Adaptive Communications and Signal Processing laboratory
Cornell University, Ithaca, NY 14853

Cross-layer Designs of Multichannel Reservation MAC under Rayleigh Fading

Atul Maharshi, Lang Tong, Ananthram Swami

Part of Technical Report No. ACSP-TR-09-02-01, September 2002

Accepted for publication in *IEEE Transactions on Signal Processing*, **August 2003**

Abstract

We consider a reservation-based medium access control (MAC) scheme where users reserve data channels through a slotted-ALOHA procedure. The base station grants access to users in a Rayleigh fading environment using measurements at the physical layer and system information at the MAC layer. This paper has two contributions pertaining to simple reservation based medium access. First, we provide a Markov chain formulation to analyze the performance (throughput/channel utilization) of multichannel slotted system. Second, a Neyman-Pearson like MAC design optimized for performance is presented. This design can serve as a benchmark in evaluating the performance of other designs based on conventional physical layer detectors such as Maximum A posteriori Probability, Maximum Likelihood and Uniformly Most Powerful detectors. Results show that utilizing system information in addition to the physical layer measurements indeed leads to a gain in performance. We discuss the issue of further improving the performance in fading by means of multiple measurements and also comment upon the delay/channel-utilization trade-off for the optimal MAC design.

I. Introduction

Standard designs of reservation-based medium access control (MAC) consist of two separate steps: a detector at the physical layer (PHY) that estimates the number of requests on a particular channel and an acknowledgement protocol at the MAC sub-layer based on the PHY layer output. Typically, if each channel can accommodate a single transmission, the detector at the physical layer tests the hypothesis that there is exactly one user requesting the channel. For example, a simple MAC design for the random access channel (RACH) of the UMTS-WCDMA [18] may acknowledge a particular channel if the strength of the measured signal exceeds certain thresholds [11], [20].

It is not obvious that treating the MAC problem as one of detecting the number of users followed by some acknowledgement protocol leads to any optimality at the MAC layer; the detector that minimizes the probability of detection error at the PHY layer need not be the one that maximizes the throughput or the one that minimizes the expected delay. In their seminal papers [9], [16], Kleinrock and Tobagi analyzed the impact of physical layer detection of the busy-tone in the context of carrier sensing multiple access (CSMA). There they showed the unusual effects of miss detection and false alarm on the MAC throughput.

Also missing in the separate design of PHY and MAC layers is the possibility of utilizing the MAC parameters at the physical layer, and in the reverse direction, the measurements at the PHY layer in the MAC acknowledgement. This interaction is particularly relevant in a multichannel MAC where the traffic statistics of the number of users requesting a channel are intertwined with the number of channels that are occupied in a particular time slot. Passing down the information on the number of available channels at a particular time to the detector may improve the performance.

In this paper, we consider a generic multichannel reservation-based MAC in a Rayleigh fading environment where users request transmissions by sending a signature randomly chosen from a pool of orthogonal codes representing the set of available channels. The receiving node grants or denies their transmissions based on the measured signal strength. A collision occurs if multiple (> 1) users send requests for a channel and that channel is acknowledged by mistake. On the other hand, if a channel is acknowledged

without any user requesting for it, it is mistakenly taken out of the pool of available channels for other users, which causes inefficient channel (code) utilization, heavier traffic, and more frequent collisions in other channels. Such random access schemes have been proposed for the UMTS-WCDMA [18].

One of the difficulties of a joint PHY and MAC design, in general, is the lack of analytical expressions that relate MAC performance to PHY layer parameters. Our first objective is to obtain such an analytical expression. We model the MAC scheme as a finite state Markov chain for which a stationary distribution exists and is parameterized by two probability vectors. When the number of codes assigned to the receiver is two ($N = 2$), the stationary distribution can be obtained exactly, which leads to an analytical expression for the MAC throughput as a function of certain MAC parameters.

The second step is to optimize the MAC function based on the throughput expression. Here we derive the optimal randomized MAC function that maps the measurements at the PHY layer and the system states to the probability that a channel is acknowledged. In a proof similar to that of the celebrated Neyman-Pearson Lemma, we give the form of the optimal MAC function.

The third step is to compare the optimal MAC with several sub-optimal but simpler MAC functions. Some of these suboptimal MAC protocols also employ the idea of cross-layer design but make less restrictive assumptions on the traffic statistics. The performance loss is evaluated through simulations.

The approach presented in this paper applies to two different types of networks. The first is the cellular network where the base-station allocates channels using some form of demand-assignment strategies. The second type is the ad hoc network that employs code division multiple access (CDMA) and receiver-based transmission protocol [15]. In such a network, a transmitting node wishing to communicate with a receiving node must know and use codes assigned to the receiver, and a request-acknowledge process may be necessary. The MAC considered here is similar to the widely used RTS-CTS protocol except that the request and acknowledgement are performed at the signal level. For ad hoc networks, the number of codes available at each node is small, which makes our exact analysis attractive. On the other hand, ad hoc networks are often half-duplex, and the ultimate performance is measured by the end-to-end throughput. Our results should be viewed as applicable to the local MAC performance when the node is in the receiving mode.

Although the literature on the joint optimization of PHY and MAC sublayers is scarce, there have been recent interest in the cross layer design of MAC for wireless networks [17]. Signal processing techniques have been used for separating colliding packets [19], [22], and more sophisticated MAC protocols are needed to take advantage of the improved PHY layer [21], [1]. The impact of PHY layer performance (fading, capture, and multipacket reception) on the MAC layer has been investigated by a number of authors [7], [13], [14], [3]. Chockalingam et. al. [6] investigated a multichannel reservation system very similar to the one presented here. However, the issue of designing an acknowledgement strategy does not arise in the set-up they consider. Kleinrock and Tobagi were, perhaps, the first to address the issue of detection error on the (CSMA) MAC protocol. The idea of combining signature detection with channel

allocation was considered by Butala and Tong[5], [4]. The optimal MAC, however, was not considered there.

The paper is organized as follows. We present the basic functions and assumptions for mobile and base stations and the fading signal model in Sec.II. In Sec.III, we present the Markov chain formulation for obtaining the throughput which is the criterion for optimization. Sec.IV presents the optimal MAC design based on the received signal power and the number of available codes. Other sub-optimal designs are considered in Sec.V. In Sec.VI we deal with issues such as delay and improving the throughput through multiple measurements. Simulation results are presented and analyzed in Sec.VII; and some concluding remarks are given in Sec.VIII.

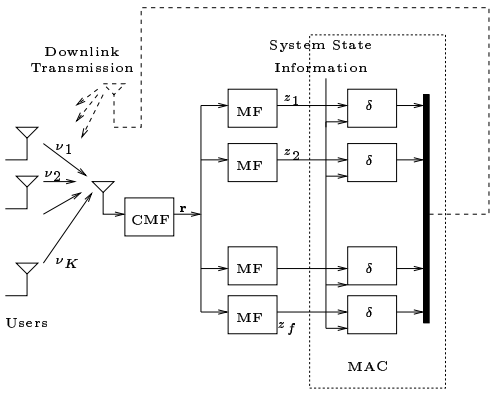

Fig. 1. A reservation-based random access CDMA scheme. CMF: Chip-matched-filtering and sampling. MF: Code matched filter.

II. System Model

The system considered here is similar to that used in the random access channel (RACH) in WCDMA [18] and is illustrated in Fig. 1. It is worth pointing out again that we use the term *base station* to include the usual cellular base station, as well as *clusterheads* or privileged nodes that have multiple codes in an *ad hoc* network.

A. The Mobile Stations

The random access scheme is based on slotted ALOHA channel reservation. At the beginning of slot t, the base station broadcasts a set $\mathcal{C}_t$ of available orthogonal preamble signatures for uplink reservation. We will denote the number of available signatures by F_t, thus $F_t = |\mathcal{C}_t| \leq N$, where N is the total number of channels in the system. An interested user transmits a randomly selected signature from $\mathcal{C}_t$ and waits for an acknowledgement. If a positive acknowledgement is received, the user proceeds to transmit data using an orthogonal code that has a one-to-one relationship with the preamble signature. The data transmission lasts for a fixed duration of L slots. If a channel is acknowledged when two or more users attempted access, a collision occurs and the channel becomes *locked*, i.e., it is unavailable to the other users even though

the channel is not contributing to the throughput. We further note that a channel might get locked when the base station transmits an ACK even when no user is attempting access. Regardless of the way a channel is occupied, we assume that the channel remains unavailable to other users for a length of L slots. The rationale for this assumption is that a base-station expects data-transmission to follow on an acknowledged channel. In case no acknowledgement is received, the user backs off and retries after a random delay. A user's back-off timer may expire when no channels are free; in such a case, it will reset its back-off timer to a new random value. We assume that no preamble power ramping is carried out *i.e.,* a user does not increase power on retries.

We make the classical assumption that the access attempts, which include new arrivals as well as retries, are points of a Poisson process with intensity λ attempts/slot. We emphasize that λ denotes the aggregate attempt rate and not the packet arrival rate. It corresponds to the parameter G used by Kleinrock and Tobagi in their analysis of CSMA [9]. In light of this assumption the resulting throughput analysis should be seen as a steady-state analysis (the input arrival rate being equal to the departure rate) with stability implicitly assumed. The Poisson assumption, of course, may not hold in practice, and it disregards the detailed retransmission mechanisms. Also, by making this assumption, we have implicitly assumed that we are working with an infinite-user, single-buffer scenario. Nonetheless, this assumption lends itself to tractable analysis which can yield sufficient insight for dealing with more realistic scenarios. See [2, Ch. 3-4],[9] for comments in this regard.

B. The Base Station

After announcing the available preamble signatures $\mathcal{C}_t$, the base station performs matched filtering for each code in $\mathcal{C}_t$. Based on the output of each matched filter, the base-station makes decisions on acknowledgement. The assumption of Poisson arrivals makes it possible for individual channels to take decisions independently, as shown in Fig. 1.

To assemble the set of available preamble signatures for the next slot, the base station first takes out codes acknowledged in the current slot. It then checks whether any codes had been allocated L slots earlier (which should now be free), and adds these released codes to the new signature pool.

