Cross Layer Design for Multiaccess Communication Over Rayleigh Fading Channels

Vidyut Naware, *Student Member, IEEE*, and Lang Tong, *Fellow, IEEE*

Abstract— An information theoretic queueing model is proposed in a wireless multiple access communication setup. The proposed symmetric N user model captures physical layer parameters such as the encoding rate, transmit power and Medium Access Control (MAC) layer metrics such as queue stability. Two alternative medium access strategies are considered: centralized scheduling and ALOHA. Next, a cross-layer approach is taken wherein the maximum stable throughput of the system is achieved by a joint optimization over the MAC parameters (*viz.,* scheduling set size with scheduling and transmission probability with ALOHA) and the encoding rate. Performance comparisons with traditional layered designs are given. It is shown that in the low and high SNR regimes, layered designs are close to optimal whereas in the moderate SNR range, cross-layer designs outperform layered schemes. Exact characterizations of the "low" and "high" SNR regimes are given quantitatively. It is also shown that ALOHA with transmission probability one is optimal in the low SNR regime.

Index Terms— ALOHA, cross-layer design, information theory, multiple access, queueing, scheduling, stability.

I. INTRODUCTION

THE problem of multiple access to a shared communication resource somehow has evaded an approach which jointly addresses the issues of noise, interference, and the "bursty" arrivals of messages [1], [2]. Broadly, there have been two main approaches to formulate the multiaccess communication problem: the network theoretic approach (initiated by Abramson [3]) and the multiaccess information theoretic approach (due to Liao [4]). The network theoretic model accounted for the burstiness of message arrivals and interference but largely neglected the effects of noise. On the other hand, the information theoretic line of thought modeled noise and interference appropriately but ignored the randomness in the arrival of messages. Taking note of this, Telatar and Gallager [5] constructed a model that combined certain elements from queueing theory and information theory to capture the essential characteristics of the multiaccess problem. Recently, there has been some work on defining better models for multiaccess communication [6], [7] that combine ideas from information theory and network theory. However, the task of

constructing a model that captures all parameters of interest in multi-access communication still remains a challenge. On the other hand, information theoretic insights and network theoretic principles have now been incorporated widely in wireless (and wireline) networks, albiet in a "layered" fashion. The modularity, inherent in the layered approach, results in simplicity of network design. However, it is being realized now that layered design might have a viable alternative *viz.,* "cross-layer" design. There has been a line of work that suggests that perhaps it is worthwhile to consider the intrinsic "coupling" that exists between different layers [6], [8]–[11] and to follow the cross-layer approach for network design (see also the special issue of IEEE Signal Processing Magazine, September 2004). Cross layer design means a joint design of two or more layers to optimize system-wide performance via an exchange of parameters across layers. The crux of the cross layer approach is to boost network performance by relaxing the layering constraints. It is therefore of interest to characterize when and where cross layer strategies should be used. At the same time, it is also of interest to see when the gain due to cross layer strategies is not enough to warrant a departure from traditional layered approaches.

A. Contributions and Context

We propose a joint information theoretic and queueing theoretic system model that captures the random arrival of messages, multiuser interference and other physical layer parameters. In this model, cross layer design manifests itself in the connection that exists between the information theoretic rate R and the capability of multipacket reception [12]. This connection is made explicit through the notion of outage *viz.,* when the rate R exceeds the instantaneous channel capacity. Next, in this symmetric N user model, we define the notion of stability and characterize the maximum stable throughput of the system under scheduling and ALOHA schemes. These schemes include layered as well as cross layer schemes. We compare these schemes and show that in low SNR regimes, a layered scheme based on information theoretic ideas is optimal whereas in high SNR regimes, a layered scheme based on network theoretic principles is optimal. In the moderate SNR regime, cross layer strategies are shown to outperform layered strategies, although not by significant margins. Further, exact characterizations of the "low" and "high" SNR regimes are given quantitatively. It is also shown that ALOHA with transmission probability one is optimal in the low SNR regime. Thus, we provide conditions under which the proposed cross-layer approach is beneficial and when it is not.

In this work, we restrict ourselves to a symmetric system model. The main reason for this is that stability is extremely

V. Naware is with Qualcomm Inc., 675 Campbell Technology Parkway, Campbell, CA 95008 USA (e-mail: vnaware@qualcomm.com).

L. Tong is with Cornell University, 384 Rhodes Hall, Ithaca, NY 14850 USA (e-mail: ltong@ece.cornell.edu).

difficult to analyze in asymmetric N user systems. For example, with ALOHA, closed form expressions for the stability region are known only for the $N = 2$ case [10]. In addition, we restrict ourselves to single user decoding schemes only. Although, this is a suboptimal approach, it is still in wide use (for example in CDMA cellular systems) because of its inherent design simplicity. Further, in our setup, it provides analytical tractability. Our work can be extended to systems with multiuser detection schemes (such as rate splitting followed by successive decoding), but that is not the focus of this work and will be addressed in a future correspondence.

B. Related Work

Perhaps the most relevant piece of work to ours is that of Telatar and Gallager [5]. In [5], the authors consider an "infinite-user" multi-access communication model over a bandlimited AWGN channel. Assuming single user decoding and poisson arrival of messages, they draw connections of their physical layer model with the processor-sharing model and show the trade-offs involved between queueing theoretic quantities and information theoretic quantities such as the trade-off between delay and probability of error. We focus on the finite user case and consider trade-offs between layered and cross-layer schemes. In [6], the authors consider an ergodic fading channel model consisting of a centralized scheduler that allocates rates and transmit powers (subject to average and peak power constraints) to users subject to the constraint that the rate vector lies in the (ergodic) MAC capacity region. It is also assumed that all the transmitters and the receiver have complete channel and queue state information. The authors find that under any power control scheme, the optimal rate allocation is a "longest weighted queue highest possible rate" scheme and under symmetric conditions it is throughput and delay optimal. In contrast, in our work we consider a non-ergodic Rayleigh fading channel model with only receiver side information of the channel state. In [7], Medard *et. al.* propose using ideas of rate splitting and broadcast codes to construct coding schemes that allow the receiver to correctly decode a subset of the transmitted bits, depending on the number of interfering users in an AWGN setup. Using ALOHA at the MAC layer, they show that as long as the average arrival rates of the users are within the MAC capacity region, there exists a family of codes such that the system is stable. The focus in [7] is on coding at the physical layer whereas our primary focus is a joint optimization of the physical and MAC layers.

