mathsf{Q}_{10}(t) & \mathsf{Q}_{11}(t) \end{bmatrix}.$$

Assume $\mathsf{P}_{01} \geq \mathsf{Q}_{01}(t)$ and $\mathsf{P}_{11} \geq \mathsf{Q}_{11}(t)$ for all t. In $\{X(t)\}$, let $\pi_X(1)$ denote the stationary probability of state 1; $\pi_X(1) = \mathsf{P}_{01}/(\mathsf{P}_{01}+\mathsf{P}_{10})$. In $\{Y(t)\}$, define

$$\pi_Y(1) \triangleq \limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} Y(t)$$

as the limiting fraction of time $\{Y(t)\}$ stays at state 1. Then we have $\pi_X(1) \geq \pi_Y(1)$.

Proof of Lemma 7: Given in Appendix E. ∎

We note that executing a scheduling policy in the network is to generate a sequence of channel selection decisions. By Lemma 7, if we apply the same sequence of channel selection decisions of some scheduling policy to the set of fictitious channels, we will get higher throughput on every channel. A direct consequence of this is that the maximum sum throughput over the fictitious channels is greater than or equal to that over the real channels.

Lemma 8. *The maximum sum throughput over the set of fictitious channels is no more than*

$$\max_{n\in\{1,2,\dots,N\}}\{c_{n,\infty}\}, \quad c_{n,\infty} \triangleq \frac{\mathsf{P}_{n,01}}{x_n \mathsf{P}_{n,10} + \mathsf{P}_{n,01}}.$$

Proof of Lemma 8: We note that finding the maximum sum throughput over fictitious channels in Fig. 5 is equivalent to solving a multi-armed bandit problem [18] with each channel acting as an arm (see Fig. 5 and note that a channel can change mode only when it is served), and one unit of reward is earned if a packet is delivered (recall that a packet is served if and only if mode M1 is visited in the upper chain of Fig. 5). The optimal solution to the multi-armed bandit system is to always play the arm (channel) with the largest average reward (throughput). The average reward over channel n is equal to the stationary probability of mode M1 in the upper chain of Fig. 5, which is

$$\frac{\pi_{n,\mathsf{ON}}}{\mathsf{P}_{n,10} + \pi_{n,\mathsf{ON}}} = \frac{\mathsf{P}_{n,01}}{x_n \mathsf{P}_{n,10} + \mathsf{P}_{n,01}}.$$

This finishes the proof. ∎

Together with the fact that throughput over any real channel n cannot exceed its stationary ON probability $\pi_{n,\mathsf{ON}}$, we have constructed an outer bound on the network capacity region Λ (the proof follows the above discussions and thus is omitted).

Theorem 2. *(Generalized Outer Capacity Bound): Any supportable throughput vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_N)$ necessarily satisfies*

$$\begin{aligned} \lambda_n &\leq \pi_{n,\mathsf{ON}}, \quad \textit{for all } n \in \{1,2,\dots,N\}, \\ \sum_{n=1}^{N} \lambda_n &\leq \max_{n\in\{1,2,\dots,N\}} \{c_{n,\infty}\} \\ &= \max_{n\in\{1,2,\dots,N\}} \left\{ \frac{\mathsf{P}_{n,01}}{x_n \mathsf{P}_{n,10} + \mathsf{P}_{n,01}} \right\}. \end{aligned}$$

These $(N+1)$ hyperplanes create an outer capacity bound Λ_{out} on Λ.

Corollary 3 (Outer Capacity Bound for Symmetric Channels)**.** *In symmetric channels with $\mathbf{P}_n = \mathbf{P}$, $c_{n,\infty} = c_\infty$, and $\pi_{n,\mathsf{ON}} = \pi_{\mathsf{ON}}$ for all n, we have*

$$\Lambda_{out} = \left\{ \boldsymbol{\lambda} \geq \mathbf{0} \mid \sum_{n=1}^{N} \lambda_n \leq c_\infty,\ \lambda_n \leq \pi_{\mathsf{ON}} \text{ for } 1 \leq n \leq N \right\}, \tag{10}$$

where $\geq$ is taken entrywise.

We note that Lemma 5 in Section III-C directly follows Corollary 3.

D. A Two-User Example on Symmetric Channels

Here we consider a two-user example on symmetric channels. For simplicity we will drop the subscript n in notations. From Corollary 3, we have the outer bound

$$\Lambda_{\text{out}} = \left\{ \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \;\middle|\; \begin{array}{c} 0 \leq \lambda_n \leq \mathsf{P}_{01}/x, \text{ for } 1 \leq n \leq 2, \\ \lambda_1 + \lambda_2 \leq \mathsf{P}_{01}/(x\mathsf{P}_{10} + \mathsf{P}_{01}), \\ x = \mathsf{P}_{01} + \mathsf{P}_{10} \end{array} \right\}.$$

For the inner bound Λ_{int}, we note that policy RandRR can execute three round robin policies $\mathsf{RR}(\boldsymbol{\phi})$ for $\boldsymbol{\phi} \in \Phi = \{(1,1),(0,1),(1,0)\}$. From Corollary 2, we have

$$\Lambda_{\text{int}} = \left\{ \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \;\middle|\; \begin{array}{c} 0 \leq \lambda_n \leq \mu_n, \text{ for } 1 \leq n \leq 2, \\ \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \in \text{conv}\left(\left\{ \begin{bmatrix} c_2/2 \\ c_2/2 \end{bmatrix}, \begin{bmatrix} c_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ c_1 \end{bmatrix} \right\} \right) \end{array} \right\}.$$

Under the special case $\mathsf{P}_{01} = \mathsf{P}_{10} = 0.2$, the two bounds λ_{int} and Λ_{out} are shown in Fig. 6.

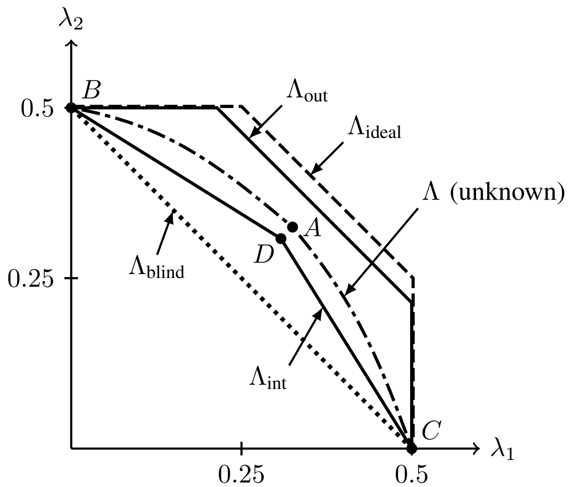

Fig. 6. Comparison of rate regions under different assumptions.

In Fig. 6, we also compare Λ_{int} and Λ_{out} with other rate regions. Set Λ_{ideal} is the ideal capacity region when instantaneous channel states are known without causing any (timing) overhead [19]. Next, it is shown in [6] that the maximum sum throughput in this network is achieved at point $A = (0.325, 0.325)$. The (unknown) network capacity region Λ is bounded between Λ_{int} and Λ_{out}, and has boundary points B, A, and C. Since the boundary of Λ is a concave curve connecting B, A, and C, we envision that Λ shall contain but be very close to Λ_{int}.

Finally, the rate region Λ_{blind} is rendered by completely neglecting channel memory and treating the channels as i.i.d. over slots [2]. We observe the throughput gain $\Lambda_{\text{int}} \setminus \Lambda_{\text{blind}}$, as much as 23% in this example, is achieved by incorporating channel memory. In general, if channels are symmetric and treated as i.i.d. over slots, the maximum sum throughput in the network is $\pi_{\mathsf{ON}} = c_1$. Then the maximum throughput gain of RandRR using channel memory is $c_N - c_1$, which as $N \to \infty$ converges to

$$c_\infty - c_1 = \frac{\mathsf{P}_{01}}{x\mathsf{P}_{10} + \mathsf{P}_{01}} - \frac{\mathsf{P}_{01}}{\mathsf{P}_{01} + \mathsf{P}_{10}},$$

which is controlled by the factor $x = \mathsf{P}_{01} + \mathsf{P}_{10}$.

E. A Heuristically Tighter Inner Bound

It is shown in [7] that the following policy maximizes the sum throughput in a symmetric network:

> *Serve channels in a circular order, where on each channel keep transmitting data packets until a NACK is received.*

In the above two-user example, this policy achieves throughput vector A in Fig. 7. If we replace our round robin policy

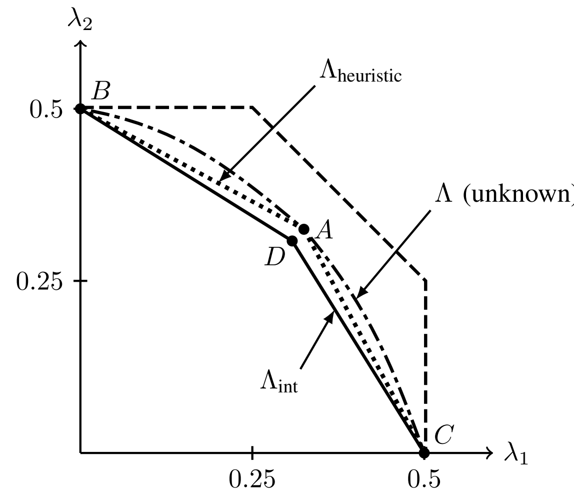

Fig. 7. Comparison of our inner bound Λ_{int}, the unknown network capacity region Λ, and a heuristically better inner bound $\Lambda_{\text{heuristic}}$.

$\mathsf{RR}(\boldsymbol{\phi})$ by this one, heuristically we are able to construct a tighter inner capacity bound. For example, we can support the tighter inner bound $\Lambda_{\text{heuristic}}$ in Fig. 7 by appropriate time sharing among the above policy that serves different subsets of channels. However, we note that this approach is difficult to analyze because the $\{L_{kn}\}$ process (see (4)) forms a high-order Markov chain. Yet, our inner bound Λ_{int} provides a good throughput guarantee for this class of heuristic policies.

