Combining Queueing Theory with Information Theory for Multiaccess

İ. Emre Telatar, *Member, IEEE,* and Robert G. Gallager, *Fellow, IEEE*

(Invited Paper)

Abstract— **We develop and analyze a multiaccess communication model over the additive Gaussian noise channel. The framework is information-theoretic; nonetheless it also incorporates some queueing-theoretic aspects of the problem.**

I. Introduction

A MULTIACCESS communication system consists of a set of transmitters sending information to a single receiver. Each transmitter is fed by an information source generating a sequence of messages; the successive messages arrive for transmission at random times. We will assume that the information sources that feed the transmitters are independent processes and that the messages generated by a given information source form an independent sequence. The signal received at the receiver is a stochastic function of the signals sent by the transmitters. We will further assume that the feedback from the receiver is limited; in particular, the possibility of any transmitter observing the received signal is ruled out.

From the description above, one sees that there are two issues of interest: (1) the random arrival of the messages to the transmitters, and (2) the noise and interference that affect the transmission of these messages. The main bodies of research in multiaccess communications seem to treat these two issues as if they were separable [1]. The collision resolution approach focuses on the random arrival of the messages but ignores noise and trivializes the interference of the transmitted signals, e.g., [2], [3]. The multiaccess information-theoretic approach, on the other hand, develops accurate models for the transmission process (noise and interference) but ignores the random arrival of the messages,[1] e.g., [4]–[7]. In addition, one can say with some oversimplification that the results generated by the two approaches are of different character. The information-theoretic results mostly state upper and lower bounds (which sometimes coincide) to the performance of the best possible scheme, whereas collision resolution results mostly analyze the performance of particular algorithms.

Manuscript received September 30, 1994; revised April 1, 1995. This work was supported by the U.S. Army Research Office, Grant DAAL03-86-K-0171.
İ. E. Telatar is with AT&T Bell Laboratories, Murray Hill, NJ 07974 USA.
R. G. Gallager is with the Electrical Engineering and Computer Science Department and the Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.
IEEE Log Number 9412646.

[1] An information theorist will rightly point out that the randomness in the arrival process should be taken care of by appropriate source coding. However, that would introduce large delays, rendering analysis of delay in communication impossible.

In this paper we present an analysis of a system with just enough complexity to exhibit both aspects of a multiaccess problem. In the analysis we use tools borrowed from both queueing theory and information theory. As a result, we are able to indicate the trade-offs between queueing-theoretic quantities and information-theoretic quantities, such as the trade-off between delay and error probability.

II. The Model and a Preliminary Analysis

Our multiaccess environment consists of an additive Gaussian noise channel with noise density $N_0/2$ and two-sided bandwidth $2W$. All the transmitters have equal power P. The messages are generated in accordance with a Poisson process of rate λ and we will assume that each message is transmitted by a different transmitter. In effect, there are an infinite number of transmitters, each handling one message. This assumption simplifies the model so that we do not have to consider message queues at individual transmitters. Each message consists of a sequence of bits of (possibly) variable length. As soon as the message arrives, the transmitter encodes it into a time signal of *infinite* duration (henceforth codeword) and starts transmitting it. However, the transmitter will not transmit for the whole duration of the signal; it will transmit only until the receiver decodes the message and instructs the transmitter to stop. (see Fig. 1.) Thus, if the system is stable, with probability one, only a finite initial segment of the infinite duration codeword will be transmitted.[2] The decoder treats each transmitter independently; each message is decoded regarding the other transmissions as noise.

If there are n active transmitters at a given time, the signal-to-noise ratio (SNR) for any of these active transmitters is $P/((n-1)P + N_0W)$. At this point let us assume that the decoder can resolve

$$W \ln\Big(1 + \frac{P}{(n-1)P + N_0W}\Big) \tag{1}$$

nats[3] per unit time for each transmitter.[4] Let us emphasize that this is not a justified assumption. Even though the SNR seen

[2] The reader will notice that we have assumed that the transmitted signals are of finite duration and also of finite bandwidth. We treat this contradiction as a technicality since the signals of real transmitters are of finite duration and bandlimited in an approximate sense. An analysis *à la* Wyner [8] would have been more desirable.