Most of the work on cross-layer design is focused on the interaction between the physical and MAC layers. In [10], the authors consider the effect of multipacket reception on stability of a finite user ALOHA system. They show that as the multipacket reception capability improves, ALOHA with transmission probability one becomes optimal in terms of achieving the maximum stable throughput. In [9], Adireddy and Tong consider the effect of having knowledge of fading at the transmitters on the design of ALOHA. They show that significant gains can be made by allowing the transmission probability to be a function of the channel state (as opposed to conventional power control). In [13], the authors propose and analyze a class of protocols called Network Diversity Multiple Access (NDMA) to resolve collisions through packet retrans-

missions using signal processing techniques. They find necessary and sufficient conditions for stability of their protocols using the Foster-Lyapunov criterion for stability (ergodicity) of Markov chains. Our work is based on a cross-layer approach, that in contrast to all of the above, incorporates information theoretic and queueing theoretic quantities of interest in multi-access communication.

The remainder of this paper is organized as follows. In Section II, we define the system model and the notions of stability and maximum stable throughput. In Section III, we describe traditional layered approaches to optimize the maximum stable throughput with two examples *viz.,* (i) single user coding with ALOHA and (ii) no MAC with single user decoding. In Section IV, we describe novel cross PHY/MAC layer approaches to jointly optimize the maximum stable throughput. These approaches include (i) joint coding with scheduling and (ii) joint coding with ALOHA. In Section V, we first provide numerical comparisons between the layered and cross layer schemes and then provide some quantitative results about the optimality of the cross layer schemes. Finally, we conclude in Section VI.

II. SYSTEM MODEL AND MAXIMUM STABLE THROUGHPUT

We define the system model at the physical and MAC layer below.

A. Physical Layer

We consider an N user symmetric slotted multiple access system modelled by a discrete Gaussian MAC channel. A slot consists of n channel uses and is indexed by $t \in \mathbb{N}$. Each user has an average power constraint P. The noise variance at the receiver is σ^2 and hence the transmit SNR $= P/\sigma^2$. We assume a Rayleigh block fading channel for every user in every slot. The fading gain for user i in time slot t is denoted by $h_i(t)$ and is assumed to have unit variance. The fading is also assumed to be independant across users and slots. We assume that only the receiver has perfect channel state information. Each user encodes information at a fixed rate of R bits per channel use[1]. The receiver employs single user decoding to decode information from users. By that we mean that the receiver decodes a user's codeword assuming every other user's signal as noise. We also assume that the slot size n is large enough to achieve the instantaneous capacity of the channel provided that there is no outage.

B. MAC layer

We assume that every user has an infinite queue to store arriving bits. Let $A_i(t)$ be the number of bits[2] arriving to user i in time slot t with $\mathbb{E}(A_i(t)) = n\lambda/N$ so that the sum arrival rate in bits per channel use is λ. We assume that the arrivals are independent across users and time slots. Let $Q_i(t)$ denote the number of bits in the ith user's queue in time slot t. Let $D_i(t)$ be the number of bits departing (or successfully decoded) from the ith user's queue in time

---

[1]We assume that nR is a positive integer.
[2]We assume that the number of arriving bits is a non-negative integer.

slot t. Thus, $D_i(t)$ is a random variable that is a function of the channel realizations and the underlying MAC protocol. When a successful departure occurs, the following happens: if $Q_i(t) \geq nR$ then $D_i(t) = nR$ and if $Q_i(t) < nR$ then $D_i(t) = Q_i(t)$. Note that as a result, $D_i(t) \in \{0, 1, \cdots, nR\}$. Assuming that there are no feedback errors, it follows that the queue for user i, $1 \leq i \leq N$ evolves as

$$Q_i(t+1) = \big(Q_i(t) - D_i(t)\big)^+ + A_i(t), \qquad (1)$$

where $(x)^+$ denotes $\max\{x, 0\}$. Define a new random variable $\tilde{D}_i(t)$ as

$$\tilde{D}_i(t) = \begin{cases} nR, & \text{if } 0 < D_i(t) \leq nR, \\ 0, & \text{if } D_i(t) = 0. \end{cases}$$

Note that if we replace $D_i(t)$ with $\tilde{D}_i(t)$ in (1), then the queue evolution is left unaltered. So, for the remainder of this paper we consider $\tilde{D}_i(t)$ as the departure process. Note also that we are implicitly assuming that if a user has less than nR bits in his queue at time t, then still he can encode his information (by say stuffing the remaining bits) so as to be decoded by the receiver provided that there is no outage. It is easy to see that the joint queue lengths $\mathbf{Q}(t) \triangleq (Q_1(t), Q_2(t), \cdots, Q_N(t))$ has a countable state space $\mathbb{N}_+^N$, where $\mathbb{N}_+$ is the set of non-negative integers.

C. Maximum Stable Throughput

In this subsection, we formalize our notion of stability and maximum stable throughput. Intuitively, a system is defined to be stable if for each node the queue size does not go to infinity. In other words, given a positive number $0 < \epsilon \leq 1$, there exists a buffer size such that the probability of buffer overflow is less than ϵ. It should be obvious that stability is an important requirement for a network. The requirement of stability can be said to impose a mild requirement on delay.

We now define the notions of stability and maximum stable throughput. We say that the system $\mathbf{Q}(t)$ is stable for a particular arrival process, if for $\mathbf{x} \in \mathbb{N}_+^N$, there exists a $H(\mathbf{x})$ such that

$$\lim_{t \to \infty} \Pr\{\mathbf{Q}(t) < \mathbf{x}\} = H(\mathbf{x}), \quad \lim_{\mathbf{x} \to \infty} H(\mathbf{x}) = 1.$$

This notion of stability has been used before [10], [14]–[16]. Note that this definition of stability does not imply that the expected delay in the system is finite (but the converse is true). We shall see that the stability of the system can be characterized by λ, the sum arrival rate alone. This allows us to define maximum stable throughput as the supremum of all sum rates λ for which the system is stable. Note that the maximum stable throughput is a function of the underlying MAC protocol.

III. LAYERED APPROACHES

The traditional layered network architecture uses a modular approach in the design of the various layers; all the layers are designed separately so as to simplify network design. The design of PHY and MAC layers are based on abstractions such as the collision channel *viz.,* the MAC layer is designed to

avoid "collisions" that result from simultaneous transmissions from more than one user. The PHY layer is assumed to be incapable of handling multi-user interference. Similarly, PHY layer designs assume users always have data to send and thus neglect the issue of source burstiness. As an example, consider the following layered schemes for the system model described in Section II.