V. Proximity of the Inner Bound to the True Capacity Region — Symmetric Case

Next we bound the closeness of the boundaries of Λ_{int} and Λ in the case of symmetric channels. In Section III-C, by choosing $M = N$, we have provided such analysis for the boundary point in the direction $(1,1,\dots,1)$. Here we generalize to all boundary points. Define

$$\mathcal{V} \triangleq \left\{ (v_1, v_2, \dots, v_N) \;\middle|\; \begin{array}{c} v_n \geq 0 \text{ for } 1 \leq n \leq N, \\ v_n > 0 \text{ for at least one } n \end{array} \right\}$$

as a set of directional vectors. For any $\boldsymbol{v} \in \mathcal{V}$, let $\boldsymbol{\lambda}^{\text{int}} = (\lambda_1^{\text{int}}, \lambda_2^{\text{int}}, \dots, \lambda_N^{\text{int}})$ and $\boldsymbol{\lambda}^{\text{out}} = (\lambda_1^{\text{out}}, \lambda_2^{\text{out}}, \dots, \lambda_N^{\text{out}})$ be the boundary point of Λ_{int} and Λ_{out} in the direction of $\boldsymbol{v}$, respectively. It is useful to compute $\sum_{n=1}^N (\lambda_n^{\text{out}} - \lambda_n^{\text{int}})$, because it upper bounds the loss of the sum throughput of Λ_{int} from Λ in the direction of $\boldsymbol{v}$.[6] We note that computing $\boldsymbol{\lambda}^{\text{int}}$ in an arbitrary

[6]Note that $\sum_{n=1}^N (\lambda_n^{\text{out}} - \lambda_n^{\text{int}})$ also bounds the closeness between Λ_{out} and Λ.

direction is difficult. Thus we will find an upper bound on $\sum_{n=1}^{N}(\lambda_n^{\text{out}} - \lambda_n^{\text{int}})$.

A. Preliminary

To have more intuitions on Λ_{int}, we start with a toy example of $N = 3$ users. We are interested in the boundary point of Λ_{int} in the direction of $\boldsymbol{v} = (1, 2, 1)$. Consider two RandRR-type policies ψ_1 and ψ_2 defined as follows.

$$\text{For } \psi_1, \text{ choose } \begin{cases} \boldsymbol{\phi}^1 = (1,0,0) & \text{with prob. } 1/4 \\ \boldsymbol{\phi}^2 = (0,1,0) & \text{with prob. } 1/2 \\ \boldsymbol{\phi}^3 = (0,0,1) & \text{with prob. } 1/4 \end{cases}$$

$$\text{For } \psi_2, \text{ choose } \begin{cases} \boldsymbol{\phi}^4 = (1,1,0) & \text{with prob. } 1/2 \\ \boldsymbol{\phi}^5 = (0,1,1) & \text{with prob. } 1/2 \end{cases}$$

Both ψ_1 and ψ_2 support data rates in the direction of $(1, 2, 1)$. However, using the analysis of Lemma 4 and Theorem 1, we know ψ_1 supports throughput vector

$$\frac{1}{4}\begin{bmatrix} c_1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ c_1 \\ 0 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} 0 \\ 0 \\ c_1 \end{bmatrix} = \frac{c_1}{4}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},$$

while ψ_2 supports

$$\frac{1}{2}\begin{bmatrix} c_2/2 \\ c_2/2 \\ 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ c_2/2 \\ c_2/2 \end{bmatrix} = \frac{c_2}{4}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \geq \frac{c_1}{4}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},$$

where c_1 and c_2 are defined in (6). We see that ψ_2 achieves data rates closer than ψ_1 does to the boundary of Λ_{int}. It is because every sub-policy of ψ_2, namely $\mathsf{RR}(\boldsymbol{\phi}^4)$ and $\mathsf{RR}(\boldsymbol{\phi}^5)$, supports sum throughput c_2 (by Lemma 4), where those of ψ_1 only support c_1. In other words, policy ψ_2 has better *multiuser diversity gain* than ψ_1 does. This example suggests that we can find a good lower bound on $\boldsymbol{\lambda}^{\text{int}}$ by exploring to what extent the multiuser diversity can be exploited. We start with the following definition.

Definition 1. *For any $\boldsymbol{v} \in \mathcal{V}$, we say $\boldsymbol{v}$ is d-*user diverse *if $\boldsymbol{v}$ can be written as a* positive *combination of vectors in Φ_d, where Φ_d denotes the set of N-dimensional binary vectors having d entries be 1. Define*

$$d(\boldsymbol{v}) \triangleq \max_{1 \leq d \leq N} \{d \mid \boldsymbol{v} \text{ is } d\text{-user diverse}\},$$

and we shall say $\boldsymbol{v}$ is maximally $d(\boldsymbol{v})$-user diverse*.*

The notion of $d(\boldsymbol{v})$ is well-defined because every $\boldsymbol{v}$ must be 1-user diverse.[7] Definition 1 is the most useful to us through the next lemma.

Lemma 9. *The boundary point of Λ_{int} in the direction of $\boldsymbol{v} \in \mathcal{V}$ has sum throughput at least $c_{d(\boldsymbol{v})}$, where*

$$c_{d(\boldsymbol{v})} \triangleq \frac{\mathsf{P}_{01}(1-(1-x)^{d(\boldsymbol{v})})}{x\,\mathsf{P}_{10} + \mathsf{P}_{01}(1-(1-x)^{d(\boldsymbol{v})})}, \quad x \triangleq \mathsf{P}_{01} + \mathsf{P}_{10}.$$

[7]The set $\Phi_1 = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$ is the collection of unit coordinate vectors where $\boldsymbol{e}_n$ has its nth entry be 1 and 0 otherwise. Any vector $\boldsymbol{v} \in \mathcal{V}$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_N)$, can be written as $\boldsymbol{v} = \sum_{v_n > 0} v_n \boldsymbol{e}_n$.

Proof of Lemma 9: If direction $\boldsymbol{v}$ can be written as a positive weighted sum of vectors in $\Phi_{d(\boldsymbol{v})}$, we can normalize the weights, and use the new weights as probabilities to randomly mix $\mathsf{RR}(\boldsymbol{\phi})$ policies for all $\boldsymbol{\phi} \in \Phi_{d(\boldsymbol{v})}$. This way we achieve sum throughput $c_{d(\boldsymbol{v})}$ in every transmission round, and overall the throughput vector will be in the direction of $\boldsymbol{v}$. Therefore the result follows. For details, see Appendix G. ∎

Fig. 8 provides an example of Lemma 9 in the two-user symmetric system in Section IV-D. We observe that

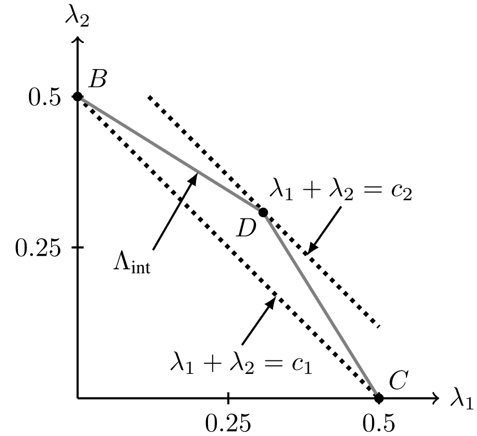

Fig. 8. An example for Lemma 9 in the two-user symmetric network. Point B and C achieve sum throughput $c_1 = \pi_{\mathsf{ON}} = 0.5$, and the sum throughput at D is $c_2 \approx 0.615$. Any other boundary point of Λ_{int} has sum throughput between c_1 and c_2.

direction $(1, 1)$, the one that passes point D in Fig. 8, is the only direction that is maximally 2-user diverse. The sum throughput c_2 is achieved at D. For all the other directions, they are maximally 1-user diverse and, from Fig. 8, only sum throughput c_1 is guaranteed along those directions. In general, geometrically we can show that a maximally d-user diverse vector, say $\boldsymbol{v}_d$, forms a smaller angle with the all-1 vector $(1, 1, \ldots, 1)$ than a maximally d'-user diverse vector, say $\boldsymbol{v}_{d'}$, does if $d' < d$. In other words, data rates along $\boldsymbol{v}_d$ are *more balanced* than those along $\boldsymbol{v}_{d'}$. Lemma 9 states that we guarantee to support higher sum throughput if the user traffic is more balanced.

B. Proximity Analysis

We use the notion of $d(\boldsymbol{v})$ to upper bound $\sum_{n=1}^{N}(\lambda_n^{\text{out}} - \lambda_n^{\text{int}})$ in any direction $\boldsymbol{v} \in \mathcal{V}$. Let $\boldsymbol{\lambda}^{\text{out}} = \theta\boldsymbol{\lambda}^{\text{int}}$ (i.e., $\lambda_n^{\text{out}} = \theta\lambda_n^{\text{int}}$ for all n) for some $\theta \geq 1$. By (10), the boundary of Λ_{out} is characterized by the interaction of the $(N + 1)$ hyperplanes $\sum_{n=1}^{N} \lambda_n = c_\infty$ and $\lambda_n = \pi_{\mathsf{ON}}$ for each $n \in \{1, 2, \ldots, N\}$. Specifically, in any given direction, if we consider the cross points on all the hyperplanes in that direction, the boundary point $\boldsymbol{\lambda}^{\text{out}}$ is the one closest to the origin. We do not know which hyperplane $\boldsymbol{\lambda}^{\text{out}}$ is on, and thus need to consider all $(N + 1)$ cases. If $\boldsymbol{\lambda}^{\text{out}}$ is on the plane $\sum_{n=1}^{N} \lambda_n = c_\infty$, i.e., $\sum_{n=1}^{N} \lambda_n^{\text{out}} = c_\infty$, we get

$$\sum_{n=1}^{N}(\lambda_n^{\text{out}} - \lambda_n^{\text{int}}) \overset{(a)}{\leq} c_\infty - c_{d(\boldsymbol{v})} \overset{(b)}{\leq} c_\infty(1-x)^{d(\boldsymbol{v})},$$

where (a) is by Lemma 9 and (b) is by (7). If $\boldsymbol{\lambda}^{\text{out}}$ is on the plane $\lambda_n = \pi_{\mathsf{ON}}$ for some n, then $\theta = \pi_{\mathsf{ON}}/\lambda_n^{\text{int}}$. It follows

$$\sum_{n=1}^{N}(\lambda_n^{\text{out}} - \lambda_n^{\text{int}}) = (\theta-1)\sum_{n=1}^{N}\lambda_n^{\text{int}} \leq \left(\frac{\pi_{\mathsf{ON}}}{\lambda_n^{\text{int}}} - 1\right) c_\infty.$$

The above discussions lead to the next lemma.

Lemma 10. *The loss of the sum throughput of Λ_{int} from Λ in the direction of $\boldsymbol{v}$ is upper bounded by*

$$\min\left[c_\infty(1-x)^{d(\boldsymbol{v})},\ \min_{1\leq n\leq N}\left\{\left(\frac{\pi_{\mathsf{ON}}}{\lambda_n^{int}}-1\right)c_\infty\right\}\right]$$
$$= c_\infty \min\left[(1-x)^{d(\boldsymbol{v})},\ \frac{\pi_{\mathsf{ON}}}{\max_{1\leq n\leq N}\{\lambda_n^{int}\}}-1\right]. \quad (11)$$

Lemma 10 shows that, if data rates are more balanced, namely, have a larger $d(\boldsymbol{v})$, the sum throughput loss is dominated by the first term in the minimum of (11) and decreases to 0 geometrically fast with $d(\boldsymbol{v})$. If data rates are biased toward a particular user, the second term in the minimum of (11) captures the throughput loss, which goes to 0 as the rate of the favored user goes to the single-user capacity π_{ON}.