[3] We will use natural units from now on; $1\,\text{nat} = \log_2 e$ b.

[4] This is the capacity of a Gaussian channel with SNR $P/((n-1)P + N_0W)$.

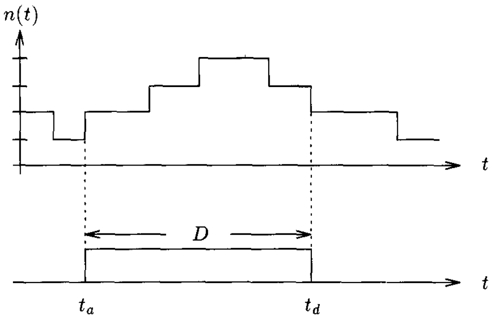

Fig. 1. Transmission of a packet in the example system. In the figure, $n(t)$ denotes the number of active transmitters. The lower illustration focuses on a particular transmitter. A message arrives at t_a and is transmitted until t_d at which time it is decoded at the receiver. The duration D of transmission is a random variable, which is dependent on the values of $n(t)$ for $t \geq t_a$.

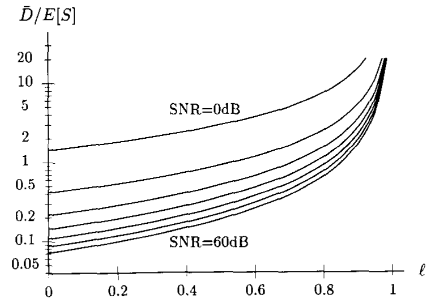

Fig. 2. Average delay as a function of loading and SNR. The seven curves correspond to different SNR values ranging from 0 dB to 60 dB in increments of 10 dB. The delay is normalized by the nat arrival rate per unit bandwidth $E[S]$.

by a particular transmitter during the time when there are a total of n active transmitters is $P/((n-1)P + N_0W)$, there is no coding theorem that guarantees the existence of a code and a decoder that can achieve the transmission rate indicated in (1). With this remark, it is clear that the analysis that follows is not rigorous. However it provides the intuitive setting in which to understand the essential ideas of a correct analysis presented in Section III. With this assumption, the decoder has a total information resolving power of

$$nW \ln\Big(1 + \frac{P}{(n-1)P + N_0W}\Big) \quad \text{nats/unit-time}$$

which it shares equally among the active transmitters. Note that the total resolving power is not a constant, but depends on the number of active transmitters.

One can liken the situation just described to that of a processor-sharing system where jobs compete for the processors time. The role of the jobs is taken by transmitters that are served by the decoder. The more transmitters that are active at a given time, lesser the rate of service each receives, since there is more interference. We can indeed formulate the problem as a classical processor-sharing system in queueing theory, with the following difference: the total service rate depends on the state of the queue through the number of jobs competing for service. This problem has been analyzed (see e.g., [9]) and we reproduce the relevant results below. Let us first define the processor-sharing model.

Suppose that jobs in a processor-sharing system arrive in accordance to a Poisson process of rate λ. Each job requires a random amount of service, S, distributed according to G

$$\Pr\{S \leq s\} = G(s).$$

The service requirements of jobs are independent. Given u jobs in the system, the server can provide service at a rate of $\phi(u) > 0$ units of service per unit time, and it divides this service rate equally among all jobs in the system. That is, whenever there are $u > 0$ jobs in the system, each will receive service at a rate of $\phi(u)/u$ per unit time. A job will depart the system when the service it has received equals its service requirement.

Theorem 1 (see e.g., [9, sec. 3.3]): For the processor-sharing model described above, the number of jobs in the system has the steady-state distribution

$$\Pr\{u \text{ jobs in the system}\} = \frac{1}{K\phi_!(u)}(\lambda E[S])^u$$

where

$$\phi_!(u) = \prod_{v=1}^{u} \phi(v) \quad \text{and} \quad K = 1 + \sum_{u=1}^{\infty} (\lambda E[S])^u / \phi_!(u)$$

provided that the infinite sum is well-defined.