A. No MAC With Single User Decoding

In this scheme, given an outage requirement ϵ, every user transmits at the highest possible rate $R_1(\epsilon, \text{SNR})$ to meet the outage requirement assuming that every other user has information to send. An outage happens when the rate chosen to encode information exceeds the instantaneous capacity of the channel in the slot. Assuming single user decoding, it is easy to see that $R_1(\epsilon, \text{SNR})$ satisfies

$$\Pr\left\{ R_1(\epsilon, \text{SNR}) \leq \frac{1}{2} \log\left(1 + \frac{|h_1|^2 \text{SNR}}{\sum_{i=2}^{N} |h_i|^2 \text{SNR} + 1}\right) \right\} = 1 - \epsilon.$$

Note that the encoding rate is so chosen as to satisfy the quality of service requirement when all users have bits and thus this is a worst case design. For such a physical layer design, a MAC protocol is not needed and indeed this is an information theoretic approach to multiple access. The idea is that multiuser interference is tackled through coding and there is no need for transmission control. It is intuitively clear that for a given ϵ and SNR, the maximum stable throughput is given by $(1 - \epsilon)NR_1(\epsilon, \text{SNR})$. We denote the maximum stable throughput of this scheme by $\rho_1(\text{SNR})$.

Proposition 1:

$$\rho_1(\text{SNR}) = \max_{\epsilon \in [0,1]} (1 - \epsilon)NR_1(\epsilon, \text{SNR}).$$

Proof: Refer to the appendix.

B. Single User Coding With ALOHA

In this scheme, given an outage requirement ϵ, every user transmits at the highest possible rate $R_2(\epsilon, \text{SNR})$ to meet the outage requirement assuming that no other user has bits to send. In this case, it is easy to see that $R_2(\epsilon, \text{SNR})$ satisfies

$$\Pr\left\{ R_2(\epsilon, \text{SNR}) \leq \frac{1}{2} \log(1 + |h_1|^2 \text{SNR}) \right\} = 1 - \epsilon.$$

In this scheme, the encoding rate is chosen so as to meet the outage requirement assuming the best case scenario *viz.,* no other user has information bits to transmit. Consequently, the need for transmission control arises and for this scheme we assume a symmetric ALOHA protocol with transmission probability $p = 1/N$. For this scheme, we denotee the maximum stable throughput to be $\rho_2(\text{SNR})$.

Proposition 2:

$$\rho_2(\text{SNR}) = \max_{\epsilon \in [0,1]} (1 - \epsilon)R_2(\epsilon, \text{SNR})\left(1 - \frac{1}{N}\right)^{N-1}.$$

Proof: Follows by arguments similar to the one used in the proof of Proposition 1.

Note that with TDMA, the maximum stable throughput is simply given by $\max_{\epsilon \in [0,1]} (1 - \epsilon)R_2(\epsilon, \text{SNR})$.

IV. CROSS-LAYER APPROACHES

In this section, we investigate cross-layer design as an alternative to the layered strategies discussed in the previous section. By "cross-layer" design we mean that parameters used in the design of PHY and MAC layers are exchanged to optimize performance jointly. In order to see how this is done, first observe that when k users transmit in a slot, nR bits from any particular user (say user 1) get through with probability (denoted by p_n^k)

$$p_n^k = \Pr\left\{ R \leq \frac{1}{2} \log\left(1 + \frac{|h_1|^2 \text{SNR}}{\sum_{i=2}^k |h_i|^2 \text{SNR} + 1} \right) \right\}.$$

Proposition 3: When the fading coefficients are i.i.d. Rayleigh with unit variance (so that $|h_i|^2$ has a unit mean exponential distribution), then

$$p_n^k = \exp\left(-\frac{(e^{2R}-1)}{\text{SNR}} \right) \frac{1}{e^{2R(k-1)}}. \tag{2}$$

Proof: Refer to the appendix.

By symmetry, the average number of bits that get through per slot when k users transmit is

$$k(nR) \exp\left(-\frac{(e^{2R}-1)}{\text{SNR}} \right) \frac{1}{e^{2R(k-1)}},$$

and hence, the average number of bits that get through per channel use when k users transmit is

$$C_k(R, \text{SNR}) \triangleq kR \exp\left(-\frac{(e^{2R}-1)}{\text{SNR}} \right) \frac{1}{e^{2R(k-1)}}. \tag{3}$$

Here, we note the difference of our formulation with that of Ghez, Verdu and Schwartz [12]. In [12], in a corresponding "packet" model, C_k is defined to be the average number of "packets" that get through per slot when k users transmit. In contrast, equation 3 captures the essence of the cross-layer approach: the dependence of C_k on physical layer parameters such as R, SNR is seen explicitly and since C_k determines MAC layer performance, throughput maximization involves both physical and MAC layer parameters.

Now we turn our attention to two cross-layer schemes.

A. Joint Coding and Scheduling

In this scheme, in every slot the receiver schedules a subset of users to transmit in order to maximize the stable throughput. By definition, scheduling is the best MAC protocol in the sense that for a given encoding rate R and SNR, it achieves the maximum stable throughput. Since we are considering a symmetric system, it follows that the optimal scheduling scheme will be of the following form: in every slot, a certain number of users (k_{opt}) will be scheduled to transmit in a cyclic manner to guarantee fairness. In order to choose k_{opt}, the receiver uses the physical layer parameters (i) encoding rate R and (ii) the transmit SNR. For a given SNR and encoding rate R, let $\rho^*(R, \text{SNR}, N)$ be the maximum stable throughput with scheduling and let $k_{\text{opt}}(R, \text{SNR}, N)$ be the optimal size of the subset of users scheduled.

Lemma 1: For a fixed SNR and encoding rate R, the maximum stable throughput with scheduling is given by

$$\rho^*(R, \text{SNR}, N) = \max\{C_1(R, \text{SNR}), C_2(R, \text{SNR}), \\ \cdots, C_N(R, \text{SNR})\}.$$

Proof: Refer to the appendix.

Intuitively, it is clear that if all users had infinite bits in their queue then indeed $\rho^*(R, \text{SNR}, N)$ would be the maximum throughput. However, the result implies that even with random arrivals of bits, $\rho^*(R, \text{SNR}, N)$ is the maximum stable arrival rate. This result is a recurring theme in symmetric finite user multiple access systems—the maximum stable throughput with i.i.d random arrivals is the throughput as if the system is infinitely backlogged [9], [10], [14]. The next theorem provides us with an analytical characterization of $\rho^*(R, \text{SNR}, N)$.

Theorem 1:

$$\rho^*(R, \text{SNR}, N) = C_N(R, \text{SNR})\mathbf{1}_{R \in [0, \, 0.5 \log(N/N-1)]} \\ + C_1(R, \text{SNR})\mathbf{1}_{R \in (0.5 \log 2, \, \infty)} \\ + \sum_{i=2}^{N-1} C_i(R, \text{SNR})\mathbf{1}_{R \in (0.5 \log(i+1/i), \, 0.5 \log(i/i-1)]}.$$

where $\mathbf{1}_{R \in [a,b]}$ is the indicator function of R over the interval $[a, b]$.