VI. Throughput-Achieving Queue-dependent Round Robin Policy

Let $a_n(t)$, for $1 \leq n \leq N$, be the number of exogenous packet arrivals destined for user n in slot t. Suppose $a_n(t)$ are independent across users, i.i.d. over slots with rate $\mathbb{E}[a_n(t)] = \lambda_n$, and $a_n(t)$ is bounded with $0 \leq a_n(t) \leq A_{\max}$, where $A_{\max}$ is a finite integer. Let $Q_n(t)$ be the backlog of user-n packets queued at the base station at time t. Define $\boldsymbol{Q}(t) \triangleq (Q_1(t), Q_2(t), \ldots, Q_N(t))$ and suppose $Q_n(0) = 0$ for all n. The queue process $\{Q_n(t)\}$ evolves as

$$Q_n(t+1) = \max\left[Q_n(t) - \mu_n(s_n(t),t), 0\right] + a_n(t), \quad (12)$$

where $\mu_n(s_n(t),t) \in \{0,1\}$ is the service rate allocated to user n in slot t. We have $\mu_n(s_n(t),t) = 1$ if user n is served and $s_n(t) = \mathsf{ON}$, and 0 otherwise. In the rest of the paper we drop $s_n(t)$ in $\mu_n(s_n(t),t)$ and use $\mu_n(t)$ for notational simplicity. We say the network is (strongly) stable if

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{n=1}^{N}\mathbb{E}\left[Q_n(\tau)\right] < \infty.$$

Consider a rate vector $\boldsymbol{\lambda}$ interior to the inner capacity region bound Λ_{int} given in Theorem 1. Namely, there exists an $\epsilon > 0$ and a probability distribution $\{\beta_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$ such that

$$\lambda_n + \epsilon < \sum_{\boldsymbol{\phi}\in\Phi}\beta_{\boldsymbol{\phi}}\eta_n^{\boldsymbol{\phi}}, \quad \text{for all } 1\leq n\leq N, \quad (13)$$

where $\eta_n^{\boldsymbol{\phi}}$ is defined in (9). By Theorem 1, there exists a RandRR policy that yields service rates equal to the right-side of (13) and thus stabilizes the network with arrival rate vector $\boldsymbol{\lambda}$ [20, Lemma 3.6]. The existence of this policy is useful and we shall denote it by RandRR^*. Recall that on each new scheduling round, the policy RandRR^* randomly picks a binary vector $\boldsymbol{\phi}$ using probabilities $\alpha_{\boldsymbol{\phi}}$ (defined over all of the (2^N-1) subsets of users). The $M(\boldsymbol{\phi})$ active users in $\boldsymbol{\phi}$ are served for one round by the round robin policy $\mathsf{RR}(\boldsymbol{\phi})$, serving the least recently used users first. However, solving for the probabilities needed to implement the RandRR^* policy that yields (13) is intractable when N is large, because we need to find (2^N-1) unknown probabilities $\{\alpha_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$, compute $\{\beta_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$ from (19), and make (13) hold. Instead of probabilistically finding the vector $\boldsymbol{\phi}$ for the current round of scheduling, we use the following simple *queue-dependent* policy.

Queue-dependent Round Robin Policy (QRR):

1) Start with $t = 0$.
2) At time t, observe the current queue backlog vector $\boldsymbol{Q}(t)$ and find the binary vector $\boldsymbol{\phi}(t) \in \Phi$ defined as[8]

$$\boldsymbol{\phi}(t) \triangleq \arg\max_{\boldsymbol{\phi}\in\Phi} f(\boldsymbol{Q}(t), \mathsf{RR}(\boldsymbol{\phi})), \quad (14)$$

where

$$f(\boldsymbol{Q}(t), \mathsf{RR}(\boldsymbol{\phi}))$$
$$\triangleq \sum_{n:\phi_n=1}\left[Q_n(t)\mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}-1\right] - \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]\sum_{n=1}^{N}Q_n(t)\lambda_n\right]$$

and $\mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right] = 1 + \mathsf{P}_{n,01}^{(M(\boldsymbol{\phi}))}/\mathsf{P}_{n,10}$ from (8). Ties are broken arbitrarily.[9]
3) Run $\mathsf{RR}(\boldsymbol{\phi}(t))$ for one round of transmission. We emphasize that active channels in $\boldsymbol{\phi}$ are served in the least-recently-used order. After the round ends, go to Step 2.

The QRR policy is a frame-based algorithm similar to RandRR, except that at the beginning of every transmission round the policy selection is no longer random but based on a queue-dependent rule. We note that QRR is a polynomial time algorithm because we can compute $\boldsymbol{\phi}(t)$ in (14) in polynomial time with the following divide and conquer approach:

1) Partition the set Φ into subsets $\{\Phi_1, \ldots, \Phi_N\}$, where Φ_M, $M \in \{1, \ldots, N\}$, is the set of N-dimensional binary vectors having exactly M entries be 1.
2) For each $M \in \{1, \ldots, N\}$, find the maximizer of $f(\boldsymbol{Q}(t), \mathsf{RR}(\boldsymbol{\phi}))$ among vectors in Φ_M. For each $\boldsymbol{\phi} \in \Phi_M$, we have

$$f(\boldsymbol{Q}(t), \mathsf{RR}(\boldsymbol{\phi})) =$$
$$\sum_{n:\phi_n=1}\left[Q_n(t)\frac{\mathsf{P}_{n,01}^{(M)}}{\mathsf{P}_{n,10}} - \left(1+\frac{\mathsf{P}_{n,01}^{(M)}}{\mathsf{P}_{n,10}}\right)\sum_{n=1}^{N}Q_n(t)\lambda_n\right],$$

and the maximizer of $f(\boldsymbol{Q}(t), \mathsf{RR}(\boldsymbol{\phi}))$ is to *activate* the M channels that yield the M largest summands of the above equation.
3) Obtain $\boldsymbol{\phi}(t)$ by comparing the maximizers from the above step for different values of M.

The detailed implementation is as follows.

Polynomial time implementation of Step 2 of QRR:

1) For each fixed $M \in \{1, \ldots, N\}$, we do the following:

[8] The vector $\boldsymbol{\phi}(t)$ is a queue-dependent decision and thus we should write $\boldsymbol{\phi}(\boldsymbol{Q}(t), t)$ as a function of $\boldsymbol{Q}(t)$. For simplicity we use $\boldsymbol{\phi}(t)$ instead.

[9] In (50) we show that as long as the queue backlog vector $\boldsymbol{Q}(t)$ is not identically zero and the arrival rate vector $\boldsymbol{\lambda}$ is interior to the inner capacity bound Λ_{int}, we always have $\max_{\boldsymbol{\phi}\in\Phi} f(\boldsymbol{Q}(t), \mathsf{RR}(\boldsymbol{\phi})) > 0$.

Compute

$$Q_n(t)\frac{\mathsf{P}_{n,01}^{(M)}}{\mathsf{P}_{n,10}} - \left(1+\frac{\mathsf{P}_{n,01}^{(M)}}{\mathsf{P}_{n,10}}\right)\sum_{n=1}^{N} Q_n(t)\lambda_n \tag{15}$$

for all $n \in \{1,\dots,N\}$. Sort these N numbers and define the binary vector $\boldsymbol{\phi}^M = (\phi_1^M,\dots,\phi_N^M)$ such that $\phi_n^M = 1$ if the value (15) of channel n is among the M largest, otherwise $\phi_n^M = 0$. Ties are broken arbitrarily. Let $\hat{f}(\boldsymbol{Q}(t), M)$ denote the sum of the M largest values of (15).

2) Define $M(t) \triangleq \arg\max_{1\le M\le N} \hat{f}(\boldsymbol{Q}(t), M)$. Then we assign $\boldsymbol{\phi}(t) = \boldsymbol{\phi}^{M(t)}$.

Using a novel variable-length frame-based Lyapunov analysis, we show in the next theorem that QRR stabilizes the network with any arrival rate vector $\boldsymbol{\lambda}$ strictly within the inner capacity bound Λ_{int}. The idea is that we compare QRR with the (unknown) policy RandRR^* that stabilizes $\boldsymbol{\lambda}$. We show that, in every transmission round, QRR finds and executes a round robin policy $\mathsf{RR}(\boldsymbol{\phi}(t))$ that yields a larger negative drift on the queue backlogs than RandRR^* does in the current round. Therefore, QRR is stable.

Theorem 3. *For any data rate vector $\boldsymbol{\lambda}$ interior to Λ_{int}, policy* QRR *strongly stabilizes the network.*

Proof of Theorem 3: See Appendix H. ∎

VII. Conclusion

The network capacity of a wireless network is practically degraded by communication overhead. In this paper, we take a step forward by studying the fundamental achievable rate region when communication overhead is kept minimum, that is, when channel probing is not permitted. While solving the original problem is difficult, we construct an inner and an outer bound on the network capacity region, with the aid of channel memory. When channels are symmetric and the network serves a large number of users, we show the inner and outer bound are progressively tight when the data rates of different users are more balanced. We also derive a simple queue-dependent frame-based policy, as a function of packet arrival rates and channel statistics, and show that this policy stabilizes the network for any data rates strictly within the inner capacity bound.

Transmitting data without channel probing is one of the many options for communication over a wireless network. Practically each option may have pros and cons on criteria like the achievable throughput, power efficiency, implementation complexity, etc. In the future it is important to explore how to combine all possible options to push the practically achievable network capacity to the limit. It is part of our future work to generalize the methodology and framework developed in this paper to more general cases, such as when limited probing is allowed and/or other QoS metrics such as energy consumption are considered. It will also be interesting to see how this framework can be applied to solve new problems in opportunistic spectrum access in cognitive radio networks, in opportunistic scheduling with delayed/uncertain channel state information, and in restless bandit problems.