C. The Signal Model

The preamble power received at the detector is a critical parameter in the MAC design problem. The results obtained in this paper apply to systems in which transmissions over different channels are orthogonal. This orthogonality can be achieved by various means we are familiar with - channels could be separated by means of codes, time, frequency or a mix of them.

We shall now obtain the distribution of received power for the specific signal model used in this paper which achieves this orthogonality through codes. As shown in Fig. 1, the detector takes as its input the sampled chip-matched filtered signal. We assume that the transmitted signal undergoes Rayleigh flat

fading, the Rayleigh parameter having the same value (σ_d^2) for each user. This corresponds to a situation in which power-control has been achieved to combat long-term (shadow) fading, but, the system is still susceptible to short-term fluctuations in the signal strength. Assume that in slot t, $F_t = f$ channels are available and K users contend for reservation. The sampled output of the chip-matched filtered can be written as:

$$\mathbf{r} = \sum_{k=1}^{K} \nu_k \mathbf{s}_k + \mathbf{w} \tag{1}$$

where ν_k are the complex amplitudes which are realizations of i.i.d. random variables (in keeping with our assumption of partial power-control) with distribution $\mathcal{CN}(0, \sigma_d^2/2)$ where σ_d^2 is the SNR; symbols in bold font denote vectors of length N_s, the signature length. The signatures, $\mathbf{s}_k$, belong to the set of available orthogonal signatures, $\mathcal{C}_t = \{\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_f\}$. The elements of $\mathcal{C}_t$ have a one-to-one relationship with the set of available channels, and $\mathbf{c}_i^H \mathbf{c}_i = 1$ for $i = 1, 2, ..., f$ (H denotes the Hermitian operator). The noise term $\mathbf{w}$ is a realization of AWGN with distribution $\mathcal{CN}(0, \frac{1}{2}\mathbf{I})$, in accordance with our definition of σ_d^2 as the SNR.

At the i^{th} detector, decorrelating with the signature $\mathbf{c}_i$ we get:

$$z_i = \mathbf{c}_i^H \mathbf{r} \tag{2}$$

$$= \sum_{k:\mathbf{s}_k = \mathbf{c}_i} N_s \nu_k + w_i = x_i + w_i \tag{3}$$

where x_i is a realization of $X_i \sim \mathcal{CN}(0, \theta_i \sigma_d^2/2)$, θ_i being the number of users selecting signature $\mathbf{c}_i$, and w_i is a realization of a random variable with distribution $\mathcal{CN}(0, 1/2)$. The assumption of the arrivals being Poisson implies that θ_i itself is a realization of $\Theta_i \sim Poisson(\lambda/f)$. We can interpret z_i to be a realization of $Z_i \sim \mathcal{CN}(0, (\theta_i \sigma_d^2 + 1)/2)$. Thus, the received signal power $Y_i = |Z_i|^2$, has the distribution:

$$p_{Y_i|\Theta_i}(y_i|\theta_i) = \frac{1}{\theta_i \sigma_d^2 + 1} \exp\left(\frac{-y_i}{\theta_i \sigma_d^2 + 1}\right) . \tag{4}$$

The MAC must use the received signal to decide whether or not a single user is requesting access, *i.e.,* if $\Theta_i = 1$, and then based upon the accuracy of this decision carry out the appropriate acknowledgement procedure. We note that as a result of the Rayleigh fading assumption, Z_i is circularly symmetric complex Gaussian and, thus, Y_i is a sufficient statistic that can be generated from $\mathbf{r}$ for this purpose. We will drop the subscript i for the detector here onwards, as, given $F_t = f$, the working of each detector is identical to that of the rest.

Now, the size of $\mathcal{C}_t$ varies from slot to slot, which makes the attempt rate time varying at each channel (even though the overall attempt rate is constant). The fluctuation of the available signatures and, therefore, fluctuations in traffic affect the distribution of the received signal power. This dictates that a MAC function should adapt to the system state in order to deliver optimal performance making the optimal MAC design problem non-trivial. In the next section, we give a formulation to compute the performance (throughput) achieved using a given MAC policy. We will then consider designs that optimize the performance in Sec.IV.

III. MAC Performance

We consider MAC functions that devise their ACK policies based on the number of free codes available and the received signal power for each of these. In this section, we will show that such MAC functions induce a Markov chain structure facilitating throughput analysis. For a MAC function δ, throughput will be seen to depend upon $\alpha_\delta(f)$ $(= 1 - \beta_\delta(f))$ and $\gamma_\delta(f)$; $\alpha_\delta(f)$ is the conditional probability of acknowledging a channel, given that there are f free channels; $\gamma_\delta(f)$ is the conditional probability of *successfully* acknowledging a channel, given that there are f free channels.

A. Markov Chain Formulation

A channel once occupied remains so for a duration of L slots. The system, thus, has a memory of L slots. We define the state vector as

$$\mathbf{n}_t = [n_{t-L+1}, \cdots, n_t] \in \{S_0, \cdots, S_{|\mathcal{S}|-1}\} = \mathcal{S} \tag{5}$$

where n_t is the number of newly locked channels at the beginning of slot t, and $\mathcal{S}$ denotes the state space. Note that $\sum_{l=1}^{L} n_{t-l+1} \leq N$ for all t and, thus, $|\mathcal{S}|$ is finite and we can enumerate the states as in (5). We must represent state i itself by a vector as $S_i = [S_i(1), \ldots, S_i(L)]$. Thus, if $\mathbf{n}_t = S_i$, it would mean that $S_i(l) = \mathbf{n}_t(l) = n_{t-L+l}$ channels got locked at time $(t - L + l)$ for $l = 1, \ldots, L$; l here denotes the index of an element in a vector. The enumeration of the states can be done arbitrarily, e.g., for $N = 2$, $L = 3$, an enumeration is shown in table I. Thus, when the system is in state 7 ([0 1 1]), it means that one of the channels got locked in the previous slot and the other channel got locked in the slot before the previous one, leaving no channels free in the current slot ($f_7 = 0$). Whereas if the current state is 2 ([0 1 0]), we can say that one of the channels got locked in the slot before the last one, but the other channel is free ($f_2 = 1$) and users with packets to send can contend for this free channel.

The Poisson traffic assumption means that the traffic statistics are known when the number of free channels is known. For our definition of states, the number of free channels with the system in state i is given by $N - \sum_{l=1}^{L} S_i(l)$. We see that if the MAC bases its decisions on the traffic statistics and the received signal power, the transition probability from current state to the next depends solely on the two states involved and is independent of the transition history leading to the current state. The system can, therefore, be modelled as a Markov chain with the states defined as above. The transition probabilities from state i are dependent on $\alpha_\delta(f_i)$ - the conditional probability of acknowledging a channel when using the MAC function δ, given that there are $f_i (> 0)$ free channels.

For an enumeration of the states, $\{0, \ldots, (|\mathcal{S}| - 1)\}$, denote the transition matrix by P. The $(i+1, j+1)^{th}$ entry of the transition matrix is the probability that the state in slot $(t + 1)$ is j, given that the state in slot t was i:

$$P_{i+1,j+1} = Prob\{\mathbf{n}_{t+1} = S_j | \mathbf{n}_t = S_i\} \ . \tag{6}$$

Symbol	State	f_i
S_0	[0 0 0]	2
S_1	[0 0 1]	1
S_2	[0 1 0]	1
S_3	[1 0 0]	1
S_4	[0 0 2]	0
S_5	[0 2 0]	0
S_6	[2 0 0]	0
S_7	[0 1 1]	0
S_8	[1 0 1]	0
S_9	[1 1 0]	0

TABLE I

STATE TABLE FOR $N = 2,\ L = 3$

Now, we have the obvious condition that $\mathbf{n}_{t+1}(l) = \mathbf{n}_t(l+1) = n_{t+l+1-L}$, for $l = 1, \ldots, L-1$. We also have the condition that $n_{t+1} \leq N - \sum_{l=1}^{L} n_{t-l+1}$, since the number of channels that get occupied can only be less than or equal to the number of available codes. Thus, $P_{i+1,j+1}$ is nonzero only if

$$S_j(l) = S_i(l+1),\ l = 1, \ldots, L-1 \tag{7}$$

$$S_j(L) \leq N - \sum_{l=1}^{L} S_i(l) \triangleq f_i\ . \tag{8}$$

When $f_i = 0$, the condition in (8) becomes:

$$S_j(L) = 0\ . \tag{9}$$

We note that for any state i with $f_i = 0$[1], there is only one j that satisfies conditions (7) and (9) and, therefore, for this pair of states:

$$P_{i+1,j+1} = 1\ . \tag{10}$$

Fig. 2 shows the state diagram of the Markov chain for $N = 2,\ L = 3$ with states as given in table I. We see that the transitions from states $4, \ldots, 9$ (states with no free channels) are fixed.

When $f_i > 0$, we note from (8) that $S_j(L) \in \{0, \ldots, f_i\}$. This means that we can go from state i to one of $(f_i + 1)$ number of states depending on how the free channels are acknowledged. Since the acknowledgement probability, $\alpha_\delta(f_i)$, is identical and independent for all free channels, the transition

[1]Note that we do not define any acknowledgement probability $\alpha_\delta(0)$ as there are no channels to acknowledge when there are no free channels.

probability from state i to state j, provided conditions (7)-(8) are satisfied, is Binomial($S_j(L)$, f_i, $\alpha_\delta(f_i)$), i.e.,

$$P_{i+1,j+1} = \binom{f_i}{S_j(L)} \alpha_\delta^{S_j(L)}(f_i)\ \beta_\delta^{(f_i - S_j(L))}(f_i)\ . \tag{11}$$

Referring to Fig. 2, from state 0 (2 free channels) we can either go to one of states 0 and 4 or come back to state 0. In order to make a transition into state 4, both the free channels will have to get locked. Since the acknowledgement probability for a free channel in this case is $\alpha(2)$, both channels get locked with a probability $\alpha^2(2)$. Similarly, we come back to state 0 if none of the free channels get locked which happens with a probability $\beta^2(2)$ as shown in Fig. 2. It can be shown that the Markov chain is aperiodic and irreducible for arbitrary N and L when $0 < \alpha_\delta(f) < 1$ for all $f > 0$ [12]. The proof is based on the facts that (i) any state is accessible from S_0, (ii) state S_0 is accessible from any other state and (iii) there is always a self-loop associated with state S_0.