Proof: Refer to the appendix.

From Theorem 1, we can infer that the optimal size of the set of scheduled users is a monotonically decreasing step function of the rate R. Thus, for small rates ($R \leq 0.5 \log(N/N - 1)$), the optimal set size is N and throughput is $C_N(R, \text{SNR})$. As the rate increases, the optimal set size decreases monotonically and once $R > 0.5 \log 2$, the optimal set size is 1.

Next, we optimize the PHY layer parameter R *i.e.,* we choose the best encoding rate R to optimize the maximum stable throughput. Let $\rho(\text{SNR}, N)$ be the maximum stable throughput with scheduling for a given SNR,

$$\rho_{SCH}(\text{SNR}, N) \triangleq \sup_R \rho^*(R, \text{SNR}, N).$$

Note that joint coding with scheduling is optimal amongst all MAC protocols for a fixed SNR and thus $\rho(\text{SNR}, N)$ is an upper bound to the maximum stable throughput of any MAC protocol. Note that although it is not easy to find closed form expressions for $\rho(\text{SNR}, N)$, it is easy to do so numerically.

B. Joint Coding with ALOHA

We consider a symmetric ALOHA MAC protocol where every user independently decides to transmit in a slot with probability p provided he has bits in his queue. In this scheme, first for a given SNR and encoding rate R, we choose the optimal transmission probability p to maximize the stable throughput of ALOHA. For a given SNR and R, let $\rho_{\text{ALOHA}}(R, \text{SNR}, N)$ be the maximum stable throughput of ALOHA and $p^*(R, \text{SNR}, N)$ be the optimal transmission probability.

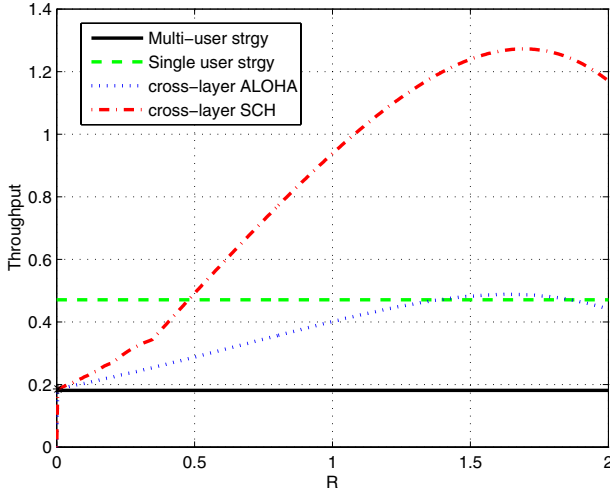*Theorem 2:* For a fixed SNR and encoding rate R, the maximum stable throughput of ALOHA is given by

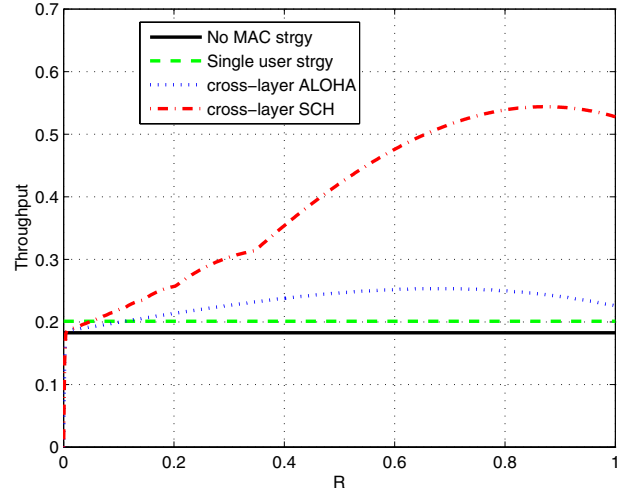Fig. 1.   Comparison of layered and cross-layer approaches with $N = 100$, SNR $= 20dB$.

Fig. 2.   Comparison of layered and cross-layer approaches with $N = 100$, SNR $= 10dB$.

$$\rho^*_{\text{ALOHA}} =$$
$$\begin{cases} \frac{NR}{e^{2R(N-1)}} \exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right) & \text{if } R \le \frac{1}{2}\log\left(\frac{N}{N-1}\right) \\ \frac{R}{1-e^{-2R}} \exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right)\left(1-\frac{1}{N}\right)^{N-1} & \text{otherwise.} \end{cases} \quad (4)$$

and the optimal transmission probability is given by

$$p^*(R, \text{SNR}, N) = \begin{cases} \frac{1}{N(1-e^{-2R})} & \text{if } R > \frac{1}{2}\log\left(\frac{N}{N-1}\right) \\ 1 & \text{otherwise.} \end{cases}$$

Proof: Refer to the appendix.

Note that as long as the encoding rate R is less than $\frac{1}{2}\log\left(\frac{N}{N-1}\right)$, the optimal transmission probability is 1 and the corresponding maximum stable throughput is actually $C_N(R, \text{SNR})$. From Theorem 1 and Theorem 2, it follows that for a given R and SNR, ALOHA with $p = 1$ is optimal if $R \le \frac{1}{2}\log\left(\frac{N}{N-1}\right)$. It is also interesting to note that in this regime, the optimal transmission probability is not a function of SNR.

Next, we optimize R to get the best possible stable throughput with ALOHA for a fixed SNR. Let $\rho_{\text{ALOHA}}(\text{SNR}, N)$ be the maximum stable throughput with ALOHA for a fixed SNR. Thus,

$$\rho_{\text{ALOHA}}(\text{SNR}, N) \triangleq \sup_R \rho^*_{\text{ALOHA}}(R, \text{SNR}, N).$$

Unfortunately, both $\rho_{\text{ALOHA}}(\text{SNR}, N)$ and $\rho(\text{SNR}, N)$ cannot be computed in closed form and so we have to resort to numerical evaluations of these functions.

V. COMPARISON OF LAYERED AND CROSS-LAYER APPROACHES

In this section, we compare the maximum stable throughput provided by all the schemes described in the previous sections for a fixed SNR.

Fig. 1 compares the four different schemes in terms of their maximum stable throughput at high SNR (20 dB). We see that

optimal cross-layer scheduling provides the maximum stable throughput and all other schemes are sub-optimal. Since the optimal R achieving maximum stable throughput with cross-layer scheduling is greater than $0.5 \log 2$, it follows that exactly one user is scheduled to transmit in each slot. However, it is interesting to note that the maximum stable throughput of single user coding with ALOHA is almost the same as that of cross-layer ALOHA. Further, the ratio of maximum stable throughput of these schemes to optimal cross-layer scheduling is approximately e^{-1}. This suggests that (i) the loss in throughput is only due to non-coordination among users' transmissions and (ii) the layered single user approach is close to optimal since the maximum stable throughput with TDMA would be close to that of cross-layer scheduling. However, the no MAC with single user decoding approach is clearly sub-optimal.