Appendix A

Proof of Lemma 2: Initially, by (3) we have $\omega_n(0) = \pi_{n,\mathsf{ON}} \ge \mathsf{P}_{n,01}^{(M)}$ for all n. Suppose the base station switches to channel n at time t, and the last use of channel n ends at slot $(t-k)$ for some $k < t$. In slot $(t-k)$, there are two possible cases:

1) Channel n turns OFF, and as a result the information state on slot t is $\omega_n(t) = \mathsf{P}_{n,01}^{(k)}$. Due to round robin, the other $(M-1)$ channels must have been used for at least one slot before t after slot $(t-k)$, and thus $k \ge M$. By (3) we have $\omega_n(t) = \mathsf{P}_{n,01}^{(k)} \ge \mathsf{P}_{n,01}^{(M)}$.
2) Channel n is ON and transmits a dummy packet. Thus $\omega_n(t) = \mathsf{P}_{n,11}^{(k)}$. By (3) we have $\omega_n(t) = \mathsf{P}_{n,11}^{(k)} \ge \mathsf{P}_{n,01}^{(M)}$. ∎

Appendix B

Proof of Lemma 6: At the beginning of a new round, suppose round robin policy $\mathsf{RR}(\boldsymbol{\phi})$ is selected. We index the $M(\boldsymbol{\phi})$ active channels in $\boldsymbol{\phi}$ as $(n_1, n_2, \dots, n_{M(\boldsymbol{\phi})})$, which is in the decreasing order of the time duration between their last use and the beginning of the current round. In other words, the last use of n_k is earlier than that of $n_{k'}$ only if $k < k'$. Fix an active channel n_k. Then it suffices to show that when this channel is served in the current round, the time duration back to the end of its last service is at least $(M(\boldsymbol{\phi})-1)$ slots (that this channel has information state no worse than $\mathsf{P}_{n_k,01}^{(M(\boldsymbol{\phi}))}$ then follows the same arguments in the proof of Lemma 2).

We partition the active channels in $\boldsymbol{\phi}$ other than n_k into two sets $\mathcal{A} = \{n_1, n_2, \dots, n_{k-1}\}$ and $\mathcal{B} = \{n_{k+1}, n_{k+2}, \dots, n_{M(\boldsymbol{\phi})}\}$. Then the last use of every channel in $\mathcal{B}$ occurs after the last use of n_k, and so channel n_k has been idled for at least $|\mathcal{B}|$ slots at the start of the current round. However, the policy in this round will serve all channels in $\mathcal{A}$ before serving n_k, taking at least one slot per channel, and so we wait at least additional $|\mathcal{A}|$ slots before serving channel n_k. The total time that this channel has been idled is thus at least $|\mathcal{A}| + |\mathcal{B}| = M(\boldsymbol{\phi}) - 1$. ∎

Appendix C

Proof of Theorem 1: Let $Z(t)$ denote the number of times Step 1 of RandRR is executed in $[0, t)$, in which we suppose vector $\boldsymbol{\phi}$ is selected $Z_{\boldsymbol{\phi}}(t)$ times. Define t_i, where $i \in \mathbb{Z}^+$, as the $(i+1)$th time instant a new vector $\boldsymbol{\phi}$ is selected. Assume $t_0 = 0$, and thus the first selection occurs at time 0. It follows that $Z(t_i^-) = i$, $Z(t_i) = i+1$, and the ith round of packet transmissions ends at time t_i^-.

Fix a vector $\boldsymbol{\phi}$. Within the time periods in which policy $\mathsf{RR}(\boldsymbol{\phi})$ is executed, denote by $L_{kn}^{\boldsymbol{\phi}}$ the duration of the kth time the base station stays with channel n. Then the time average throughput that policy $\mathsf{RR}(\boldsymbol{\phi})$ yields on its active channel n over $[0, t_i)$ is

$$\frac{\sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} \left(L_{kn}^{\boldsymbol{\phi}} - 1\right)}{\sum_{\boldsymbol{\phi}\in\Phi} \sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} \sum_{n:\phi_n=1} L_{kn}^{\boldsymbol{\phi}}}. \tag{16}$$

For simplicity, here we focus on discrete time instants $\{t_i\}$ large enough so that $Z_{\boldsymbol{\phi}}(t_i) > 0$ for all $\boldsymbol{\phi} \in \Phi$ (so that the sums in (16) make sense). The generalization to arbitrary time t can be done by incorporating fractional transmission rounds, which are amortized over time. Next, rewrite (16) as

$$\frac{\sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} \sum_{n:\phi_n=1} L_{kn}^{\boldsymbol{\phi}}}{\sum_{\boldsymbol{\phi}\in\Phi} \sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} \sum_{n:\phi_n=1} L_{kn}^{\boldsymbol{\phi}}} \underbrace{\frac{\sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} \left(L_{kn}^{\boldsymbol{\phi}} - 1\right)}{\sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} \sum_{n:\phi_n=1} L_{kn}^{\boldsymbol{\phi}}}}_{(*)}. \quad (17)$$

As $t \to \infty$, the second term $(*)$ of (17) satisfies

$$\begin{aligned}(*) &= \frac{\frac{1}{Z_{\boldsymbol{\phi}}(t_i)} \sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} \left(L_{kn}^{\boldsymbol{\phi}} - 1\right)}{\sum_{n:\phi_n=1} \frac{1}{Z_{\boldsymbol{\phi}}(t_i)} \sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} L_{kn}^{\boldsymbol{\phi}}} \\ &\overset{(a)}{\to} \frac{\mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}} - 1\right]}{\sum_{n:\phi_n=1} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]} \overset{(b)}{=} \eta_n^{\boldsymbol{\phi}},\end{aligned}$$

where (a) is by the Law of Large Numbers (we have shown in Corollary 1 that $L_{kn}^{\boldsymbol{\phi}}$ are i.i.d. for different k) and (b) by (9).

Denote the first term of (17) by $\beta_{\boldsymbol{\phi}}(t_i)$, where we note that $\beta_{\boldsymbol{\phi}}(t_i) \in [0, 1]$ for all $\boldsymbol{\phi} \in \Phi$ and $\sum_{\boldsymbol{\phi}\in\Phi} \beta_{\boldsymbol{\phi}}(t_i) = 1$. We can rewrite $\beta_{\boldsymbol{\phi}}(t_i)$ as

$$\beta_{\boldsymbol{\phi}}(t_i) = \frac{\left[\frac{Z_{\boldsymbol{\phi}}(t_i)}{Z(t_i)}\right] \sum_{n:\phi_n=1} \left[\frac{1}{Z_{\boldsymbol{\phi}}(t_i)} \sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} L_{kn}^{\boldsymbol{\phi}}\right]}{\sum_{\boldsymbol{\phi}\in\Phi} \left[\frac{Z_{\boldsymbol{\phi}}(t_i)}{Z(t_i)}\right] \sum_{n:\phi_n=1} \left[\frac{1}{Z_{\boldsymbol{\phi}}(t_i)} \sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} L_{kn}^{\boldsymbol{\phi}}\right]}.$$

As $t \to \infty$, we have

$$\beta_{\boldsymbol{\phi}} \triangleq \lim_{i\to\infty} \beta_{\boldsymbol{\phi}}(t_i) = \frac{\alpha_{\boldsymbol{\phi}} \sum_{n:\phi_n=1} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]}{\sum_{\boldsymbol{\phi}\in\Phi} \alpha_{\boldsymbol{\phi}} \sum_{n:\phi_n=1} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]}, \quad (18)$$

where by the Law of Large Numbers we have

$$\frac{Z_{\boldsymbol{\phi}}(t_i)}{Z(t_i)} \to \alpha_{\boldsymbol{\phi}}, \quad \frac{1}{Z_{\boldsymbol{\phi}}(t_i)} \sum_{k=1}^{Z_{\boldsymbol{\phi}}(t_i)} L_{kn}^{\boldsymbol{\phi}} \to \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right].$$

From (16)(17)(18), we have shown that the throughput contributed by policy $\mathsf{RR}(\boldsymbol{\phi})$ on its active channel n is $\beta_{\boldsymbol{\phi}}\eta_n^{\boldsymbol{\phi}}$. Consequently, RandRR parameterized by $\{\alpha_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$ supports any data rate vector $\boldsymbol{\lambda}$ that is entrywise dominated by $\boldsymbol{\lambda} \leq \sum_{\boldsymbol{\phi}\in\Phi} \beta_{\boldsymbol{\phi}}\boldsymbol{\eta}^{\boldsymbol{\phi}}$, where $\{\beta_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$ is defined in (18) and $\boldsymbol{\eta}^{\boldsymbol{\phi}}$ in (9).

The above analysis shows that every RandRR policy achieves a boundary point of Λ_{int} defined in Theorem 1. Conversely, the next lemma, proved in Appendix D, shows that every boundary point of Λ_{int} is achievable by some RandRR policy, and the proof is complete.

Lemma 11. *For any probability distribution $\{\beta_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$, there exists another probability distribution $\{\alpha_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$ that solves the linear system*

$$\beta_{\boldsymbol{\phi}} = \frac{\alpha_{\boldsymbol{\phi}} \sum_{n:\phi_n=1} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]}{\sum_{\boldsymbol{\phi}\in\Phi} \alpha_{\boldsymbol{\phi}} \sum_{n:\phi_n=1} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right]}, \quad \textit{for all } \boldsymbol{\phi} \in \Phi. \quad (19)$$

■

Appendix D

Proof of Lemma 11: For any probability distribution $\{\beta_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$, we prove the lemma by inductively constructing the solution $\{\alpha_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi}\in\Phi}$ to (19). The induction is on the cardinality of Φ. Without loss of generality, we index elements in Φ by $\Phi = \{\boldsymbol{\phi}^1, \boldsymbol{\phi}^2, \ldots\}$, where $\boldsymbol{\phi}^k = (\phi_1^k, \ldots, \phi_N^k)$. We define $\chi_k \triangleq \sum_{n:\phi_n^k=1} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}^k}\right]$ and redefine $\beta_{\boldsymbol{\phi}^k} \triangleq \beta_k$ and $\alpha_{\boldsymbol{\phi}^k} \triangleq \alpha_k$. Then we can rewrite (19) as

$$\beta_k = \frac{\alpha_k \chi_k}{\sum_{1\leq k\leq|\Phi|} \alpha_k \chi_k}, \quad \text{for all } k \in \{1, 2, \ldots, |\Phi|\}. \quad (20)$$

We first note that $\Phi = \{\boldsymbol{\phi}^1\}$ is a degenerate case where β_1 and α_1 must both be 1. When $\Phi = \{\boldsymbol{\phi}^1, \boldsymbol{\phi}^2\}$, for any probability distribution $\{\beta_1, \beta_2\}$ with *positive elements*,[10] it is easy to show

$$\alpha_1 = \frac{\chi_2\beta_1}{\chi_1\beta_2 + \chi_2\beta_1}, \quad \alpha_2 = 1 - \alpha_1.$$

Let $\Phi = \{\boldsymbol{\phi}^k : 1 \leq k \leq K\}$ for some $K \geq 2$. Assume that for any probability distribution $\{\beta_k > 0 : 1 \leq k \leq K\}$ we can find $\{\alpha_k : 1 \leq k \leq K\}$ that solves (20).