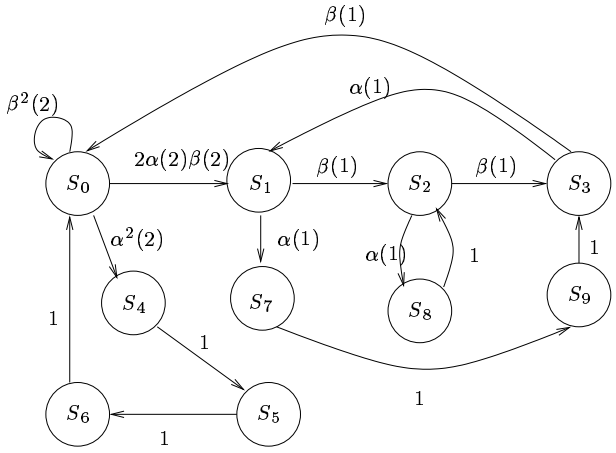

Fig. 2. Markov Chain for $N = 2$, $L = 3$

Since the chain is irreducible, aperiodic and finite state, a unique stationary distribution exists. The calculation of the stationary distribution from the transition matrix can be simplified by noting that there exist groups of states which have the same stationary distribution. For example, with $N = 3$ (with arbitrary L), all the states of the form $[0 \ldots 030 \ldots 0]$ will have the same stationary probability as the state $[0 \ldots 03]$. In the appendix, we derive the stationary distribution for $N = 2$ for arbitrary L. So far, $N = 2$ is the case for which closed form expression for the stationary distribution for general L could be obtained. For other cases, it is, of course, possible to obtain the stationary distribution numerically. The state space size increases exponentially with N and L with $|\mathcal{S}| = \sum_{k=0}^{\min(N,L)} \binom{N}{k}\binom{L}{k}$. However, the transition matrix is sparse, e.g., for $N = 2$, it is easy to show that the number of non-zero entries is $(L+2)(L+3)/2$, whereas the number of states is $(L+1)(L+2)/2$, so that the fractional number of non-zero entries is of order $1/L^2$. The sparsity holds for general N, L.

B. Throughput and Channel Utilization

Having obtained the stationary distribution (which we will denote by $\boldsymbol{\pi}_\delta = [\pi_{\delta,0}, ..., \pi_{\delta,(|\mathcal{S}|-1)}]$) one can now obtain figures of merit for network performance. We consider two figures of merit, throughput and channel utilization. Throughput is defined as the average number of successful access attempts per slot and channel utilization is the number of successful transmissions per slot per channel. Channel utilization can also be thought of as the fraction of the slots that actually get utilized for data transmission.

With f free channels in a slot, the expected number of successful access attempts is $f\gamma(f)$. Thus, the throughput can be written as:

$$\eta_\delta = \sum_{i=0}^{|\mathcal{S}|-1} f_i \gamma_\delta(f_i) \pi_{\delta,i} \ . \tag{12}$$

We can rewrite the last equation in the form:

$$\eta_\delta = \sum_{f=0}^{N} f \gamma_\delta(f) \pi(f, \boldsymbol{\alpha}_\delta) \tag{13}$$

where $\pi(f, \boldsymbol{\alpha}_\delta)$ is given by

$$\pi(f, \boldsymbol{\alpha}_\delta) = \sum_{i \in \{j: f_j = f\}} \pi_{\delta,i} \ . \tag{14}$$

We can interpret $\pi(f, \boldsymbol{\alpha}_\delta)$ as the stationary distribution of F_t. In (14), we have tried to emphasize the dependance of the stationary distribution on $\boldsymbol{\alpha}_\delta$. Note that $\boldsymbol{\gamma}_\delta$ does not affect the stationary distribution but affects the throughput. This apparent "decoupling" between $\boldsymbol{\alpha}_\delta$ and $\boldsymbol{\gamma}_\delta$ has consequences in the derivation of the optimal MAC function as will be seen later.

Each successful user occupies the channel for L slots. Thus, the average number of successful transmissions per slot per channel, *i.e.,* the channel utilization, is given by:

$$\zeta_\delta = L\eta_\delta / N \ . \tag{15}$$

Thus, for a given L, the detector strategy that maximizes η also maximizes the channel occupancy. In the appendix, we obtain the throughput and channel utilization expressions for the case of $N = 2$ for arbitrary L.

In the next section, we give the form of the optimal MAC function which is the principal contribution of this paper along with the Markov chain performance analysis. The function optimizes the performance in terms of throughput as derived in this section. We give proof of its optimality and existence.

IV. The Optimal MAC

As pointed out before, MAC functions should base their ACK policies on the number of free codes available and the received signal power for each of these. We define the MAC function as:

$$\delta \stackrel{\Delta}{=} \mathbb{R}_+ \times \mathbb{F} \to [0, 1] \tag{16}$$

where $\mathbb{R}_+$ is the set of non-negative reals, $\mathbb{F} = \{1, 2, \ldots, N\}$ is the observation space of F_t and $\delta(y, f) = q$ means that the channel is acknowledged with probability q when $Y = y$ and $F_t = f$. This definition of a MAC function helps us evaluate the probabilities: $\boldsymbol{\alpha}_\delta = [\alpha_\delta(1), \ldots, \alpha_\delta(N)]$ and $\boldsymbol{\gamma}_\delta = [\gamma_\delta(1), \ldots, \gamma_\delta(N)]$, where $\alpha_\delta(f)$ and $\gamma_\delta(f)$ are as defined in Sec.III. We have:

$$\begin{aligned} \alpha_\delta(f) &= E(\delta | F_t = f) && (17) \\ &= \int_0^\infty \delta \sum_{\theta=0}^{\infty} p_{Y|\Theta}(y|\theta) p_{\Theta|F_t}(\theta|f) dy && (18) \end{aligned}$$

and

$$\begin{aligned} \gamma_\delta(f) &= E(\delta 1_{\Theta=1} | F_t = f) && (19) \\ &= \int_0^\infty \delta p_{Y|\Theta}(y|1) p_{\Theta|F_t}(1|f) dy && (20) \end{aligned}$$

Here, $p_{Y|\Theta}(y|\theta)$ denotes the p.d.f. of the received power given the number of users and is given in (4), and $p_{\Theta|F}(\theta|f)$ is the p.m.f. of the number of users given the number of free channels. The stationary distribution can then be obtained as in the previous section from which the throughput can be computed using (12).

A. The Problem

We can formulate the problem as: Given the total number of channels N, packet length L and overall arrival rate λ, and given that $p_{Y|\Theta}(y|\theta)$ and $p_{\Theta|F}(\theta|f)$ are known, determine the MAC function δ which maximizes the throughput (12), i.e., find δ_o such that

$$\delta_o = \arg\max_\delta \sum_{f=1}^{N} f \gamma_\delta(f) \pi(f, \boldsymbol{\alpha}_\delta) \ . \tag{21}$$

We first define the aposteriori probability functions as:

$$\begin{aligned} l_0(y; f) &\triangleq \sum_{\theta \neq 1} p_{Y|\Theta}(y|\theta) p_{\Theta|F}(\theta|f) && (22) \\ l_1(y; f) &\triangleq p_{Y|\Theta}(y|1) p_{\Theta|F}(1|f) \ . && (23) \end{aligned}$$

The principal result concerning the optimal MAC function is given in the following proposition modelled on the Neyman-Pearson Lemma.

Proposition 1: For a system with $l_0(y; f)$ and $l_1(y; f)$ as given in (22), (23) and $\boldsymbol{\alpha} \geq 0$, the following statements are true.

Optimality: Let δ be a MAC function such that $\boldsymbol{\alpha}_\delta = \boldsymbol{\alpha}$ and let δ^* be a MAC function of the form:

$$\delta^*(y, i) = \begin{cases} 1 & \text{when } l_1(y; f) > \rho_f l_0(y; f) \\ \xi(y, i) & \text{when } l_1(y; f) = \rho_f l_0(y; f) \\ 0 & \text{when } l_1(y; f) < \rho_f l_0(y; f) \end{cases} \tag{24}$$

with $\rho_f \geq 0$ and $0 \leq \xi(y,i) \leq 1$ such that $\boldsymbol{\alpha}_{\delta^*} = \boldsymbol{\alpha}$. Then $\eta_{\delta^*} \geq \eta_\delta$.

Existence: For every $\boldsymbol{\alpha} \in (0,1)^N$, there exists a MAC function δ_o of the form in (24) with $\xi(y;f) = \xi_0(f)$ such that $\boldsymbol{\alpha}_{\delta_o} = \boldsymbol{\alpha}$.

Proof: Optimality

We first prove that among MACs with $\boldsymbol{\alpha}_\delta = \boldsymbol{\alpha}$, the MAC of the form in (24), if it exists, gives the highest throughput.