Fig. 2 compares the four different schemes at medium SNR (10 dB). Here, we see that both layered approaches are sub-optimal. The cross-layer ALOHA scheme offers higher throughput than the corresponding single user coding with ALOHA scheme.

Fig. 3 shows the stable throughput in the low SNR regime ($-20\ dB$). Note that the range of R shown in Fig. 3 is much smaller than that in Figs. 1 and 2. In the low SNR regime, the single user coding with ALOHA strategy is sub-optimal. On the other hand, the stable throughput of both cross-layer schemes is almost the same. For $R \le 0.5 \log(100/99)$, they are identical and thereafter the difference is marginal. The interesting observation is that the no MAC with single user decoding strategy is optimal. Since the optimal encoding rate for cross-layer scheduling/ALOHA is less than $0.5 \log(100/99)$, it follows that both schemes correspond to allowing all users to transmit whenever they have bits. This is exactly the information theoretic approach to multiple access.

To summarize, we see that in both the high and low SNR regimes, one of the two layered approaches is optimal: In the high SNR regime, single user coding with TDMA is optimal whereas in the low SNR regime, no MAC with single
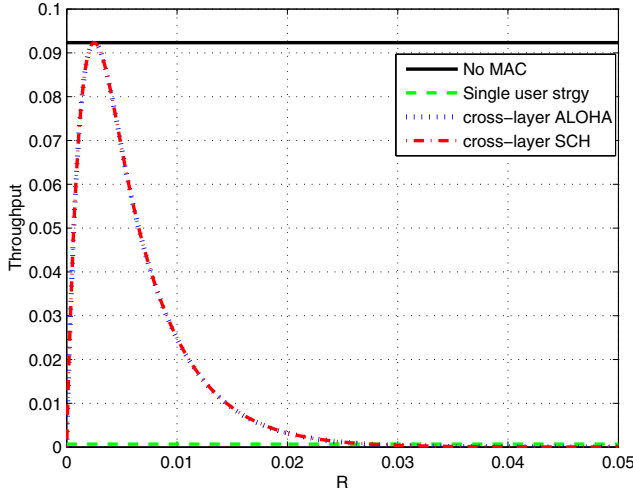
Fig. 3.   Comparison of layered and cross-layer approaches with $N = 100$, SNR $= -20dB$.

Fig. 4.   Comparision of optimal encoding rates with cross-layer schemes. $N = 10$

user decoding is optimal. Intuitively, this can be explained as follows. In the high SNR regime, the system is interference limited due to sub-optimal single user decoding and hence medium access control is necessary. So, the optimal strategy is to encode at a high rate and schedule every user to transmit once in every N slots. On the other hand in the low SNR regime, the dominant interference term is noise and not multiuser interference. Hence, the optimal scheme corresponds to no transmission control with low rate encoding.

Next, we show that there is actually a sharp transition to the "low" SNR regime asymptotically as $N \to \infty$ and that this transition occurs between -1.76 to $1.98\ dB$ with scheduling. Precisely,

Theorem 3: If SNR $< \frac{2}{3}$, then $\forall N$,

$$\arg \max_R \left(\rho^*(R, \text{SNR}, N) \right) \leq 0.5 \log \left(\frac{N}{N-1} \right).$$

If SNR $> \frac{1}{1+\log(\log(2))}$, then $\exists\ N_1^*$ such that $\forall N > N_1^*$,

$$\arg \max_R \left(\rho^*(R, \text{SNR}, N) \right) > 0.5 \log \left(\frac{N}{N-1} \right).$$

Proof: Refer to [17].

Theorem 3 says that as long as the transmit SNR is less than $2/3$, scheduling all users to transmit simultaneously is optimal or in other words, ALOHA with transmission probability one is optimal. On the other hand, if the transmit SNR is greater than approximately 1.54, for a large enough network scheduling all users to transmit is suboptimal. Note that although the optimal rate cannot be expressed in closed form, it can be computed numerically. From numerical calculations we conjecture that this phase transition actually occurs at $0\ dB$ asymptotically in N. This is quite surprising since for any fixed SNR, we expect medium access control to be necessary as the number of users increases. The reason for this behavior when SNR $< \frac{2}{3}$ is that the optimal rate goes to zero as $N \to \infty$ as well.
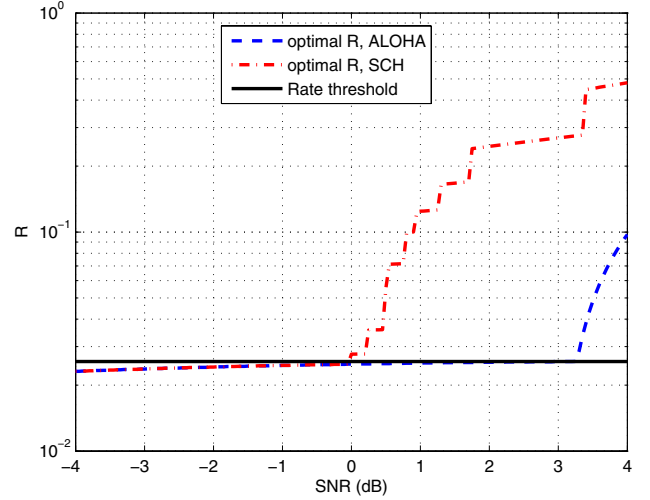
Next, we restrict ourselves to using ALOHA as the MAC protocol. Here, again we find a sharp phase transition in the point at which ALOHA with transmission probability $p = 1$ ceases to be optimal asymptotically as $N \to \infty$.

Theorem 4: If SNR < 2, then $\forall N$,

$$\arg \max_R \left(\rho_{\text{ALOHA}}^*(R, \text{SNR}, N) \right) \leq 0.5 \log \left(\frac{N}{N-1} \right).$$

If SNR > 2, then $\exists\ N_2^*$ such that $\forall N > N_2^*$,

$$\arg \max_R \left(\rho_{\text{ALOHA}}^*(R, \text{SNR}, N) \right) > 0.5 \log \left(\frac{N}{N-1} \right).$$

Proof: Refer to [17].