For the case $\Phi = \{\boldsymbol{\phi}^k : 1 \leq k \leq K+1\}$ and any $\{\beta_k > 0 : 1 \leq k \leq K+1\}$, we construct the solution $\{\alpha_k : 1 \leq k \leq K+1\}$ to (18) as follows. Let $\{\gamma_2, \gamma_3, \ldots, \gamma_{K+1}\}$ be the solution to the linear system

$$\frac{\gamma_k\chi_k}{\sum_{k=2}^{K+1} \gamma_k\chi_k} = \frac{\beta_k}{\sum_{k=2}^{K+1} \beta_k}, \quad 2 \leq k \leq K+1. \quad (21)$$

By the induction assumption, the set $\{\gamma_2, \gamma_3, \ldots, \gamma_{K+1}\}$ exists and satisfies $\gamma_k \in [0, 1]$ for $2 \leq k \leq K+1$ and $\sum_{k=2}^{K+1} \gamma_k = 1$. Define

$$\alpha_1 \triangleq \frac{\beta_1 \sum_{k=2}^{K+1} \gamma_k\chi_k}{\chi_1(1-\beta_1) + \beta_1 \sum_{k=2}^{K+1} \gamma_k\chi_k} \quad (22)$$

$$\alpha_k \triangleq (1-\alpha_1)\gamma_k, \quad 2 \leq k \leq K+1. \quad (23)$$

It remains to show (22) and (23) are the desired solution. It is easy to observe that $\alpha_k \in [0, 1]$ for $1 \leq k \leq K+1$, and

$$\sum_{k=1}^{K+1} \alpha_k = \alpha_1 + (1-\alpha_1) \sum_{k=2}^{K+1} \gamma_k = \alpha_1 + (1-\alpha_1) = 1.$$

By rearranging terms in (22) and using (23), we have

$$\beta_1 = \frac{\alpha_1\chi_1}{\alpha_1\chi_1 + \sum_{k=2}^{K+1}(1-\alpha_1)\gamma_k\chi_k} = \frac{\alpha_1\chi_1}{\sum_{k=1}^{K+1} \alpha_k\chi_k}. \quad (24)$$

For $2 \leq k \leq K+1$,

$$\begin{aligned}\frac{\alpha_k\chi_k}{\sum_{k=1}^{K+1} \alpha_k\chi_k} &= \left[\frac{\alpha_k\chi_k}{\sum_{k=2}^{K+1} \alpha_k\chi_k}\right] \left[\frac{\sum_{k=2}^{K+1} \alpha_k\chi_k}{\sum_{k=1}^{K+1} \alpha_k\chi_k}\right] \\ &\overset{(a)}{=} \left[\frac{(1-\alpha_1)\gamma_k\chi_k}{\sum_{k=2}^{K+1}(1-\alpha_1)\gamma_k\chi_k}\right] \left[1 - \frac{\alpha_1\chi_1}{\sum_{k=1}^{K+1} \alpha_k\chi_k}\right] \\ &\overset{(b)}{=} \left[\frac{\gamma_k\chi_k}{\sum_{k=2}^{K+1} \gamma_k\chi_k}\right] (1-\beta_1)\end{aligned}$$

[10] If one element of $\{\beta_1, \beta_2\}$ is zero, say $\beta_2 = 0$, we can show necessarily $\alpha_2 = 0$ and it degenerates to the one-policy case $\Phi = \{\boldsymbol{\phi}^1\}$. Such degeneration happens in general cases. Thus in the rest of the proof we will only consider probability distributions that only have positive elements.

$$\overset{(c)}{=} \left(\frac{\beta_k}{\sum_{k=2}^{K+1} \beta_k} \right) (1-\beta_1) \overset{(d)}{=} \beta_k,$$

where (a) is by plugging in (23), (b) uses (24), (c) uses (21), and (d) is by $\sum_{k=1}^{K+1} \beta_k = 1$. The proof is complete. ■

Appendix E

Proof of Lemma 7: Let $\mathcal{N}_1(T) \subseteq \{0, 1, \ldots, T-1\}$ be the subset of time instants in which $Y(t) = 1$. Note that $\sum_{t=0}^{T-1} Y(t) = |\mathcal{N}_1(T)|$. For each $t \in \mathcal{N}_1(T)$, let $1_{[1\to 0]}(t)$ be an indicator function which is 1 if $Y(t)$ transits from 1 to 0 at time t, and 0 otherwise. We define $\mathcal{N}_0(T)$ and $1_{[0\to 1]}(t)$ similarly.

In $\{0, 1, \ldots, T-1\}$, since state transitions of $\{Y(t)\}$ from 1 to 0 and from 0 to 1 differ by at most 1, we have

$$\left| \sum_{t\in\mathcal{N}_1(T)} 1_{[1\to 0]}(t) - \sum_{t\in\mathcal{N}_0(T)} 1_{[0\to 1]}(t) \right| \leq 1, \tag{25}$$

which is true for all T. Dividing (25) by T, we get

$$\left| \frac{1}{T} \sum_{t\in\mathcal{N}_1(T)} 1_{[1\to 0]}(t) - \frac{1}{T} \sum_{t\in\mathcal{N}_0(T)} 1_{[0\to 1]}(t) \right| \leq \frac{1}{T}. \tag{26}$$

Consider the subsequence $\{T_k\}$ such that

$$\lim_{k\to\infty} \frac{1}{T_k} \sum_{t=0}^{T_k-1} Y(t) = \pi_Y(1) = \lim_{k\to\infty} \frac{|\mathcal{N}_1(T_k)|}{T_k}. \tag{27}$$

Note that $\{T_k\}$ exists because $(1/T)\sum_{t=0}^{T-1} Y(t)$ is a bounded sequence indexed by integers T. Moreover, there exists a subsequence $\{T_n\}$ of $\{T_k\}$ so that each of the two averages in (26) has a limit point with respect to $\{T_n\}$, because they are bounded sequences, too. In the rest of the proof we will work on $\{T_n\}$, but we drop subscript n for notational simplicity. Passing $T \to \infty$, we get from (26) that

$$\begin{aligned} &\underbrace{\left(\lim_{T\to\infty} \frac{|\mathcal{N}_1(T)|}{T} \right)}_{\overset{(a)}{=} \pi_Y(1)} \underbrace{\left(\lim_{T\to\infty} \frac{1}{|\mathcal{N}_1(T)|} \sum_{t\in\mathcal{N}_1(T)} 1_{[1\to 0]}(t) \right)}_{\triangleq \beta} \\ &= \underbrace{\left(\lim_{T\to\infty} \frac{|\mathcal{N}_0(T)|}{T} \right)}_{\overset{(b)}{=} 1-\pi_Y(1)} \underbrace{\left(\lim_{T\to\infty} \frac{1}{|\mathcal{N}_0(T)|} \sum_{t\in\mathcal{N}_0(T)} 1_{[0\to 1]}(t) \right)}_{\triangleq \gamma}, \end{aligned} \tag{28}$$

where (a) is by (27) and (b) is by $|\mathcal{N}_1(T)| + |\mathcal{N}_0(T)| = T$. From (28) we get

$$\pi_Y(1) = \frac{\gamma}{\beta + \gamma}.$$

The next lemma, proved in Appendix F, helps to show $\gamma \leq \mathsf{P}_{01}$.

Lemma 12 (Stochastic coupling of random binary sequences)**.** *Let $\{I_n\}_{n=1}^{\infty}$ be an infinite sequence of binary random variables. Suppose for all $n \in \{1, 2, \ldots\}$ we have*

$$\Pr\left[I_n = 1 \mid I_1 = i_1, \ldots, I_{n-1} = i_{n-1}\right] \leq \mathsf{P}_{01} \tag{29}$$

for all possible values of $i_1, \ldots, i_{n-1}$. Then we can construct a new sequence $\{\hat{I}_n\}_{n=1}^{\infty}$ of binary random variables that are i.i.d. with $\Pr\left[\hat{I}_n = 1\right] = \mathsf{P}_{01}$ for all n and satisfy $\hat{I}_n \geq I_n$ for all n. Consequently, we have

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_n \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \hat{I}_n = \mathsf{P}_{01}.$$

To use Lemma 12 to prove $\gamma \leq \mathsf{P}_{01}$, let t_n denote the nth time $Y(t) = 0$ and let $I_n = 1_{[0\to 1]}(t_n)$. For simplicity assume $\{t_n\}$ is an infinite sequence so that state 0 is visited infinitely often in $\{Y(t)\}$. By the assumption that $\mathsf{Q}_{01}(t) \leq \mathsf{P}_{01}$ for all t, we know (29) holds. Therefore by Lemma 12 we have

$$\gamma \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} 1_{[0\to 1]}(t_n) \leq \mathsf{P}_{01}.$$

Similarly as Lemma 12, we can show $\beta \geq \mathsf{P}_{10}$ by stochastic coupling. Therefore

$$\pi_Y(1) = \frac{\gamma}{\beta + \gamma} \leq \frac{\gamma}{\mathsf{P}_{10} + \gamma} \leq \frac{\mathsf{P}_{01}}{\mathsf{P}_{01} + \mathsf{P}_{10}} = \pi_X(1).$$

■

Appendix F

Proof of Lemma 12: For simplicity, we assume

$$\Pr\left[I_n = 0 \mid I_1 = i_1, \ldots, I_{n-1} = i_{n-1}\right] > 0$$

for all n and all possible values of $i_1, \ldots, i_{n-1}$. For each $n \in \{1, 2, \ldots\}$, define $\hat{I}_n$ as follows: If $I_n = 1$, define $\hat{I}_n = 1$. If $I_n = 0$, observe the *history* $I_1^{n-1} \triangleq (I_1, \ldots, I_{n-1})$ and independently choose $\hat{I}_n$ as follows:

$$\hat{I}_n = \begin{cases} 1 & \text{with prob. } \frac{\mathsf{P}_{01} - \Pr\left[I_n = 1 \mid I_1^{n-1}\right]}{\Pr\left[I_n = 0 \mid I_1^{n-1}\right]} \\ 0 & \text{with prob. } 1 - \frac{\mathsf{P}_{01} - \Pr\left[I_n = 1 \mid I_1^{n-1}\right]}{\Pr\left[I_n = 0 \mid I_1^{n-1}\right]}. \end{cases} \tag{30}$$

The probabilities in (30) are well-defined because $\mathsf{P}_{01} \geq \Pr\left[I_n = 1 \mid I_1^{n-1}\right]$ by (29), and

$$\mathsf{P}_{01} \leq 1 = \Pr\left[I_n = 1 \mid I_1^{n-1}\right] + \Pr\left[I_n = 0 \mid I_1^{n-1}\right]$$

and therefore

$$\mathsf{P}_{01} - \Pr\left[I_n = 1 \mid I_1^{n-1}\right] \leq \Pr\left[I_n = 0 \mid I_1^{n-1}\right].$$

With the above definition of $\hat{I}_n$, we have $\hat{I}_n = 1$ whenever $I_n = 1$. Therefore $\hat{I}_n \geq I_n$ for all n. Further, for any n and any binary vector $i_1^{n-1} \triangleq (i_1, \ldots, i_{n-1})$, we have

$$\begin{aligned} &\Pr\left[\hat{I}_n = 1 \mid I_1^{n-1} = i_1^{n-1}\right] \\ &= \Pr\left[I_n = 1 \mid I_1^{n-1} = i_1^{n-1}\right] + \Pr\left[I_n = 0 \mid I_1^{n-1} = i_1^{n-1}\right] \\ &\quad \times \frac{\mathsf{P}_{01} - \Pr\left[I_n = 1 \mid I_1^{n-1} = i_1^{n-1}\right]}{\Pr\left[I_n = 0 \mid I_1^{n-1} = i_1^{n-1}\right]} = \mathsf{P}_{01}. \end{aligned} \tag{31}$$

Therefore, for all n we have

$$\Pr\left[\hat{I}_n = 1\right]$$

$$= \sum_{i_1^{n-1}} \Pr\left[\hat{I}_n = 1 \mid I_1^{n-1} = i_1^{n-1}\right] \Pr\left[I_1^{n-1} = i_1^{n-1}\right] = \mathsf{P}_{01},$$

and thus the $\hat{I}_n$ variables are identically distributed. It remains to prove that they are independent.