Let $\delta^*(y;f)$ have the specified form with $\boldsymbol{\alpha}_{\delta^*} = \boldsymbol{\alpha}$. The existence of such a MAC will be established later. Let $\delta(y;f)$ be any other MAC function with $\boldsymbol{\alpha}_\delta = \boldsymbol{\alpha}$. Define

$$\begin{aligned}\Gamma_1^*(f) &\triangleq \{y : l_1(y;f) > \rho_f l_0(y;f)\}, \quad \partial\Gamma_1^*(f) \triangleq \{y : l_1(y;f) = \rho_f l_0(y;f)\} \\ \Gamma_0^*(f) &\triangleq \{y : l_1(y;f) < \rho_f l_0(y;f)\}\end{aligned}$$

Compare the probabilities of successful reservation:

$$\begin{aligned}\gamma_{\delta^*}(f) - \gamma_\delta(f) &= \int_0^\infty (\delta^*(y;f) - \delta(y;f)) l_1(y;f) dy \\ &= \int_{\Gamma_1^*(f)} (1 - \delta(y;f)) l_1(y;f) dy + \int_{\partial\Gamma_1^*(f)} (\xi(y;f) - \delta(y;f)) l_1(y;f) dy \\ &\quad - \int_{\Gamma_0^*(f)} \delta(y;f) l_1(y;f) dy \\ &> \rho_f \int_{\Gamma_1^*(f)} (1 - \delta(y;f)) l_0(y;f) dy + \rho_f \int_{\partial\Gamma_1^*(f)} (\xi(y;f) - \delta(y;f)) l_0(y;f) dy \\ &\quad - \rho_f \int_{\Gamma_0^*(f)} \delta(y;f) l_0(y;f) dy \\ &= \rho_f(\alpha_{\delta^*}(f) - \alpha_\delta(f)) - \rho_f(\gamma_{\delta^*}(f) - \gamma_\delta(f)) \\ &= -\rho_f(\gamma_{\delta^*}(f) - \gamma_\delta(f))\end{aligned}$$

We have that $\gamma_{\delta^*}(f) - \gamma_\delta(f) \geq 0$. Now, since $\boldsymbol{\alpha}_{\delta^*} = \boldsymbol{\alpha}_\delta$, we have $\boldsymbol{\pi}_{\delta^*} = \boldsymbol{\pi}_\delta$, thus implying that $\eta_{\delta^*} \geq \eta_\delta$.

Existence

We now show that for any $\boldsymbol{\alpha}$, there exists a MAC δ of the specified form with $\boldsymbol{\alpha}_\delta = \boldsymbol{\alpha}$. With this, we conclude that the optimal MAC function can be obtained using the specified form. We only need to consider the case when $l_0(y;f) > 0$. Now, let

$$a(\rho) \triangleq 1 - \Pr[\frac{l_1(Y,f)}{l_0(Y;f)} \leq \rho]. \tag{25}$$

Since $\Pr[\frac{l_1(Y;f)}{l_0(Y;f)} \leq \rho]$ is a cumulative distribution function, $a(\rho)$ is right-continuous and monotonically decreasing. Hence for any $\alpha(f)$, we can always find a ρ_f and a ξ such that

$$\alpha(f) = a(\rho_f) + \xi \Pr[l_1(Y;f) = \rho_f l_0(Y;f)] \tag{26}$$

This completes the proof. □

B. Optimization

Proposition 1 above suggests that we consider MAC designs as in (24) and optimize for $\boldsymbol{\alpha}$, through $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_N]$, to obtain the design which gives highest throughput. Unfortunately, the throughput needs to be obtained via the Markov chain formulation and we have to search for this optimal design numerically. We will first see how to obtain the throughput for one set of parameters.

We know that the conditional distribution of Y given $\Theta = \theta$ is

$$p(y|\theta) = \frac{1}{\theta\sigma_d^2 + 1} \exp\left(\frac{-y}{\theta\sigma_d^2 + 1}\right) . \tag{27}$$

Since the arrivals are Poisson, given that f channels are free, the access attempt rate for a particular channel is $\lambda_f = \lambda/f$. Thus, the prior probability for Θ given the arrival rate can be written as

$$p(\theta|f) = \exp(-\lambda_f)\frac{(\lambda_f)^\theta}{\theta!} . \tag{28}$$

The ratio of the aposteriori probability functions is, thus, given by:

$$\frac{l_1(y;f)}{l_0(y;f)} = \frac{\frac{1}{\sigma_d^2+1}\exp\left(\frac{-y}{\sigma_d^2+1}\right)\lambda_f}{\sum_{\theta\neq 1}\frac{1}{\theta\sigma_d^2+1}\exp\left(\frac{-y}{\theta\sigma_d^2+1}\right)\frac{\lambda_f^\theta}{\theta!}} . \tag{29}$$

Given for $F = f$, we can numerically determine the decision regions[2], $\Gamma_1(\rho_f, \lambda_f)$ and $\Gamma_0(\rho_f, \lambda_f)$ corresponding to the two hypotheses. Note that for a system operating at an SNR of σ_d^2, the decision regions are dependent on the number of free channels through ρ_f and λ_f and we have chosen to emphasize this dependence by denoting the decision regions as $\Gamma_i(\rho_f, \lambda_f)$ rather than $\Gamma_i(f)$. Note that for the region where the aposteriori ratios are equal, $\partial\Gamma_1$, is of measure zero. The decision regions are of the form:

$$y \in \Gamma_1(\rho_f, \lambda_f) \quad \text{if } \tau_1(\rho_f, \lambda_f) \leq y \leq \tau_2(\rho_f, \lambda_f) \tag{30}$$

$$y \in \Gamma_0(\rho_f, \lambda_f) \qquad \text{otherwise} \tag{31}$$

where $\tau_1(\rho_f, \lambda_f)$ and $\tau_2(\rho_f, \lambda_f)$ can be interpreted as power thresholds based on which the detector makes its decisions. The decision regions were obtained numerically as no closed form expressions could be found. Intuitively, we would expect the decision regions to be of the form given in (30)-(31), so that power falling below the lower threshold corresponds to the case of no user attempting access, while power falling above the upper threshold corresponds to the case of two or more users attempting access.

In Fig. 3, we show the thresholds as a function of access rate λ_f for SNR's of 5 dB and 10 dB with $\rho_f = 1$. Also plotted in Fig. 3 is the variation of thresholds vs. SNR for $\lambda_f = 0.5$, $\rho_f = 1$. The upper threshold τ_2 decreases as the SNR decreases and as λ_f increases, which is intuitive. What is surprising is the lack of variation of the lower threshold τ_1 which is not very sensitive to λ_f or SNR. Contrary to intuition, it does not go down as λ_f increases.

[2]Note that in the current problem setting, $\Pr[l_1(Y;f) = \rho_f l_0(Y;f)] = 0$, and thus, the randomized nature of the MAC function is not conspicuous.

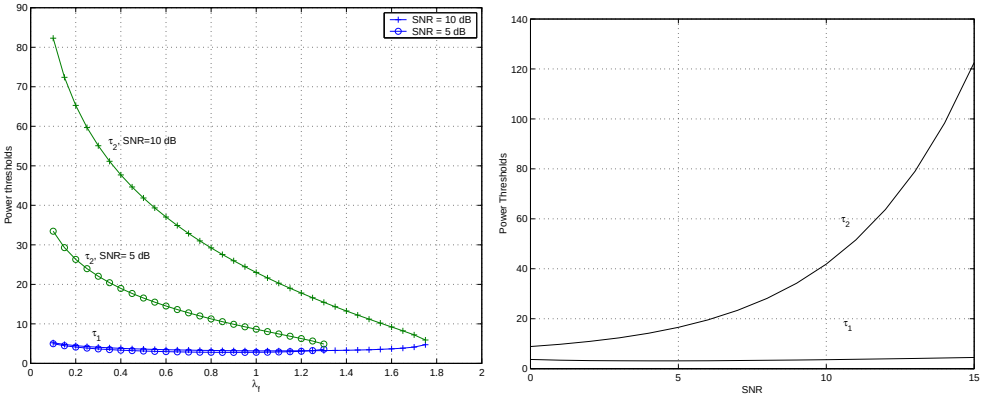

Fig. 3. Thresholds vs. access rate λ_f for SNR = 5 dB and 10 dB(left), Thresholds vs. SNR for $\lambda_f = 0.5$(right), $\rho_f = 1$

Having obtained the decision regions, we can now obtain the event probabilities needed for the Markov chain formulation. Let $\tau_{1f} = \tau_1(\lambda_f, \rho_f)$ and $\tau_{2f} = \tau_2(\lambda_f, \rho_f)$ be the lower and upper thresholds given that f channels are free. From the thresholds we can compute the probabilities needed in the Markov chain formulation using:

$$\alpha_\delta(f) = \sum_{\theta=0}^{\infty} \frac{e^{-\lambda_f}\lambda_f^\theta}{\theta!}\left(e^{-\tau_{1f}/(\theta\sigma_d^2+1)} - e^{-\tau_{2f}/(\theta\sigma_d^2+1)}\right) \tag{32}$$

$$\gamma_\delta(f) = e^{-\lambda_f}\lambda_f(e^{-\tau_{1f}/(\sigma_d^2+1)} - e^{-\tau_{2f}/(\sigma_d^2+1)}) \tag{33}$$

Optimization now involves searching for the optimal vector of cost-ratios:

$$\boldsymbol{\rho}_0 = \arg\max_{\boldsymbol{\rho}} \ \eta_\delta(\boldsymbol{\rho}) \ . \tag{34}$$

The highest throughput that can be obtained is, thus, $\eta_\delta(\boldsymbol{\rho}_0)$. We would like to see how other designs compare to the optimal one. In the next section, we present some alternate designs and obtain expressions through which we can determine the throughput they deliver.

V. Sub-optimal Designs

The cross-layer function can be viewed in split form; a simple MAC layer entity which merely ACKs or NACKs depending on the decisions of a physical detector. The physical layer detector performs hypothesis testing on the pair,

$$\begin{aligned} H_0 &: \ \Theta \in \{0, 2, 3, \cdots\} \stackrel{\Delta}{=} \Lambda_0 \\ H_1 &: \ \Theta \in \{1\} \stackrel{\Delta}{=} \Lambda_1 \ . \end{aligned} \tag{35}$$

In this setting, the function δ can be viewed as a physical layer decision rule, a detector[3], on the binary hypotheses pair. The MAC entity follows the procedure of acknowledging when hypothesis H_1 is held to be true and NACKing when H_0 is held true.

[3]We will be using the term *detector* interchangeably with the phrase *decision rule*.

We could also consider decision rules on multiple hypotheses -

$$\mathcal{H}_\theta : \Theta = \theta \ . \tag{36}$$

Such decision rules could prove useful in cases in which a single channel can support multiple users. For a collision channel, a MAC based on multi-hypotheses decision rules ACKs a channel only when $\mathcal{H}_1$ is true.

By considering the problem on split-layers, we sacrifice the optimality achieved with the truly cross-layer design of the previous section. However, the schemes that we consider in this section have certain advantages as we shall see.