Theorem 4 says that as long as the transmit SNR is less than $3\ dB$, no transmission control is optimal ($p^* = 1$) and above $3\ dB$, for a large enough network, transmission control is necessary ($p^* < 1$). This result, similar to Theorem 4, is also rather counter-intuitive. It is still not clear to us why the thresholds for the "low" SNR regime with both scheduling and ALOHA are in the 0 to 5 dB range.

Figs. 4, 5, 6, 7 show the numerically computed optimal encoding rates vs. SNR for (i) cross-layer scheduling and (ii) cross-layer ALOHA with $N = 10, 20, 30, 60$. The black horizontal line marks $0.5 \log(N/N-1)$, the encoding rate below which transmission control is not necessary. These figures illustrate Theorem 3 and Theorem 4 quantitatively. We see that as N increases, with scheduling, the optimal encoding rate crosses the black line around $0\ dB$ whereas with ALOHA, the crossover happens around $3\ dB$. We also notice that the scheduling curve has several kinks. Each kink corresponds to a change in the optimal scheduling set size. Since, we do not have closed form expressions for the optimal encoding rates, it is difficult to infer if the kinks are actually discontinuities in the curve. To summarize, we have shown that with both scheduling and ALOHA, as long as the SNR is below a *fixed* threshold, there is no need for transmission control *regardless of the number of users in the system*. Conversely, if the SNR
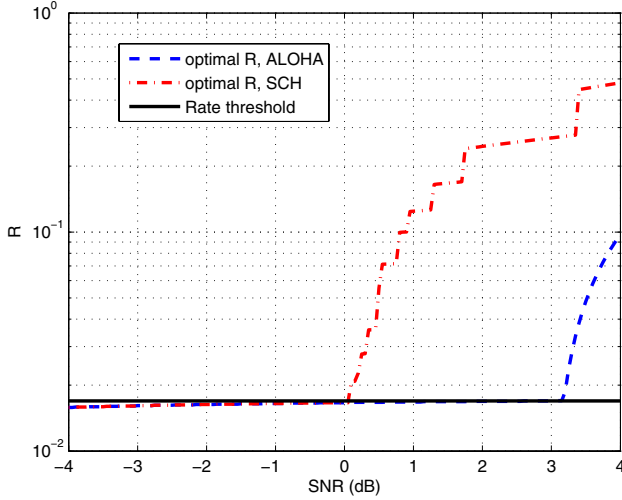
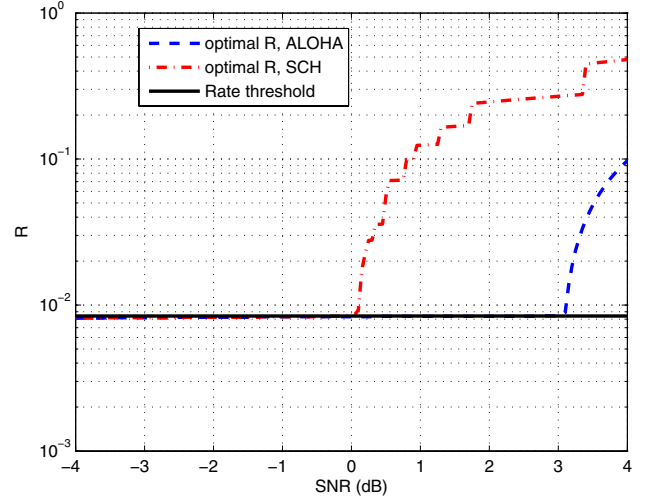Fig. 5. Comparison of optimal encoding rates with cross-layer schemes. $N = 20$

Fig. 7. Comparision of optimal encoding rates with cross-layer schemes. $N = 60$
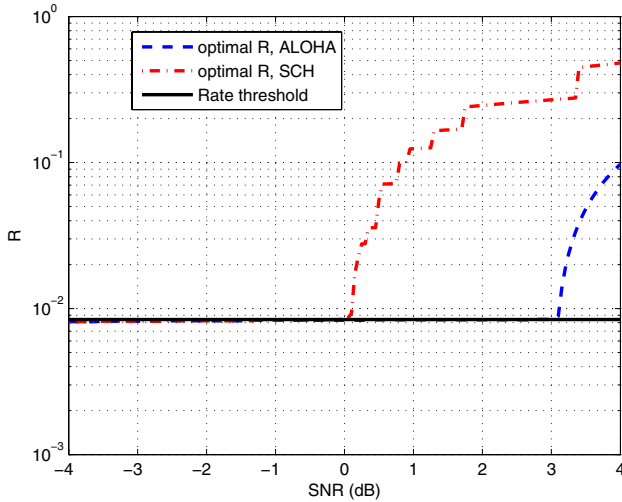

Fig. 6. Comparision of optimal encoding rates with cross-layer schemes. $N = 30$

exceeds a *fixed* threshold and if the number of users in the system is large enough, *transmission control is necessary.*

Next, we characterize the maximum stable throughput in the low SNR regime.

Theorem 5: For SNR $< \frac{2}{3}$,

$$\lim_{N \to \infty} \rho(\text{SNR}, N) = \frac{1}{2e},$$

and for SNR < 2,

$$\lim_{N \to \infty} \rho_{\text{ALOHA}}(\text{SNR}, N) = \frac{1}{2e}.$$

Proof: Refer to [17].

From Theorems 3 and 4, note that in the low SNR regime, the optimal encoding rate goes to zero with both ALOHA and scheduling. However, Theorem 5 implies that as $N \to \infty$, in the low SNR regime, even though the maximum stable throughput per user with both scheduling and ALOHA goes to zero, the total maximum stable throughput converges to

$(2e)^{-1}$. In general, it is difficult to give closed form characterizations of the asymptotic performance of these cross-layer schemes in the "high" SNR regime.

VI. CONCLUSIONS

In this work, we proposed a model for multi-access communication that combines ideas from information theory and queueing theory. In this symmetric N user model, we defined the notion of stability and characterized the maximum stable throughput of the system under scheduling and ALOHA MAC schemes. These schemes included layered as well as cross layer schemes. We compared these schemes and showed that in low SNR regimes, a layered scheme based on information theoretic ideas is optimal whereas in high SNR regimes, a layered scheme based on network theoretic principles is optimal. Thus, one of our main finding is that it is not always necessary to take a cross-layer approach—layered strategies can indeed be optimal in certain regimes. However, in the moderate SNR regime, we showed that cross layer strategies outperform layered strategies. We also showed that ALOHA with transmission probability one is optimal in the low SNR regime. Future work would be directed along two lines: *(i)* generalizing results to other common fading distributions and *(ii)* incorporating optimal multiuser detection strategies based on rate splitting and successive decoding in the model.