Suppose components in $\hat{I}_1^n \triangleq (\hat{I}_1, \dots, \hat{I}_n)$ are independent. We prove that components in $\hat{I}_1^{n+1} = (\hat{I}_1, \dots, \hat{I}_{n+1})$ are also independent. For any binary vector $\hat{i}_1^{n+1} \triangleq (\hat{i}_1, \dots, \hat{i}_{n+1})$, since

$$\begin{aligned}
&\Pr\left[\hat{I}_1^{n+1} = \hat{i}_1^{n+1}\right] \\
&= \Pr\left[\hat{I}_{n+1} = \hat{i}_{n+1} \mid \hat{I}_1^n = \hat{i}_1^n\right] \Pr\left[\hat{I}_1^n = \hat{i}_1^n\right] \\
&= \Pr\left[\hat{I}_{n+1} = \hat{i}_{n+1} \mid \hat{I}_1^n = \hat{i}_1^n\right] \prod_{k=1}^{n} \Pr\left[\hat{I}_k = \hat{i}_k\right],
\end{aligned}$$

it suffices to show

$$\Pr\left[\hat{I}_{n+1} = 1 \mid \hat{I}_1^n = \hat{i}_1^n\right] = \Pr\left[\hat{I}_{n+1} = 1\right] = \mathsf{P}_{01}.$$

Indeed,

$$\begin{aligned}
&\Pr\left[\hat{I}_{n+1} = 1 \mid \hat{I}_1^n = \hat{i}_1^n\right] \\
&= \sum_{i_1^n} \Pr\left[\hat{I}_{n+1} = 1 \mid I_1^n = i_1^n, \hat{I}_1^n = \hat{i}_1^n\right] \\
&\qquad \times \Pr\left[I_1^n = i_1^n \mid \hat{I}_1^n = \hat{i}_1^n\right] \\
&= \sum_{i_1^n} \Pr\left[\hat{I}_{n+1} = 1 \mid I_1^n = i_1^n\right] \Pr\left[I_1^n = i_1^n \mid \hat{I}_1^n = \hat{i}_1^n\right] \\
&\overset{(a)}{=} \sum_{i_1^n} \mathsf{P}_{01} \Pr\left[I_1^n = i_1^n \mid \hat{I}_1^n = \hat{i}_1^n\right] = \mathsf{P}_{01},
\end{aligned}$$

where (a) is by (31), and the proof is complete. ■

APPENDIX G

Proof of Lemma 9: By definition of $d(\boldsymbol{v})$, there exists a nonempty subset $A \subseteq \Phi_{d(\boldsymbol{v})}$, and for every $\boldsymbol{\phi} \in A$ a positive real number $\hat{\beta}_{\boldsymbol{\phi}} > 0$, such that $\boldsymbol{v} = \sum_{\boldsymbol{\phi} \in A} \hat{\beta}_{\boldsymbol{\phi}} \boldsymbol{\phi}$. For each $\boldsymbol{\phi} \in A$, we have $M(\boldsymbol{\phi}) = d(\boldsymbol{v})$ and thus $c_{M(\boldsymbol{\phi})} = c_{d(\boldsymbol{v})}$. Define

$$\beta_{\boldsymbol{\phi}} \triangleq \frac{\hat{\beta}_{\boldsymbol{\phi}}}{\sum_{\boldsymbol{\phi} \in A} \hat{\beta}_{\boldsymbol{\phi}}}$$

for each $\boldsymbol{\phi} \in A$ and $\{\beta_{\boldsymbol{\phi}}\}_{\boldsymbol{\phi} \in A}$ is a probability distribution. Consider a RandRR policy that in every round selects $\boldsymbol{\phi} \in A$ with probability $\beta_{\boldsymbol{\phi}}$. By Lemma 4, this RandRR policy achieves throughput vector $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_N)$ that satisfies

$$\begin{aligned}
\boldsymbol{\lambda} = \sum_{\boldsymbol{\phi} \in A} \beta_{\boldsymbol{\phi}} \frac{c_{M(\boldsymbol{\phi})}}{M(\boldsymbol{\phi})} \boldsymbol{\phi} &= \frac{c_{d(\boldsymbol{v})}}{d(\boldsymbol{v})} \sum_{\boldsymbol{\phi} \in A} \frac{\hat{\beta}_{\boldsymbol{\phi}}}{\sum_{\boldsymbol{\phi} \in A} \hat{\beta}_{\boldsymbol{\phi}}} \boldsymbol{\phi} \\
&= \frac{c_{d(\boldsymbol{v})}}{d(\boldsymbol{v}) \sum_{\boldsymbol{\phi} \in A} \hat{\beta}_{\boldsymbol{\phi}}} \sum_{\boldsymbol{\phi} \in A} \hat{\beta}_{\boldsymbol{\phi}} \boldsymbol{\phi} \\
&= \left(\frac{c_{d(\boldsymbol{v})}}{d(\boldsymbol{v}) \sum_{\boldsymbol{\phi} \in A} \hat{\beta}_{\boldsymbol{\phi}}} \right) \boldsymbol{v},
\end{aligned}$$

which is in the direction of $\boldsymbol{v}$. In addition, the sum throughput

$$\sum_{n=1}^{N} \lambda_n = \sum_{\boldsymbol{\phi} \in A} \beta_{\boldsymbol{\phi}} \frac{c_{M(\boldsymbol{\phi})}}{M(\boldsymbol{\phi})} \left(\sum_{n=1}^{N} \phi_n \right) = \sum_{\boldsymbol{\phi} \in A} \beta_{\boldsymbol{\phi}} c_{M(\boldsymbol{\phi})} = c_{d(\boldsymbol{v})}$$

is achieved. ■

APPENDIX H

Proof of Theorem 3: (**A Related** RandRR **Policy**) For each randomized round robin policy RandRR, it is useful to consider a *renewal reward process* where renewal epochs are defined as time instants at which RandRR starts a new round of transmission.[11] Let T denote the renewal period. We say one unit of reward is earned by a user if RandRR serves a packet to that user. Let R_n denote the sum reward earned by user n in one renewal period T, representing the number of successful transmissions user n receives in one round of scheduling. Conditioning on the round robin policy $\mathsf{RR}(\boldsymbol{\phi})$ chosen by RandRR for the current round of transmission, we have from Corollary 1:

$$\mathbb{E}[T] = \sum_{\boldsymbol{\phi} \in \Phi} \alpha_{\boldsymbol{\phi}} \mathbb{E}[T \mid \mathsf{RR}(\boldsymbol{\phi})] \tag{32}$$

$$\mathbb{E}[T \mid \mathsf{RR}(\boldsymbol{\phi})] = \sum_{n : \phi_n = 1} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}}\right], \tag{33}$$

and for all $n \in \{1, 2, \dots, N\}$,

$$\mathbb{E}[R_n] = \sum_{\boldsymbol{\phi} \in \Phi} \alpha_{\boldsymbol{\phi}} \mathbb{E}[R_n \mid \mathsf{RR}(\boldsymbol{\phi})] \tag{34}$$

$$\mathbb{E}[R_n \mid \mathsf{RR}(\boldsymbol{\phi})] = \begin{cases} \mathbb{E}\left[L_{1n}^{\boldsymbol{\phi}} - 1\right] & \text{if } \phi_n = 1 \\ 0 & \text{if } \phi_n = 0. \end{cases} \tag{35}$$

Consider the round robin policy $\mathsf{RR}((1, 1, \dots, 1))$ that serves all N channels in one round. We define T_{max} as its renewal period. From Corollary 1, we know $\mathbb{E}[T_{\text{max}}] < \infty$ and $\mathbb{E}\left[(T_{\text{max}})^2\right] < \infty$. Further, for any RandRR, including using a $\mathsf{RR}(\boldsymbol{\phi})$ policy in every round as special cases, we can show that T_{max} is stochastically larger than the renewal period T, and $(T_{\text{max}})^2$ is stochastically larger than T^2. It follows that

$$\mathbb{E}[T] \le \mathbb{E}[T_{\text{max}}], \ \mathbb{E}\left[T^2\right] \le \mathbb{E}\left[(T_{\text{max}})^2\right]. \tag{36}$$

We have denoted by RandRR* (in the discussion after (13)) the randomized round robin policy that achieves a service rate vector strictly larger than the target arrival rate vector $\boldsymbol{\lambda}$ entrywise. Let T^* denote the renewal period of RandRR*, and R_n^* the sum reward (the number of successful transmissions) received by user n over the renewal period T^*. Then we have

$$\begin{aligned}
\frac{\mathbb{E}[R_n^*]}{\mathbb{E}[T^*]} &\overset{(a)}{=} \frac{\sum_{\boldsymbol{\phi} \in \Phi} \alpha_{\boldsymbol{\phi}} \mathbb{E}[R_n^* \mid \mathsf{RR}(\boldsymbol{\phi})]}{\sum_{\boldsymbol{\phi} \in \Phi} \alpha_{\boldsymbol{\phi}} \mathbb{E}[T^* \mid \mathsf{RR}(\boldsymbol{\phi})]} \\
&\overset{(b)}{=} \sum_{\boldsymbol{\phi} \in \Phi} \left(\frac{\alpha_{\boldsymbol{\phi}}}{\sum_{\boldsymbol{\phi} \in \Phi} \alpha_{\boldsymbol{\phi}} \mathbb{E}[T^* \mid \mathsf{RR}(\boldsymbol{\phi})]} \right) \mathbb{E}[R_n^* \mid \mathsf{RR}(\boldsymbol{\phi})]
\end{aligned}$$

[11] We note that the renewal reward process is defined solely with respect to RandRR, and is only used to facilitate our analysis. At these renewal epochs, the state of the network, including the current queue state $\boldsymbol{Q}(t)$, does *not* necessarily renew itself.