A. The Multi-hypotheses MAP

The optimal MAC function does not admit a closed form expression for thresholds; numerical optimization must be carried out for different traffic rates and available free channels. We consider a detector for which the decision regions can be determined in closed form. The detector is actually a multi-hypotheses MAP detector which optimally detects the number of users attempting access based on the a posteriori probabilities of each $\theta \in \Theta$. The detector gives $\hat{\theta} = \arg\max_\theta p(y|\theta)p(\theta|f)$. The MAC protocol can then make a decision based on $\hat{\theta}$. $\mathcal{H}_1$ will be held to be true when $\theta = 1$ has the maximum a posteriori probability amongst all $\theta \in \Theta$, i.e., when:

$$\arg\max_\theta p(y|\theta)p(\theta|f) = 1 \ . \tag{37}$$

The Multi-hypotheses MAP detector also leads to decision region as given in (30)-(31); it can be shown that $\tau_1(\lambda_f)$ and $\tau_2(\lambda_f)$ are determined by (see [12] for a proof):

$$\tau_1(\lambda_f) = \max\{\tilde{\tau}_1, 0\}, \tau_2(\lambda_f) = \inf\{\tilde{\tau}_\theta : \theta > 1\} \tag{38}$$

where

$$\tilde{\tau}_1(\lambda_f) = \frac{\sigma_d^2+1}{\sigma_d^2}\log\left(\frac{\sigma_d^2+1}{\lambda_f}\right) \tag{39}$$

$$\tilde{\tau}_\theta(\lambda_f) = \frac{(\sigma_d^2+1)(\theta\sigma_d^2+1)}{(\theta-1)\sigma_d^2}\log\left(\frac{(\theta\sigma_d^2+1)\theta!}{\lambda_f^{(\theta-1)}(\sigma_d^2+1)}\right) \ . \tag{40}$$

The probabilities $\alpha(f)$ and $\gamma(f)$ can now be computed (via (32)-(33)) with $\alpha(f)$ and $\gamma(f)$ depending only on λ_f to obtain the throughput. Since the thresholds are fixed directly by σ_d^2 and λ_f, this detector does not involve any optimization. One problem is that if λ_f is large enough, τ_2 could become negative. In this case, the detector never ACKs a channel request, essentially leading to a system breakdown.

B. A Single Threshold Detector

We consider now the class of single threshold detectors which acknowledge a channel when the power exceeds a given threshold (the upper threshold $\tau_2 = \infty$). Let τ_{sf} be the single threshold when f channels

are free. Then the detector is given by:

$$\delta(y,f) = \begin{cases} 1 & \tau_{sf} < y \\ 0 \text{ or } 1 & y = \tau_{sf} \\ 0 & y > \tau_{sf} \ . \end{cases} \tag{41}$$

In this case, we want to find the optimal among the $\boldsymbol{\tau} = (\tau_{s1}, \ldots, \tau_{sN})$, so that

$$\boldsymbol{\tau}_0 = \arg\max_{\boldsymbol{\tau}} \ \eta(\boldsymbol{\tau}) \ . \tag{42}$$

Given $\boldsymbol{\tau}$, we can obtain $\alpha(f)$ and γ_f required in the throughput expression. We can also evaluate the detector operating characteristics under the constraint of having a single threshold. The false-alarm and detection probabilities are directly related to $\boldsymbol{\tau}$ by:

$$P_D(f) = e^{-\tau_{sf}/(\sigma_d^2+1)} \tag{43}$$

$$P_F(f) = \frac{1}{1-\lambda_f e^{-\lambda_f}} \sum_{\theta \neq 1} \frac{\lambda_f^\theta e^{-\lambda_f}}{\theta!} e^{-\tau_{sf}/(\theta\sigma_d^2+1)} \ . \tag{44}$$

We can consider the optimization in terms of the false-alarm probabilities as against the thresholds themselves. The optimization problem now becomes finding $\mathbf{P}_{F_o} = (P_{F_o}(1), \ldots, P_{F_o}(N))$ such that:

$$\mathbf{P}_{F_o} = \arg\max_{\mathbf{P}_F} \ \eta(\mathbf{P}_F) \ . \tag{45}$$

where $\mathbf{P}_F = (P_F(1), \ldots, P_F(N))$.

C. UMP and ML detectors

In determining the thresholds for the optimal and Multihypotheses MAP designs, we require knowledge of the traffic statistics. We could use detectors which do not require prior probabilities when the knowledge of traffic statistics becomes unreliable. A Uniformly Most Powerful (UMP) test with parameter μ given in [10], does not require the priors to be known. The test can be written as:

$$\delta(y) = \begin{cases} 1 & \tau_1(\mu) < y < \tau_2(\mu) \\ 0 \text{ or } 1 & y = \tau_1(\mu) \text{ or } y = \tau_2(\mu) \\ 0 & y > \tau_1(\mu) \text{ or } y < \tau_2(\mu) \end{cases} \tag{46}$$

where τ_1 and τ_2 satisfy:

$$E(\delta(y)|\theta = 0) = E(\delta(y)|\theta = 2) = \mu \ . \tag{47}$$

The condition above leads to the following expressions from which we can evaluate the thresholds:

$$(\mu + e^{-\tau_2})^{1/(2\sigma^2+1)} = \mu + e^{-\tau_2/(2\sigma^2+1)} \tag{48}$$

$$e^{-\tau_1} = \mu + e^{-\tau_2} \ . \tag{49}$$

We can then compute the throughput having obtained $\alpha(f)$ and $\gamma(f)$. Again, the parameter might depend on the number of free channels. The search must, therefore, be made over $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)$. The optimization involves finding:

$$\boldsymbol{\mu}_o = \arg\max_{\boldsymbol{\mu}} \ \eta(\boldsymbol{\mu}) \ . \tag{50}$$

We could also use the Maximum Likelihood test for multihypotheses to determine thresholds when the priors are not known. The thresholds for the ML detector are given by:

$$\tau_1(\lambda_f) = \frac{\sigma_d^2+1}{\sigma_d^2}\log(\sigma_d^2+1) \tag{51}$$

$$\tau_2(\lambda_f) = \frac{(\sigma_d^2+1)(2\sigma_d^2+1)}{\sigma_d^2}\log\left(\frac{2\sigma_d^2+1}{\sigma_d^2+1}\right) \ . \tag{52}$$

As in the case of the Multihypotheses MAP, we do not have any degrees of freedom to optimize throughput. The number of free channels immediately fixes $\alpha(f)$ and $\gamma(f)$.

VI. Extensions

A. Multiple Measurements

We expect throughput to increase with SNR. However, as will be seen through simulation results, under fading, the throughput saturates without reaching the ideal value. This is because, for a high SNR, we can only expect to make no error in judging the presence or absence of user(s). However, errors will still be made in distinguishing the presence of exactly one user.

We can hope to achieve ideal detection by making decisions based on multiple independent measurements of the reservation requests. Such multiple measurements could be obtained in the same slot or be spread out over consecutive slots depending on how fast the fading occurs.

Let the sampled, despread and match-filtered received vector obtained after n measurements be $\mathbf{Z} = [Z_1 \ Z_2 \ldots Z_n]$, Z_i's are i.i.d. $\mathcal{CN}(0, \theta\sigma_d^2+1)$ where σ_d^2 is the SNR (the number of users attempting access is assumed to be θ). Since Z_i's are normal, a sufficient statistic is the sum of power of the received components:

$$Y = T(\mathbf{Z}) = \sum_{i=1}^{n} |Z_i|^2 \ . \tag{53}$$

Conditioned on $\Theta = \theta$, $|Z_i|^2 \sim Exp(1/(\theta\sigma_d^2+1))$, thus, the conditional distribution of y is given by:

$$p(y|\theta) = \frac{y^{(n-1)}}{(\theta\sigma_d^2+1)^n(n-1)!}e^{\left(\frac{-y}{\theta\sigma_d^2+1}\right)} \ . \tag{54}$$

The ratio of aposteriori probabilities for the optimal MAC design is given by:

$$\frac{l_1(y;f))}{l_0(y;f)} = \frac{\frac{1}{(\sigma_d^2+1)^n}\exp\left(\frac{-y}{\sigma_d^2+1}\right)\lambda_f}{\sum_{\theta\neq 1}\frac{1}{(\theta\sigma_d^2+1)^n}\exp\left(\frac{-y}{\theta\sigma_d^2+1}\right)\frac{\lambda_f^\theta}{\theta!}} \ . \tag{55}$$

For this case too, the decision regions are of the form in (30),(31) and have to be obtained numerically.

We can also consider the Multihypotheses MAP detector for the multiple measurements case for which the thresholds are given by:

$$\tilde{\tau}_1(\lambda_f) = \frac{\sigma_d^2+1}{\sigma_d^2}\log\left(\frac{(\sigma_d^2+1)^n}{\lambda_f}\right) \quad (56)$$

$$\tilde{\tau}_\theta(\lambda_f) = \frac{(\sigma_d^2+1)(\theta\sigma_d^2+1)}{(\theta-1)\sigma_d^2}\log\left(\frac{(\theta\sigma_d^2+1)^n\theta!}{\lambda_f^{(\theta-1)}(\sigma_d^2+1)^n}\right) . \quad (57)$$

and

$$\tau_1 = max\{\tilde{\tau}_1, 0\}, \quad \tau_2 = \inf\{\tilde{\tau}_\theta : \theta > 1\} . \quad (58)$$

For the ML detector, the thresholds are simply n times of the ones for the single measurement case

$$\tau_1(\lambda_f) = n\frac{\sigma_d^2+1}{\sigma_d^2}\log(\sigma_d^2+1) \quad (59)$$

$$\tau_2(\lambda_f) = n\frac{(\sigma_d^2+1)(2\sigma_d^2+1)}{\sigma_d^2}\log\left(\frac{2\sigma_d^2+1}{\sigma_d^2+1}\right) . \quad (60)$$

Having obtained the two thresholds, the probabilities required for evaluating η can be obtained from:

$$\gamma(f) = \lambda_f e^{-\lambda_f}\left(e^{\frac{-\tau_{1f}}{(\sigma_d^2+1)}}\sum_{j=0}^{n-1}\frac{1}{j!}\frac{(\tau_{1f})^j}{(\sigma_d^2+1)^j} - e^{\frac{-\tau_{2f}}{(\sigma_d^2+1)}}\sum_{j=0}^{n-1}\frac{1}{j!}\frac{(\tau_{2f})^j}{(\sigma_d^2+1)^j}\right) \quad (61)$$

$$\alpha(f) = \sum_{\theta=0}^{\infty}\frac{e^{-\lambda_f}(\lambda_f)^\theta}{\theta!}\left(e^{\frac{-\tau_{1f}}{(\theta\sigma_d^2+1)}}\sum_{j=0}^{n-1}\frac{1}{j!}\frac{(\tau_{1f})^j}{(\theta\sigma_d^2+1)^j} - e^{\frac{-\tau_{2f}}{(\theta\sigma_d^2+1)}}\sum_{j=0}^{n-1}\frac{1}{j!}\frac{(\tau_{2f})^j}{(\theta\sigma_d^2+1)^j}\right) . \quad (62)$$

B. Delay

When L is a system design parameter, one has to deal with a trade-off that exists between throughput and channel utilization. Channel utilization increases with increasing L. But an increasing L implies that the base station cannot service as many access requests per slot as before, leading to a decrease in throughput. This in turn would lead to longer delays for newly generated packets.