APPENDIX

A. Proof of Proposition 1

Note that for a fixed ϵ and SNR, the maximum stable throughput of the no MAC with single user decoding (denote by $\rho_1(\epsilon, \text{SNR})$) is given by equation (14) with $p = 1$ and $R = R_1(\epsilon, \text{SNR})$. Thus,

$$\rho_1(\epsilon, \text{SNR}) = C_N(R_1(\epsilon, \text{SNR}), \text{SNR}),$$
$$= (1 - \epsilon)N R_1(\epsilon, \text{SNR}), \qquad (5)$$

where, (5) follows from equation (3) and the definition of no MAC with single user decoding scheme. Finally, the maximum stable throughput of this scheme is achieved by taking the supremum over the outage probability ϵ. $\square$

B. Proof of Proposition 3

We proceed step by step. Denote the left hand side of equation (2) by LHS. Then,

$$
\begin{aligned}
&\text{LHS}\\
&=\Pr\left\{|h_1|^2 \geq (\sum_{i=2}^{k}|h_i|^2)(e^{2R}-1)+\frac{(e^{2R}-1)}{\text{SNR}}\right\},\\
&=\int_0^{\infty}\Pr\left\{|h_1|^2 \geq t(e^{2R}-1)+\frac{(e^{2R}-1)}{\text{SNR}}\right\}\frac{t^{k-2}e^{-t}}{(k-2)!}\,dt\text{(6)}\\
&=\int_0^{\infty}\exp\left(-t(e^{2R}-1)-\frac{(e^{2R}-1)}{\text{SNR}}\right)\frac{t^{k-2}e^{-t}}{(k-2)!}\,dt,\quad\text{(7)}\\
&=\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right)\int_0^{\infty}\frac{\exp(-te^{2R})t^{k-2}}{(k-2)!}\,dt,\\
&=\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right)\frac{1}{e^{2R(k-1)}},\quad\quad\text{(8)}
\end{aligned}
$$

where, we get equations (6), (7) since $\sum_{i=2}^{k}|h_i|^2$ has a gamma distribution (sum of unit mean i.i.d. exponential distributions). □

C. Proof of Lemma 1

We sketch an outline of the proof. By Corollary 1 in [10], it follows that the maximum stable throughput of any MAC protocol is less than $\max\{C_1(R,\text{SNR}),\cdots,C_N(R,\text{SNR})\}$. Thus, it follows that

$$\rho^*(R,\text{SNR},N) \leq \max\{C_1(R,\text{SNR}),\cdots,C_N(R,\text{SNR})\}.$$

On the other hand, note that $\max\{C_1(R,\text{SNR}),\cdots,C_N(R,\text{SNR})\}$ is the maximum stable throughput with scheduling *assuming all users have infinite bits to begin with*. Note also that

$$C_1(R,\text{SNR}) \geq \frac{C_2(R,\text{SNR})}{2} \geq \cdots \geq \frac{C_N(R,\text{SNR})}{N},\quad\text{(9)}$$

and hence the probability of success for a user increases as the number of transmitting users decreases. This implies that the assumption that all users have infinite bits to begin with gives the worst case stability analysis and hence

$$\rho^*(R,\text{SNR},N) \geq \max\{C_1(R,\text{SNR}),\cdots,C_N(R,\text{SNR})\}.\quad\text{(10)}$$

For technical details regarding equation (10) which use the concept of stochastic dominance, refer to [18]. □

D. Proof of Theorem 1

Define a function $g_x(R)$ for $x \in \mathbb{R}$ as

$$g_x(R) = xe^{-2R(x-1)}.\quad\quad\text{(11)}$$

Note that this function attains its supremum at $x = 1/2R$. Also note that $g_x(R)$ is strictly increasing for $x < 1/2R$ and strictly decreasing for $x > 1/2R$. Now, note that

$$C_k(R,\text{SNR}) = g_k(R)\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right),\quad\text{(12)}$$

and hence (13) (see top of next page). Next, since $g_x(R)$ is strictly increasing for $x < 1/2R$ and strictly decreasing for $x > 1/2R$, it follows that for any R, $\max_{1\leq k\leq N}\{g_k(R)\}$ occurs at $k = \lfloor 1/2R \rfloor$ or $k = \lfloor 1/2R \rfloor + 1$ where $\lfloor r \rfloor$ denotes the smallest integer less than or equal to r. Thus, it follows that it suffices to compare $g_{\lfloor 1/2R \rfloor}(R)$ and $g_{\lfloor 1/2R \rfloor+1}(R)$ for a given R. Now, $g_k(R) \geq g_{k+1}(R)$ if and only if $R \geq 0.5\log(k+1/k)$. Also note that at $R = 1/2$, the optimal k equals 1. But, $g_1(R) \geq g_2(R)$ if $R \geq 0.5\log 2$ and so the optimal k equals 1 for $R \geq 0.5\log(2)$. After that point $g_2(R)$ is the maximum and continuing this argument we see that the optimal k equals n for $0.5\log(n+1/n) \leq R \leq 0.5\log(n/n-1)$. Thus, it follows that

$$
\begin{aligned}
\rho^*(R,\text{SNR},N) &= C_N(R,\text{SNR})\mathbf{1}_{R\in[0,\,0.5\log(N/N-1)]}\\
&+C_1(R,\text{SNR})\mathbf{1}_{R\in(0.5\log 2,\,\infty)}\\
&+\sum_{i=2}^{N-1}C_i(R,\text{SNR})\mathbf{1}_{R\in(0.5\log(i+1/i),\,0.5\log(i/i-1)]},
\end{aligned}
$$

and we are done. □

E. Proof of Theorem 2

Note that with ALOHA, the joint queue lengths $\mathbf{Q}(t)$ is a countable state space Markov chain. Further, under loose conditions (for eg. $\Pr\{A_i(t) = 1\} > 0$), it is irreducible and aperiodic. Then, from [19], it can be shown as a special case that for a finite user symmetric system with transmission probability p satisfying the condition (9), the maximum stable throughput of ALOHA is given by

$$\rho_{\text{ALOHA}}(p,R,\text{SNR},N) = \sum_{k=1}^{N}\binom{N}{k}p^k(1-p)^{N-k}C_k(R,\text{SNR}).\quad\text{(14)}$$

Then, by definition,

$$\rho_{\text{ALOHA}}^*(R,\text{SNR},N) = \max_{p\in[0,1]}\rho_{\text{ALOHA}}(p,R,\text{SNR},N).\quad\text{(15)}$$