$$= \sum_{\phi\in\Phi} \underbrace{\frac{\alpha_{\phi}\,\mathbb{E}\left[T^* \mid \mathsf{RR}(\phi)\right]}{\sum_{\phi\in\Phi}\alpha_{\phi}\,\mathbb{E}\left[T^* \mid \mathsf{RR}(\phi)\right]}}_{(c)=\beta_n} \underbrace{\frac{\mathbb{E}\left[R_n^* \mid \mathsf{RR}(\phi)\right]}{\mathbb{E}\left[T^* \mid \mathsf{RR}(\phi)\right]}}_{(d)=\eta_n^{\phi}}$$

$$= \sum_{\phi\in\Phi} \beta_{\phi}\eta_n^{\phi} \overset{(e)}{>} \lambda_n + \epsilon, \tag{37}$$

where (a) is by (32)(34), (b) is by rearranging terms, (c) is by plugging (33) into (19), (d) is by plugging (33) and (35) into (9) in Section IV-B, and (e) is by (13). From (37) we get

$$\mathbb{E}\left[R_n^*\right] > (\lambda_n + \epsilon)\mathbb{E}\left[T^*\right], \quad \text{for all } n \in \{1, \dots, N\}. \tag{38}$$

(**Lyapunov Drift**) From (12), in a frame of size T (which is possibly random), we can show that for all n

$$Q_n(t+T) \le \max\left[Q_n(t) - \sum_{\tau=0}^{T-1}\mu_n(t+\tau), 0\right] + \sum_{\tau=0}^{T-1} a_n(t+\tau). \tag{39}$$

We define a Lyapunov function $L(\boldsymbol{Q}(t)) \triangleq (1/2)\sum_{n=1}^{N} Q_n^2(t)$ and the T-slot Lyapunov drift

$$\Delta_T(\boldsymbol{Q}(t)) \triangleq \mathbb{E}\left[L(\boldsymbol{Q}(t+T) - L(\boldsymbol{Q}(t)) \mid \boldsymbol{Q}(t)\right],$$

where in the last term the expectation is with respect to the randomness of the whole network in frame T, including the randomness of T. By taking square of (39) and then conditional expectation on $\boldsymbol{Q}(t)$, we can show

$$\Delta_T(\boldsymbol{Q}(t)) \le \frac{1}{2}N(1 + A_{\max}^2)\mathbb{E}\left[T^2 \mid \boldsymbol{Q}(t)\right] - \mathbb{E}\left[\sum_{n=1}^{N} Q_n(t)\left[\sum_{\tau=0}^{T-1}(\mu_n(t+\tau) - a_n(t+\tau))\right] \mid \boldsymbol{Q}(t)\right]. \tag{40}$$

Define $f(\boldsymbol{Q}(t), \theta)$ as the last term of (40), where θ represents a scheduling policy that controls the service rates $\mu_n(t+\tau)$ and the frame size T. In the following analysis, we only consider θ in the class of RandRR policies, and the frame size T is the renewal period of a RandRR policy. By (36), the second term of (40) is less than or equal to the constant $B_1 \triangleq (1/2)N(1 + A_{\max}^2)\mathbb{E}\left[(T_{\max})^2\right] < \infty$. It follows that

$$\Delta_T(\boldsymbol{Q}(t)) \le B_1 - f(\boldsymbol{Q}(t), \theta). \tag{41}$$

In $f(\boldsymbol{Q}(t), \theta)$, it is useful to consider $\theta = \mathsf{RandRR}^*$ and T is the renewal period T^* of RandRR^*. Assume t is the beginning of a renewal period. For each $n \in \{1, 2, \dots, N\}$, because R_n^* is the number of successful transmissions user n receives in the renewal period T^*, we have

$$\mathbb{E}\left[\sum_{\tau=0}^{T^*-1}\mu_n(t+\tau) \mid \boldsymbol{Q}(t)\right] = \mathbb{E}\left[R_n^*\right].$$

Combining with (38), we get

$$\mathbb{E}\left[\sum_{\tau=0}^{T^*-1}\mu_n(t+\tau) \mid \boldsymbol{Q}(t)\right] > (\lambda_n + \epsilon)\mathbb{E}\left[T^*\right]. \tag{42}$$

By the assumption that packet arrivals are i.i.d. over slots and independent of the current queue backlogs, we have for all n

$$\mathbb{E}\left[\sum_{\tau=0}^{T^*-1} a_n(t+\tau) \mid \boldsymbol{Q}(t)\right] = \lambda_n\mathbb{E}\left[T^*\right]. \tag{43}$$

Plugging (42) and (43) into $f(\boldsymbol{Q}(t), \mathsf{RandRR}^*)$, we get

$$f(\boldsymbol{Q}(t), \mathsf{RandRR}^*) \ge \epsilon\mathbb{E}\left[T^*\right]\sum_{n=1}^{N} Q_n(t). \tag{44}$$

It is also useful to consider θ as a round robin policy $\mathsf{RR}(\phi)$ for some $\phi \in \Phi$. In this case frame size T is the renewal period T^{ϕ} of $\mathsf{RR}(\phi)$ (note that $\mathsf{RR}(\phi)$ is a special case of RandRR). From Corollary 1, we have

$$\mathbb{E}\left[T^{\phi} \mid \boldsymbol{Q}(t)\right] = \mathbb{E}\left[T^{\phi}\right] = \sum_{n:\phi_n=1}\mathbb{E}\left[L_{1n}^{\phi}\right], \tag{45}$$

where $\mathbb{E}\left[L_{1n}^{\phi}\right]$ can be expanded by (8). Let t be the beginning of a transmission round. If channel n is active, we have

$$\mathbb{E}\left[\sum_{\tau=0}^{T^{\phi}-1}\mu_n(t+\tau) \mid \boldsymbol{Q}(t)\right] = \mathbb{E}\left[L_{1n}^{\phi}\right] - 1,$$

and 0 otherwise. It follows that

$$f(\boldsymbol{Q}(t), \mathsf{RR}(\phi)) = \left(\sum_{n:\phi_n=1} Q_n(t)\mathbb{E}\left[L_{1n}^{\phi} - 1\right]\right) - \mathbb{E}\left[T^{\phi}\right]\sum_{n=1}^{N} Q_n(t)\lambda_n \overset{(a)}{=} \sum_{n:\phi_n=1}\left[Q_n(t)\mathbb{E}\left[L_{1n}^{\phi} - 1\right] - \mathbb{E}\left[L_{1n}^{\phi}\right]\sum_{n=1}^{N} Q_n(t)\lambda_n\right], \tag{46}$$

where (a) is by (45) and rearranging terms.

(**Design of** QRR) Given the current queue backlogs $\boldsymbol{Q}(t)$, we are interested in the policy that maximizes $f(\boldsymbol{Q}(t), \theta)$ over all RandRR policies in one round of transmission. Although the RandRR policy space is uncountably large and thus searching for the optimal solution could be difficult, next we show that the optimal solution is a round robin policy $\mathsf{RR}(\phi)$ for some $\phi \in \Phi$ and can be found by maximizing $f(\boldsymbol{Q}(t), \mathsf{RR}(\phi))$ in (46) over $\phi \in \Phi$. To see this, we denote by $\phi(t)$ the binary vector associated with the $\mathsf{RR}(\phi)$ policy that maximizes $f(\boldsymbol{Q}(t), \mathsf{RR}(\phi))$ over $\phi \in \Phi$, and we have

$$f(\boldsymbol{Q}(t), \mathsf{RR}(\phi(t))) \ge f(\boldsymbol{Q}(t), \mathsf{RR}(\phi)), \text{ for all } \phi \in \Phi. \tag{47}$$

For any RandRR policy, conditioning on the policy $\mathsf{RR}(\phi)$ chosen for the current round of scheduling, we have

$$f(\boldsymbol{Q}(t), \mathsf{RandRR}) = \sum_{\phi\in\Phi}\alpha_{\phi} f(\boldsymbol{Q}(t), \mathsf{RR}(\phi)), \tag{48}$$

where $\{\alpha_{\phi}\}_{\phi\in\Phi}$ is the probability distribution associated with RandRR. By (47)(48), for any RandRR we get

$$\begin{aligned} f(\boldsymbol{Q}(t), \mathsf{RR}(\phi(t))) &\ge \sum_{\phi\in\Phi}\alpha_{\phi} f(\boldsymbol{Q}(t), \mathsf{RR}(\phi)) \\ &= f(\boldsymbol{Q}(t), \mathsf{RandRR}). \end{aligned} \tag{49}$$

We note that as long as the queue backlog vector $\boldsymbol{Q}(t)$ is not identically zero and the arrival rate vector $\boldsymbol{\lambda}$ is strictly within the inner capacity bound Λ_{int}, we get

$$\begin{aligned}\max_{\boldsymbol{\phi}\in\Phi} f(\boldsymbol{Q}(t),\mathsf{RR}(\boldsymbol{\phi})) &\overset{(a)}{=} f(\boldsymbol{Q}(t),\mathsf{RR}(\boldsymbol{\phi}(t))) \\ &\overset{(b)}{\geq} f(\boldsymbol{Q}(t),\mathsf{RandRR}^*) \overset{(c)}{>} 0,\end{aligned} \tag{50}$$

where (a) is from the definition of $\boldsymbol{\phi}(t)$, (b) from (49), and (c) from (44).

The policy QRR is designed to be a frame-based algorithm where at the beginning of each round we observe the current queue backlog vector $\boldsymbol{Q}(t)$, find the binary vector $\boldsymbol{\phi}(t)$ whose associated round robin policy $\mathsf{RR}(\boldsymbol{\phi}(t))$ maximizes $f(\boldsymbol{Q}(t),\mathsf{RandRR})$ over RandRR policies, and execute $\mathsf{RR}(\boldsymbol{\phi}(t))$ for one round of transmission. We emphasize that in every transmission round of QRR, active channels are served in the order that the least recently used channel is served first, and the ordering may change from one round to another.