We can get a measure of the delay incurred by calculating the expected number of (re)transmissions of the reservation request a user has to make before it transmits data [16]. Since the rate of requests (λ) and the rate of those that are successful (η) are known the expected number of (re)transmissions required can be computed from:

$$R = \lambda/\eta . \quad (63)$$

Obviously, for the same arrival rate, having a better throughput also means less number of access attempts before data transmission. Note that in the infinite user single buffer scenario, queueing considerations do not arise and, therefore, we do not deal with delay introduced by queueing.

VII. Numerical Results

In this section we present the results of numerical evaluation of the throughput and channel utilization obtained with the various designs. We will also consider aspects such as the dependence of throughput on

the SNR (σ_d^2), the packet length, L, and multiple measurements. Finally, we will look into the trade-off between channel utilization and throughput for variations in L. We will be working with $N = 2$ receiver codes throughout this section.

A. Comparison of designs

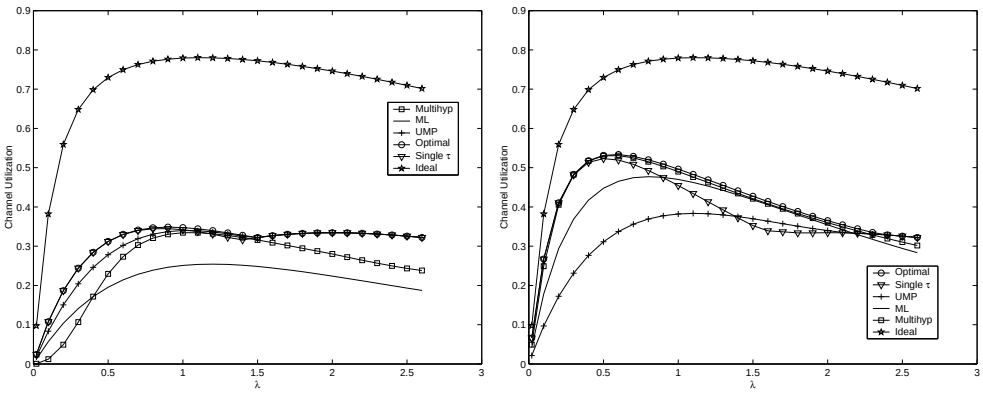

Fig. 4. Normalized Channel Utilization vs. Arrival rate for various designs with $N = 2$, L equal to 10 SNR = 0dB(left), SNR = 10dB (right). Optimal -$\circ$, Multi-hypotheses MAP-$\square$, Single threshold -$\bigtriangledown$, UMP -+, ML-plain

Plotted in Fig. 4 is the channel utilization for various designs for various SNRs with L equal to 10. In this case, the ideal scenario would be for a succession of user pairs to occupy the two channels for the duration of L slots corresponding to an arrival rate of about 0.2 users per slot. We restrict our simulation results to arrival rates of upto 2.5.

For low SNR, the single threshold and UMP detectors are close to the optimal achievable. The performance of the multihypotheses MAP is not encouraging for low SNR, but for high SNR, it gives channel utilization close to that achieved using the optimal MAC design. We would expect the decisions made by the Multihypotheses MAP detector to be reliable at high SNR, which is what we see. Practically speaking, we would like to keep the SNR for the reservation requests to be as high as possible bearing in mind how crucial the reservation phase is to the performance of the whole system.

Knowledge of the arrival rate does improve the performance as is seen in the difference in performance of the UMP/ML based schemes and the others. Neither UMP nor ML based designs takes into consideration the knowledge of the arrival rate. As such, these schemes can be considered to be acting based solely on the received signal power. For the single threshold plots, we can say that knowledge of the arrival rate has been assumed while optimizing to find the best single threshold.

B. Increasing the packet length

The channel utilization can be expected to go up as L increases. However, the utilization obtained need not be arbitrarily close to the ideal. The limit as L grows large depends on the fading conditions and

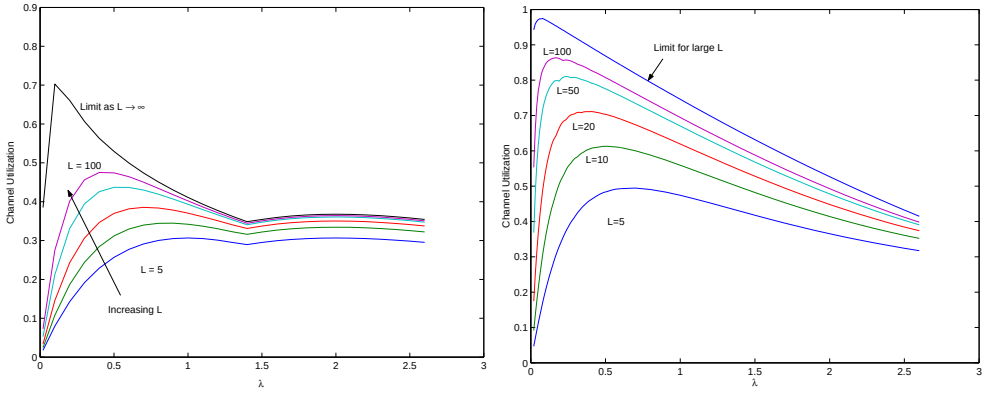

Fig. 5. Effect of increasing L on channel utilization ($N = 2$). Optimal design, SNR = 0dB(left), SNR = 20dB (right)

the detector ROCs. From Fig. 5, it is evident that increasing L leads to increased channel utilization. However, the channel utilization does not increase beyond a limit as computed in (73) of the appendix, and the limit is reached only gradually, as can be seen from the figure.

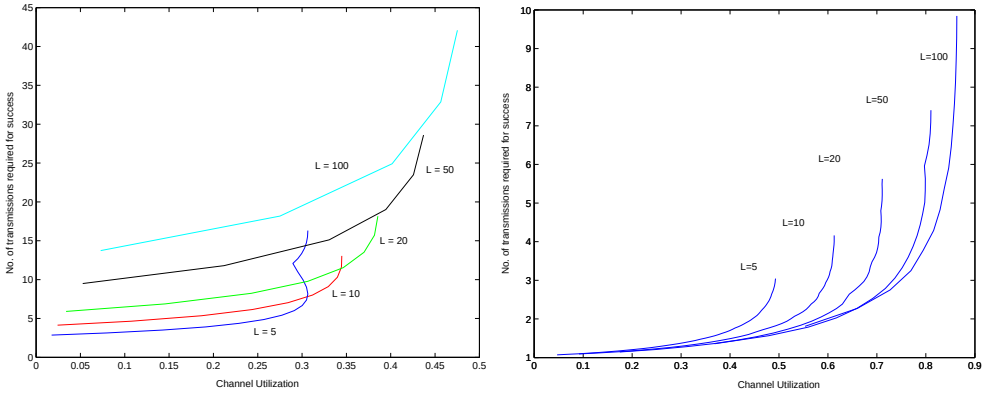

Fig. 6. Optimal design: Trade-off of channel utilization and number of retransmissions required. $N = 2$, SNR = 0dB (left), SNR = 20dB(right)

Fig. 6 depicts the trade-off that exists between channel utilization and the number of (re)transmissions required. The plots are related to the plots in Fig. 5 and have been obtained for arrival rates less than or equal to the ones corresponding to the peak channel utilization for the respective L. For arrival rates higher than the one with peak channel utilization, the number of retransmissions required will be higher with less channel utilization. The points on the trade-off curves for rates higher than those corresponding to peak channel utilization have not been plotted. The trade-off is especially severe when the SNR is low, higher channel utilization coming at the price of increased number of retransmissions required and therefore, incurring more delay. The trade-off is almost non-existent for higher SNR, though the channel utilization peaks for a lower offered rate, λ, as is seen in Fig. 5.

C. Increasing SNR and Number of Measurements

As we commented in Sec.VI-A, throughput and, hence, channel utilization saturate with increasing SNR, as seen in Fig. 7. With increased number of measurements, channel utilization close to the ideal can be reached. Fig. 8 shows how the channel utilization increases with the number of measurements for Optimal and Multihypotheses designs at SNRs of 10dB and 20dB. It is can be seen that increasing the number of measurements does not change the saturation effect w.r.t. SNR, as the plots for SNRs of 10dB and 20dB are very much close together. Again, the Multihypotheses detector gives performance close to the optimal one at high SNR.

In Fig. 9 we see the how the channel utilization converges towards the ideal for the Multihypotheses detector operating at SNR 20dB. Plotted is the ratio of the Multihypotheses channel utilization to the ideal channel utilization versus the number of measurements. The convergence seems to be of the form $(1-e^{-x})$ with the exponent decreasing for increasing λ. The exponential form of convergence is to be expected, as with multiple number of measurements, the decision error probabilities go down exponentially. What is surprising is the varying exponent for different λ's. The ideal channel utilization is achieved for $\lambda = 0.02$ with less than 5 measurements. However, for $\lambda = 2.3$, ideal channel utilization has not been reached even with 40 measurements.