Substituting $C_k(R,\text{SNR})$ using equation (3) in equation (14), we get

$$
\begin{aligned}
\rho_{\text{ALOHA}} &= \sum_{k=1}^{N}\binom{N}{k}p^k(1-p)^{N-k}kR\frac{1}{e^{2R(k-1)}}\\
&\quad\times\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right),\\
&= \left(\sum_{k=1}^{N}\binom{N}{k}k\left(\frac{p}{e^{2R}}\right)^k(1-p)^{N-k}\right)\\
&\quad\times Re^{2R}\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right),\\
&= \left(\sum_{k=1}^{N}\frac{(N-1)!}{(k-1)!(N-k)!}\left(\frac{p}{e^{2R}}\right)^k(1-p)^{N-k}\right)\\
&\quad\times NRe^{2R}\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right),\\
&= \left(\sum_{k=1}^{N}\frac{(N-1)!}{(k-1)!(N-k)!}\left(\frac{p}{e^{2R}}\right)^{k-1}(1-p)^{N-k}\right)
\end{aligned}
$$

$$\max\left\{C_1(R,\text{SNR}),\cdots,C_N(R,\text{SNR})\right\} = \max\left\{g_1(R),\cdots,g_N(R)\right\}\cdot\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right). \tag{13}$$

$$\times NR\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right)p,$$
$$= NR\exp\left(-\frac{(e^{2R}-1)}{\text{SNR}}\right)$$
$$\times p(1-p+pe^{-2R})^{N-1}. \tag{16}$$

From equation (16), it can be easily shown that the optimal p for maximizing the stable throughput of ALOHA is

$$p^*(R,\text{SNR},N) = \begin{cases} \frac{1}{N(1-e^{-2R})} & \text{if } R > \frac{1}{2}\log(\frac{N}{N-1}) \\ 1 & \text{otherwise.} \end{cases}$$

Substituting $p^*(R,\text{SNR},N)$ back in equation (16), gives (4) and we are done. □

ACKNOWLEDGEMENT

REFERENCES

[1] R. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Inform. Theory*, vol. 31, pp. 124–142, Mar. 1985.
[2] A. Ephremides and B. Hajek, "Information theory and communication networks: An unconsummated union," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2416–2434, Oct. 1998.
[3] N. Abramson, "The ALOHA system–Another alternative for computer communications," in *Proc. Fall Joint Comput. Conf., AFIPS*, 1970, vol. 44, pp. 281–285.
[4] H. Liao, "Multiple access channels," PhD thesis, University of Hawaii, 1972.
[5] I. E. Telatar and R. G. Gallager, "Combining queueing theory with information theory for multiaccess," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 963–969, Aug. 1995.
[6] R. Berry and E. Yeh, "Cross-layer wireless resource allocation: Fundamental performance limits," *IEEE Signal Processing Mag., Special Issue on Signal Processing for Networking: An integrated Approach*, vol. 21, pp. 59–68, July 2004.
[7] M. Medard, J. Huang, A. Goldsmith, S. Meyn, and T. Coleman, "Capacity of time-slotted ALOHA packetized multiple access systems over the AWGN channel," *IEEE Trans. Wirless Commun.*, vol. 3, pp. 486–499, Mar. 2004.
[8] L. Tong, Q. Zhao, and G. Mergen, "Multipacket reception in random access wireless networks: From signal processing to optimal medium access control," *IEEE Commun. Mag., Special Issue on Design Methodologies for Adaptive and Multimedia Networks*, vol. 39, pp. 108–112, Nov. 2001.
[9] S. Adireddy and L. Tong, "Exploiting decentralized channel state information for random access," *IEEE Trans. Inform. Theory*, vol. 51, pp. 537–561, Feb. 2005. Available: http://acsp.ece.cornell.edu/pubJ.html
[10] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite user slotted ALOHA with multipacket reception," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2636–2656, July 2005.
[11] L. Tong, V. Naware, and P. Venkitasubramaniam, "Signal processing in random access," *IEEE Signal Processing Mag., Special Issue on Signal Processing for Networking: An Integrated Approach*, vol. 21, pp. 29–39, Sep. 2004.
[12] S. Ghez, S. Verdú, and S. Schwartz, "Stability properties of slotted ALOHA with multipacket reception capability," *IEEE Trans. Automat. Cont.*, vol. 33, pp. 640–649, July 1988.
[13] G. Dimic, N. Sidiropoulas, and L. Tassiulas, "Wireless networks with retransmission diversity access mechanisms: Stable throughput and delay properties," *IEEE Trans. Signal Processing, Special Issue on Signal Processing in Networking*, vol. 51, pp. 2019–2030, Aug. 2003.
[14] B. Tsybakov and W. Mikhailov, "Ergodicity of slotted ALOHA systems," *Probl. Inform. Transmission*, vol. 15, no. 4, pp. 73–87, 1979.
[15] W. Szpankowski, "Stability conditions for some multiqueue distributed systems: Buffered random access systems," *Adv. Appl. Probab.*, vol. 26, pp. 498–515, 1994.
[16] W. Luo and A. Ephremides, "Stability of N interacting queues in random-access systems," *IEEE Tran. Inform. Theory*, vol. 45, pp. 1579–1587, 1999.
[17] V. Naware and L. Tong, "Cross layer design for multi-access communication over Rayleigh fading channels," Tech. Rep. ACSP-TR-03-05-01, Cornell University, Mar. 2005. Available: http://acsp.ece.cornell.edu/pubR.html
[18] G. Mergen and L. Tong, "Stability and capacity of wireless networks with probabilistic perceptions: Part I—General topologies, Part II—Regular networks," Tech. Rep. ACSP-TR-01-03-01, Cornell University, Jan. 2003.
[19] S. Adireddy and L. Tong, "Optimal transmission probabilities for slotted ALOHA in fading channels," in *Proc. CISS*, Mar. 2002.

Vidyut Naware (S'03) was born in India in 1978. He received the B.Tech degree in Electrical Engineering from the Indian Institute of Technology, Bombay in 2000. He received his M.S and Ph.D degrees from the School of Electrical and Computer Engineering at Cornell University in 2004 and 2005 respectively. He is currently employed with Qualcomm Inc., Campbell, California. His research interests broadly include wireless communications, information theory, and networking.

Lang Tong (S'87-M'91-SM'01-F'05) is a Professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. He received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1991, respectively, from the University of Notre Dame, Notre Dame, Indiana. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University in 1991. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology.

Dr. Tong received the *Young Investigator Award* from the Office of Naval Research in 1996, the *Outstanding Young Author Award* from the IEEE Circuits and Systems Society in 1991, the 2004 IEEE Signal Processing Society Best Paper Award (with M. Dong), and the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society (with P. Venkitasubramaniam and S. Adireddy). He serves as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE SIGNAL PROCESSING LETTERS. His areas of interest include statistical signal processing, wireless communications, communication networks and sensor networks, and information theory.