(**Stability Analysis**) Again, policy QRR comprises of a sequence of transmission rounds, where in each round QRR finds and executes policy $\mathsf{RR}(\boldsymbol{\phi}(t))$ for one round, and different $\boldsymbol{\phi}(t)$ may be used in different rounds. In the kth round, let T_k^{QRR} denote its time duration. Define $t_k = \sum_{i=1}^k T_i^{\mathsf{QRR}}$ for all $k\in\mathbb{N}$ and note that $t_k - t_{k-1} = T_k^{\mathsf{QRR}}$. Let $t_0 = 0$. Then for each $k\in\mathbb{N}$, from (41) we have

$$\begin{aligned}\Delta_{T_k^{\mathsf{QRR}}}(\boldsymbol{Q}(t_{k-1})) &\overset{(a)}{\leq} B_1 - f(\boldsymbol{Q}(t_{k-1}),\mathsf{QRR}) \\ &\overset{(b)}{\leq} B_1 - f(\boldsymbol{Q}(t_{k-1}),\mathsf{RandRR}^*) \\ &\overset{(c)}{\leq} B_1 - \epsilon\,\mathbb{E}\left[T^*\right]\sum_{n=1}^N Q_n(t_{k-1}),\end{aligned} \tag{51}$$

where (a) is by (41), (b) is because QRR is the maximizer of $f(\boldsymbol{Q}(t_{k-1}),\mathsf{RandRR})$ over all RandRR policies, and (c) is by (44). By taking expectation over $\boldsymbol{Q}(t_{k-1})$ in (51) and noting that $\mathbb{E}\left[T^*\right]\geq 1$, for all $k\in\mathbb{N}$ we get

$$\begin{aligned}&\mathbb{E}\left[L(\boldsymbol{Q}(t_k))\right] - \mathbb{E}\left[L(\boldsymbol{Q}(t_{k-1}))\right] \\ &\leq B_1 - \epsilon\,\mathbb{E}\left[T^*\right]\sum_{n=1}^N \mathbb{E}\left[Q_n(t_{k-1})\right] \\ &\leq B_1 - \epsilon\sum_{n=1}^N \mathbb{E}\left[Q_n(t_{k-1})\right].\end{aligned} \tag{52}$$

Summing (52) over $k\in\{1,2,\ldots,K\}$, we have

$$\begin{aligned}&\mathbb{E}\left[L(\boldsymbol{Q}(t_K))\right] - \mathbb{E}\left[L(\boldsymbol{Q}(t_0))\right] \\ &\quad\leq KB_1 - \epsilon\sum_{k=1}^K\sum_{n=1}^N \mathbb{E}\left[Q_n(t_{k-1})\right].\end{aligned}$$

Since $\boldsymbol{Q}(t_K)\geq\boldsymbol{0}$ entrywise and by assumption $\boldsymbol{Q}(t_0) = \boldsymbol{Q}(0) = \boldsymbol{0}$, we have

$$\epsilon\sum_{k=1}^K\sum_{n=1}^N \mathbb{E}\left[Q_n(t_{k-1})\right] \leq KB_1. \tag{53}$$

Dividing (53) by ϵK and passing $K\to\infty$, we get

$$\limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^K\sum_{n=1}^N \mathbb{E}\left[Q_n(t_{k-1})\right] \leq \frac{B_1}{\epsilon} < \infty. \tag{54}$$

Equation (54) shows that the network is stable when sampled at renewal time instants $\{t_k\}$. Then that it is also stable when sampled over all time follows because T_k^{QRR}, the renewal period of the $\mathsf{RR}(\boldsymbol{\phi})$ policy chosen in the kth round of QRR, has finite first and second moments for all k (see (36)), and in every slot the number of packet arrivals to a user is bounded. These details are provided in Lemma 13, which is proved in Appendix I.

Lemma 13. *Given that*

$$\mathbb{E}\left[T_k^{\mathsf{QRR}}\right]\leq\mathbb{E}\left[T_{max}\right],\quad \mathbb{E}\left[(T_k^{\mathsf{QRR}})^2\right]\leq\mathbb{E}\left[(T_{max})^2\right] \tag{55}$$

for all $k\in\{1,2,\ldots\}$, packets arrivals to a user is bounded by A_{max} in every slot, and the network sampled at renewal epochs $\{t_k\}$ is stable from (54)*, we have*

$$\limsup_{K\to\infty}\frac{1}{t_K}\sum_{\tau=0}^{t_K-1}\sum_{n=1}^N \mathbb{E}\left[Q_n(\tau)\right] < \infty.$$

■

APPENDIX I

Proof of Lemma 13: In $[t_{k-1}, t_k)$, it is easy to see for all $n\in\{1,\ldots,N\}$

$$Q_n(t_{k-1}+\tau)\leq Q_n(t_{k-1}) + \tau A_{\max},\quad 0\leq\tau<T_k^{\mathsf{QRR}}. \tag{56}$$

Summing (56) over $\tau\in\{0,1,\ldots,T_k^{\mathsf{QRR}}-1\}$, we get

$$\sum_{\tau=0}^{T_k^{\mathsf{QRR}}-1} Q_n(t_{k-1}+\tau)\leq T_k^{\mathsf{QRR}} Q_n(t_{k-1}) + (T_k^{\mathsf{QRR}})^2 A_{\max}/2. \tag{57}$$

Summing (57) over $k\in\{1,2,\ldots,K\}$ and noting that $t_K = \sum_{k=1}^K T_k^{\mathsf{QRR}}$, we have

$$\begin{aligned}\sum_{\tau=0}^{t_K-1} Q_n(\tau) &= \sum_{k=1}^K\sum_{\tau=0}^{T_k^{\mathsf{QRR}}-1} Q_n(t_{k-1}+\tau) \\ &\overset{(a)}{\leq}\sum_{k=1}^K\left[T_k^{\mathsf{QRR}} Q_n(t_{k-1}) + (T_k^{\mathsf{QRR}})^2 A_{\max}/2\right],\end{aligned} \tag{58}$$

where (a) is by (57). Taking expectation of (58) and dividing it by t_K, we have

$$\begin{aligned}\frac{1}{t_K}\sum_{\tau=0}^{t_K-1}\mathbb{E}\left[Q_n(\tau)\right] &\overset{(a)}{\leq}\frac{1}{K}\sum_{\tau=0}^{t_K-1}\mathbb{E}\left[Q_n(\tau)\right] \\ &\overset{(b)}{\leq}\frac{1}{K}\sum_{k=1}^K\mathbb{E}\left[T_k^{\mathsf{QRR}} Q_n(t_{k-1}) + \frac{(T_k^{\mathsf{QRR}})^2 A_{\max}}{2}\right],\end{aligned} \tag{59}$$

where (a) follows $t_K\geq K$ and (b) is by (58). Next, we have

$$\begin{aligned}\mathbb{E}\left[T_k^{\mathsf{QRR}} Q_n(t_{k-1})\right] &= \mathbb{E}\left[\mathbb{E}\left[T_k^{\mathsf{QRR}} Q_n(t_{k-1})\mid Q_n(t_{k-1})\right]\right] \\ &\overset{(a)}{\leq}\mathbb{E}\left[\mathbb{E}\left[T_{\max} Q_n(t_{k-1})\mid Q_n(t_{k-1})\right]\right]\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left[\mathbb{E}\left[T_{\max}\right] Q_n(t_{k-1})\right] \\
&= \mathbb{E}\left[T_{\max}\right] \mathbb{E}\left[Q_n(t_{k-1})\right], \qquad (60)
\end{aligned}
$$

where (a) is because $\mathbb{E}\left[T_k^{\mathsf{QRR}}\right] \leq \mathbb{E}\left[T_{\max}\right]$. Using (55)(60) to upper bound the last term of (59), we have

$$
\frac{1}{t_K}\sum_{\tau=0}^{t_K-1} \mathbb{E}\left[Q_n(\tau)\right] \leq B_2 + \mathbb{E}\left[T_{\max}\right] \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}\left[Q_n(t_{k-1})\right], \qquad (61)
$$

where $B_2 \triangleq \frac{1}{2}\mathbb{E}\left[(T_{\max})^2\right] A_{\max} < \infty$. Summing (61) over $n \in \{1, \ldots, N\}$ and passing $K \to \infty$, we get

$$
\begin{aligned}
&\limsup_{K\to\infty} \frac{1}{t_K}\sum_{\tau=0}^{t_K-1}\sum_{n=1}^{N} \mathbb{E}\left[Q_n(\tau)\right] \\
&\leq NB_2 + \mathbb{E}\left[T_{\max}\right]\left(\limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\sum_{n=1}^{N}\mathbb{E}\left[Q_n(t_{k-1})\right]\right) \\
&\overset{(a)}{\leq} NB_2 + \mathbb{E}\left[T_{\max}\right] B_1/\epsilon < \infty,
\end{aligned}
$$

where (a) is by (54). The proof is complete. ■

References

[1] C.-P. Li and M. J. Neely, "Exploiting Channel Memory for Multi-User Wireless Scheduling without Channel Measurement: Capacity Regions and Algorithms," in *IEEE Int. Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Avignon, France, May 2010.

[2] ——, "Energy-optimal scheduling with dynamic channel acquisition in wireless downlinks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 4, pp. 527–539, Apr. 2010.

[3] H. S. Wang and P.-C. Chang, "On verifying the first-order markovian assumption for a rayleigh fading channel model," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 353–357, May 1996.

[4] M. Zorzi, R. R. Rao, and L. B. Milstein, "A Markov model for block errors on fading channels," in *Personal, Indoor and Mobile Radio Communications Symp. PIMRC*, Oct. 1996.

[5] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Athena Scientific, 2005, vol. I.

[6] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multi-channel opportunistic access: Structure, optimality, and preformance," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5431–5440, Dec. 2008.

[7] S. H. A. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multichannel opportunistic access," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4040–4050, Sep. 2009.

[8] S. Ahmad and M. Liu, "Multi-channel opportunistic access: A case of restless bandits with multiple plays," in Proc. Allerton Conf. Commun. Control Comput., Allerton, IL, Oct. 2009, pp. 1361–1368.

[9] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of whittles index for dynamic multichannel access," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5547–5567, Nov. 2010.

[10] M. J. Neely, "Stochastic optimization for markov modulated networks with application to delay constrained wireless scheduling," in *IEEE Conf. Decision and Control (CDC)*, 2009.

[11] Q. Zhao and A. Swami, "A decision-theoretic framework for opportunistic spectrum access," *IEEE Wireless Commun. Mag.*, vol. 14, no. 4, pp. 14–20, Aug. 2007.

[12] A. Pantelidou, A. Ephremides, and A. L. Tits, "Joint scheduling and routing for ad-hoc networks under channel state uncertainty," in *IEEE Int. Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Apr. 2007.

[13] L. Ying and S. Shakkottai, "On throughput optimality with delayed network-state information," in *Information Theory and Application Workshop (ITA)*, 2008, pp. 339–344.

[14] ——, "Scheduling in mobile ad hoc networks with topology and channel-state uncertainty," in *IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009.

[15] P. Chaporkar, A. Proutiere, H. Asnani, and A. Karandikar, "Scheduling with limited information in wireless systems," in *ACM Int. Symp. Mobile Ad Hoc Networking and Computing (MobiHoc)*, New Orleans, LA, May 2009.

[16] R. G. Gallager, *Discrete Stochastic Processes*. Kluwer Academic Publishers, 1996.

[17] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Probab.*, vol. 25, pp. 287–298, 1988.

[18] J. C. Gittins, *Multi-Armed Bandit Allocation Indices*. New York, NY: Wiley, 1989.

[19] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 466–478, Mar. 1993.

[20] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, 2006.

[21] C.-P. Li and M. J. Neely, "Network Utility Maximization over Partially Observable Markovian Channels," arXiv report:1008.3421v1, Aug. 2010.