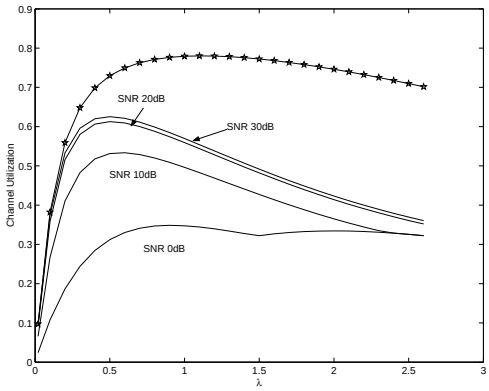

Fig. 7. Saturation of channel utilization with increasing SNR. Optimal MAC design, $L = 10$.(Ideal-$\circ$)

D. A Comment on the Kinks

Notice the kinks in the plots for the optimal, UMP and Single threshold designs for 0dB SNR (see Fig. 4). The behaviour is unusual as we expect the variation of the performance to be smooth with respect to the arrival rate for an optimal design. We checked the correctness of our results by carrying an extended simulation, where the actual contention process itself was simulated, not just the random variables pertaining to the decisions at the PHY/MAC. The set-up had $M = 100$ users each having a (re)transmission probability of g chosen so as to correspond to a given arrival rate λ, i.e., g was chosen

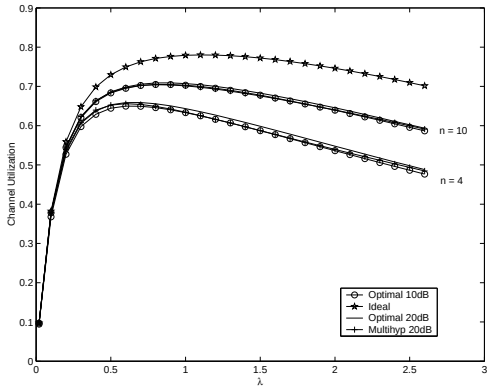

Fig. 8. Approaching the ideal - Channel utilization with multiple measurements for the optimal and Multihypotheses designs $(L = 10, N = 2)$.

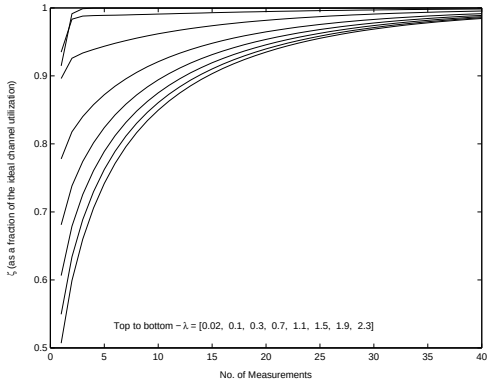

Fig. 9. Approaching the ideal - Channel utilization as a fraction of the ideal vs. the number of measurements for $\lambda = [0.02\ 0.1\ 0.3\ 0.7\ 1.1\ 1.5\ 1.9\ 2.3]$, SNR = 20dB.

such that $Mg = \lambda$. The reservation signal strength at each free channel was generated to have a Rayleigh distribution based on the number of users selecting the correspoding signature. A channel was ACKed if the signal strength fell within the two thresholds obtained through the optimization process as described in Sec.IV. Fig. 10 shows an excellent agreement between the plots obtained from the extended simulation and numerical computation. Similar agreement was obtained for the other plots that have been presented here.

It is observed that for arrival rates greater than the point corresponding to the kink, the optimal policy is to always NACK when only one channel is free ($\alpha_1 \approx 0$) and always ACK when both channels are free ($\alpha_2 \approx 1$). This fact is also depicted in Fig. 10. For high arrival rates, the number of cases of multiple users attempting access increases, but low SNR means that the detector cannot reliably decide whether exactly one user or multiple users are attempting access. It is as if, for low SNR, the detectors give up on the information available in the signal strength and let the system revert back to the elementary slotted ALOHA random access to achieve optimum performance. The peak throughput (and hence channel

utilization) for such a policy occurs when $\lambda_2 = 1$[4], *i.e.,* when $\lambda = 2$, as is seen in the figure. Note that such a policy ($\alpha_1 \approx 0, \alpha_2 \approx 1$) is achievable with single threshold or UMP detectors, making it possible for them to give optimal performance for high arrival rates at low SNR. With the single threshold detector, such a policy seems to be the best at even higher SNR (notice the slight kinkiness in the plot for the single threshold detector for SNR = 10dB).

The kinks seem to result because the "optimal" MAC design we have considered does not take the *entire* system information into consideration. The system state is described by $\mathbf{n}_t$ as we described during the Markov chain formulation. However, the MAC designs that we have considered utilize only the number of free channels into consideration. Consider two situations for the case $N = 2$ both of which have one free channel at the time of observation: (a) the other channel was occupied only in the previous slot, (b) the other occupied channel will become free by the end of the next slot. The MAC designs that we have considered will treat the above situations similarly. However, in case of situation (b) one might want to delay making decisions until the next slot when both channels will be free, thus reducing the risk of collision. We, therefore, should have treated the two situations differently. Unfortunately, that would entail considering an optimization on a far larger scale. The resulting suboptimality is, we conjecture, seen in the presence of the kinks in the plots.

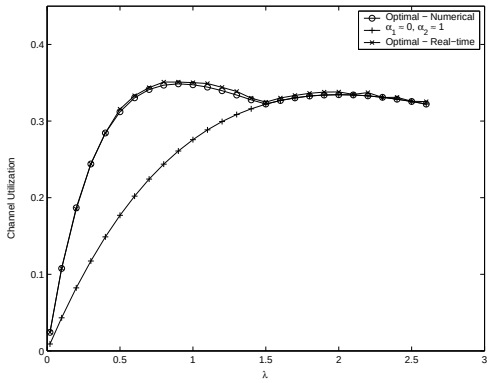

Fig. 10. The kinks in the plots - For SNR = 0dB, Optimal channel utilization (Numerical and Extended (real-time)) and the channel utilization obtained with $\alpha_1 \approx 0, \alpha_2 \approx 1$

VIII. Conclusion

For a system employing reservation for multi-access over multiple channels, we have given a framework wherein the performance at MAC level can be analyzed and optimized under fading channel conditions. Based on this framework we have given an optimal Neyman-Pearson-like MAC design which utilizes knowledge about the number of free codes in its decision making process. The design is characterized by the acknowledgement probability given the number of free channels optimized with respect to the

[4]Recall that it is with arrival rate equal to unity that the peak throughput of slotted ALOHA is achieved [2].

throughput function. The design is not truly optimal as it does not use the system state information in its entirety. The design still provides a more realistic benchmark of performance (as compared to the performance in a ideal situation) to compare the performance of various other designs including MAC designs based on classical physical layer detectors such as UMP, ML, MAP etc. Knowledge of the traffic statistics and the partial system state (number of free channels) improves the performance as seen by comparing the performance of the optimal design with that of a design based on the ML detector.

In this paper, we have given the closed form expression for the throughput with $N = 2$ receiver codes and an arbitrary packet length L. Throughput with other parameters can be obtained, but must be evaluated numerically. However, the transition probability matrix is sparse and structured because of the fact that the memory over L slots is incorporated in the definition of the states. This fact may be used in evaluating the performance through methods employing sparse matrices. The dimensionality of optimization involved while computing the optimal performance may be reduced by appropriate classification of the states.

A number of issues are not addressed in this paper. For example, the framework used in this paper does not allow analytical treatment of stability either in the finite user infinite buffer or the infinite user single buffer regime. One hopes that stability results similar to the case of slotted ALOHA can be obtained. Also missing is a justification for the aggregate attempt rate being Poisson. For the case of slotted ALOHA Ghez et.al. [8] proved that it is indeed possible to obtain Poisson aggregate attempts (which, in fact, optimize the throughput) for any input traffic statistics. Unfortunately, we cannot claim to have achieved any such connection between the input traffic (new arrivals) and aggregate traffic. The Poisson assumption also means that we have implicitly assumed an infinite user population restricting, perhaps, the applicability of the results to networks with a large number of nodes. In computing the channel utilization, we have also ignored the failure in data transmission. This omision, however, does not affect the optimality of the MAC protocol, and it is easy to take into account the effect of fading in the computation using existing results.

ACKNOWLEDGMENTS

We thank the reviewers for their detailed comments. Special thanks are due to one of the reviewers for pointing out the limitations of the critical assumption of Poisson traffic.

APPENDIX

I. STATIONARY DISTRIBUTION FOR $N = 2$, GENERAL L

For $N = 2$, we can partition the state space into four groups:

$$\mathcal{S} = \mathcal{S}_0 \oplus \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \mathcal{S}_3 \ . \tag{64}$$

$\mathcal{S}_0$ contains the state S_0. $\mathcal{S}_0 \stackrel{\Delta}{=} \{[0 \dots 0]\}$. $\mathcal{S}_1$ contains states with a single locked channel:

$$\mathcal{S}_1 \stackrel{\Delta}{=} \{[0 \dots 01], \dots, [10 \dots 0]\} = \{S_1, \dots, S_L\} \ . \tag{65}$$

$\mathcal{S}_2$ contains states with two simultaneously locked channels:

$$\mathcal{S}_2 \stackrel{\Delta}{=} \{[0\dots02],\dots,[20\dots0]\} = \{S_{L+1},\dots,S_{2L}\} \ . \tag{66}$$

$\mathcal{S}_3$ contains states with two channels locked at different slots:

$$\begin{aligned}\mathcal{S}_3 &\stackrel{\Delta}{=} \{[0\dots,011],[0,\dots0101],\dots,[10\dots01],[0\dots0110],\dots,\dots,[10\dots10]\} \\ &= \{S_{2L+1},\dots,S_{3L-1},S_{3L},\dots,S_{|\mathcal{S}|-1}\} \ .\end{aligned}$$

It will later be shown that the states within the same group have equal probabilities in the stationary distribution. It can also be seen that there is a 'translational invariance' between the components describing the states within each group. For example, with $L=3$, [011] and [101] belong to the same group - $\mathcal{S}_3$[5].

It can be verified that the Markov chain described by $\mathbf{n}_t$ (see (5) in Sec.III) is aperiodic and irreducible. Consider Fig. 2 showing the Markov chain for $N=2$ and $L=3$, with the states being numbered as in Table I. The self-loop on S_0 makes the chain aperiodic, and it is possible to go from one state to any other in a finite number of steps with positive probability when $0<\alpha_1,\alpha_2<1$. The chain is also finite and, thus, has a stationary distribution. Let $\boldsymbol{\pi}_\delta = [\pi_0,\cdots,\pi_{|\mathcal{S}|-1}]$ be the stationary distribution. The stationary distribution must satisfy the following conditions (obtained by looking at the transitions into the states on the LHS for each equation):

$$\begin{aligned}
\pi_0 &= \beta_2^2\pi_0 + \beta_1\pi_L + \pi_{2L} & \\
\pi_1 &= 2\alpha_2\beta_2\pi_0 + \alpha_1\pi_L & \\
\pi_i &= \beta_1\pi_{i-1} + \pi_j & \forall\ 1<i\le L, \text{for some } j(i)>2L \\
\pi_{L+1} &= \alpha_2^2\pi_0 & \\
\pi_i &= \pi_{i-1} & \forall\ L+1<i\le 2L \\
\pi_i &= \alpha_1\pi_{i-2L} & \forall\ 2L<i<3L \\
\pi_i &= \pi_j & \forall\ 3L\le i<|\mathcal{S}|, \text{for some } 2L<j(i)<3L
\end{aligned} \tag{67}$$

Claim: The stationary distribution is given by:

$$\pi_i = \begin{cases} \frac{1}{L\alpha_2^2+2L\alpha_2\beta_2/\beta_1+2\binom{L}{2}\alpha_1\alpha_2\beta_2/\beta_1+1} \stackrel{\Delta}{=} p_0 & i=0 \\ 2\alpha_2\beta_2\pi_0/\beta_1 \stackrel{\Delta}{=} p_1 & \forall\ 1\le i\le L \\ \alpha_2^2\pi_0 \stackrel{\Delta}{=} p_2 & \forall\ L<i\le 2L \\ 2\alpha_1\alpha_2\beta_2\pi_0/\beta_1 \stackrel{\Delta}{=} p_3 & \forall\ 2L<i<|\mathcal{S}| \ . \end{cases} \tag{68}$$

Proof: It is easy to see that the distribution above satisfies the identities listed in (67). For example, plugging the values in the second identity, we have to check if:

$$p_1 = 2\alpha_2\beta_2 p_0 + \alpha_1 p_1 \ . \tag{69}$$

[5]For general N, though groups with states having the same stationary probabilities will certainly exist, states with translational invariance need not necessarily have the same stationary probabilities.

This is equivalent to checking if $\beta_1 p_1 = 2\alpha_2\beta_2 p_0$. But since $p_1 = 2\alpha_2\beta_2 p_0/\beta_1$, we know that it holds.

Thus, the distribution in (68) is a stationary distribution. We know that for an aperiodic, irreducible, finite state Markov chain, there exists a unique stationary distribution. Thus, the unique stationary distribution is as given in (68).

Note that a stationary distribution does not exist when $\alpha_1 = 0, \alpha_2 = 1$ but (68) gives us $p_0 = p_2 = \frac{1}{L+1}$ which is not entirely meaningless considering that the fact that, for this case, the chain merely cycles through $S_0, S_{L+1}, \cdots, S_{2L}$ in order. The reader can verify that, fortunately, the other cases, where the stationary distribution does not exist, do not make practical sense from a system point of view.

II. Throughput with $N = 2$

Plugging in the values obtained from the stationary distribution in (12), we get for $N = 2$:

$$\eta = 2\gamma_2\pi_0 + \sum_{i=1}^{L}\gamma_1\pi_i = 2\gamma_2 p_0 + L\gamma_1 p_1 \tag{70}$$

$$= \frac{2\gamma_2 + 2L\gamma_1\alpha_2\beta_2/\beta_1}{L\alpha_2^2 + 2L\alpha_2\beta_2/\beta_1 + 2\binom{L}{2}\alpha_1\alpha_2\beta_2/\beta_1 + 1} \tag{71}$$

where it can be seen that $2\gamma_2$ is the expected number of access attempts that are successful when in state S_0 and γ_1 is the corresponding value when in a state belonging to $\mathcal{S}_1$. Note that only p_0 and p_1 appear in the expression for throughput, p_2 and p_3 do not figure because no contention occurs when the system state belongs to $\mathcal{S}_2$ or $\mathcal{S}_3$.

A. Channel Utilization for large L

Substituting the expression for η obtained in (71) we can directly evaluate the limit for (15) as $L \to \infty$. We have

$$\zeta = \frac{L(2\gamma_2 + 2L\gamma_1\alpha_2\beta_2/\beta_1)}{2(L\alpha_2^2 + 2L\alpha_2\beta_2/\beta_1 + 2\binom{L}{2}\alpha_1\alpha_2\beta_2/\beta_1 + 1)} \tag{72}$$

Note that we get two different limits for the cases $\alpha_1 \neq 0$ and $\alpha_1 = 0$. When $\alpha_1 > 0$, collecting the terms with L^2 in the numerator and denominator, we have $\zeta = \gamma_1/\alpha_1$. When $\alpha_1 = 0$, we have $\gamma_1 = 0$ and, therefore, $\zeta = \gamma_2/(\alpha_2(1+\beta_2))$. Optimizing over the ROCs we get:

$$\zeta \to \max\left(\max\left(\frac{\gamma_1}{\alpha_1}\right), \max\left(\frac{\gamma_2}{\alpha_2(1+\beta_2)}\right)\right) \tag{73}$$

References

[1] S. Adireddy and L. Tong. Medium Access Control using Channel State Information for Large Sensor Networks. In *Accepted to MMSP'02*, St. Thomas, US Virgin Islands, December 2002. `http://people.ece.cornell.edu/ltong/`.

[2] D.P. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, 1992.

[3] B.Hajek, A.Krishna, and R.O.LaMaire. On the Capture Probability for a Large Number of Stations. *IEEE Trans. Communications*, 45(2):254–260, February 1997.

[4] A. Butala and L. Tong. Dynamic Channel Allocation and Optimal Detection for MAC in CDMA Ad hoc Networks. In *Proc. 36th Asilomar Conf. Signals, Systems & Computers*, Pacific Grove, CA, Nov 2002.

[5] A. P. Butala. A medium access control protocol for cdma ad hoc networks. Master's thesis, Cornell University, Ithaca, NY, Aug 2001.

[6] A. Chockalingam, W. Xu, M. Zorzi, and L.B. Milstein. Throughput-Delay Analysis of a Multichannel Wireless Access Protocol. *IEEE Transactions on Vehicular Technology*, 49:661–671, March 2000.

[7] S. Ghez, S. Verdú, and S. Schwartz. Stability Properties of Slotted Aloha with Multipacket Reception Capability. *IEEE Trans. Automatic Control*, 33(7):640–649, July 1988.

[8] S. Ghez, S. Verdú, and S. Schwartz. Optimal Decentralized Control in the Random Access Multipacket Channel. *IEEE Trans. Automatic Control*, 34(11):1153–1163, Nov. 1989.

[9] L. Kleinrock and F.A. Tobagi. Packet Switching in Radio Channels: Part I–Carrier Sensing Multiple-Access Modes and Their Throughput-Delay Characteristics. *IEEE Trans. Comm.*, COM-23(12):1400–1416, Dec. 1975.

[10] E.L. Lehmann. *Testing Statistical Hypotheses*. John Wiley & Sons, Inc., New York, NY, 1959.

[11] Che-Le Lin. Investigation of 3^{rd} Generation Mobile Communication RACH Transmission. In *Proc. of Vehicular Tech. Conf.*, volume 2, pages 662 –667, Fall 2000.

[12] A. Maharshi, L. Tong, and A. Swami. Cross-layer Designs of Multichannel Reservation MAC under Rayleigh Fading. Technical Report ACSP-TR-09-02-01, Cornell Univeristy ACSP Lab, September 2002. `http://acsp.ece.cornell.edu/pubJ.html`.

[13] M.Zorzi. Mobile Radio Slotted ALOHA with Capture, Diversity and Retransmission Control in the Presence of Shadowing. *Wireless Networks*, 4:379–388, August 1998.

[14] M.Zorzi and R.Rao. Capture and Retransmission Control in Mobile Radio. *IEEE JSAC*, 12(8):1289–1298, October 1994.

[15] M. B. Pursley. The Role of Spread Spectrum in Packet Radio Networks. *Processings of The IEEE*, 75(1):116–134, Jan. 1987.

[16] F.A. Tobagi and L. Kleinrock. Packet Switching in Radio Channels: Part II–The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution. *IEEE Trans. Comm.*, COM-23(12):1417–1433, Dec. 1975.

[17] L. Tong, Q. Zhao, and G. Mergen. Multipacket reception in random access wireless networks: from signal processing to optimal medium access control. *IEEE Communication Magazine*, 39(12):108–112, Nov 2001. *Special issue on Design Methodologies for Adaptive and Multimedia Networks.*

[18] 3^{rd} Generation Partnership Project TS 25.321. *MAC Protocol Specification*. `http://www.etsi.org`, 2001.

[19] M. K. Tsatsanis, R. Zhang, and S. Banerjee. Network Assisted Diversity for Random Access Wireless Networks. *IEEE Trans. Signal Processing*, SP-48(3):702–711, March 2000.

[20] I. N. Vukovic and T. Brown. Performance Analysis of the Random Access Channel (RACH) in WCDMA. In *Proc. of Vehicular Tech. Conf.*, volume 1, pages 532–536, Spring 2001.

[21] Q. Zhao and L. Tong. A Multiqueue Service Room MAC protocol for Wireless Networks with Multipacket Reception. To appear in *IEEE/ACM Trans. on Networking*, Feb., 2003, available at `http://people.ece.cornell.edu/ltong/pubJ.html`.

[22] Q. Zhao and L. Tong. Semi-blind collision resolution in random access ad hoc wireles networks. *IEEE Trans. Signal Processing*, 48(9), Sept. 2000.

List of Figures

List of Tables