In the multiaccess problem we described we can take

$$S = \text{message length in nats}/W$$

$$\phi(u) = u \ln\Big(1 + \frac{1}{(u-1) + \text{SNR}^{-1}}\Big)$$

where $\text{SNR} \stackrel{\text{def}}{=} P/(N_0W)$. The normalization by W will make the results easier to present. Since we can compute the steady-state statistics of the number of active transmitters, N, we can use Little's law to compute the average transmission duration $\bar{D}$

$$\lambda \bar{D} = E[N].$$

For a given value of SNR, the average number of active transmitters is a function of $\ell \stackrel{\text{def}}{=} \lambda E[S]$, which is the loading of the multiaccess system in terms of nats per unit time per unit bandwidth: λ is the arrival rate of the messages, and $E[S]$ is the average message length in nats per unit bandwidth, thus $\lambda E[S]$ is the nat arrival rate per unit bandwidth. Fig. 2 shows the dependence of average waiting time to the loading ℓ and the SNR. Note that since

$$\lim_{u \to \infty} u \ln\Big(1 + \frac{1}{u - 1 + \text{SNR}^{-1}}\Big) = 1$$

the sum

$$K = 1 + \sum_{u=1}^{\infty} (\lambda E[S])^u / \phi_!(u)$$

exists only when ℓ is strictly less than unity. Thus the system is stable if and only if $\ell < 1$; equivalently the throughput of the system is 1 nat per unit time per unit bandwidth. As long as the rate of information flow normalized by the bandwidth is less than 1 nat/s/Hz, the average delay will be finite and the system will eventually clear all the messages. If, on the other hand, the normalized rate of information flow is larger than 1 nat/s/Hz, the average delay will become unbounded and messages will keep accumulating in the system. This limit of 1 nat/s/Hz is a consequence of independent decoding: if n transmitters are active and we decode each as if others were noise, then the transmission rate per unit bandwidth of any transmitter cannot exceed $\ln(1+1/(n-1+\text{SNR}^{-1}))$ nats per unit time. Using the inequality $\ln(1+x) \le x$, the aggregate transmission rate per unit bandwidth is then upper bounded by $n/(n-1+\text{SNR}^{-1})$, which for large n approaches unity.

III. Analysis and Results

Recall that the heuristic analysis given in the previous section suffers from our unjustified assumption as to the decoding rate of the receiver. Here we will set things right. We had assumed that the receiver decodes each transmitter independently; we may imagine that there are as many receivers as there are transmitters, each receiver decoding its corresponding transmitter, regarding other transmitters as noise. We will choose the codewords of each transmitter as samples of bandlimited white Gaussian noise. Each receiver will know the codebook of the transmitter that it is going to decode; the signals transmitted by the other transmitters are indistinguishable from those generated by a Gaussian noise source. However, we will assume that each receiver is aware of the total number of active transmitters at all times. This is a sensible assumption: we may imagine that there is a separate channel on which the transmitters announce the start of their transmissions so that a decoder will be assigned to them. To be able to cast our system as a processor-sharing queue and use Theorem 1 to analyze it, we must identify the service demand of each transmitter and the service rate offered by the receiver to the transmitters. Intuitive candidates for these quantities are, respectively, the number of nats in the transmitters' messages and the average mutual information over the channel. The intuition behind these candidates is that the number of nats of the transmitters' messages decoded per unit time should be related to the mutual information over the channel, and thus the rate of information flow should constitute the provided service rate. In the previous section this intuitive idea was treated as fact and the results derived were based on it. It will turn out that this intuitive idea is too simplistic and we will define the demand and service differently in the following. Nonetheless, we will see that the intuitive candidates can be interpreted as the limiting case of the ones we will define.

Let us focus on a single transmitter–receiver pair, and condition on the process $u(t)$, $t \ge 0$, the number of active transmitter–receiver pairs at all times t.

The samples of the process $u(t)$ will be integer valued step functions. Let $0 \stackrel{\text{def}}{=} t_0 < t_1 < t_2 < \cdots$ be the times the process changes value and let u_k, $k \ge 1$ be the value of the process in the interval $[t_{k-1}, t_k)$.

We can model our Gaussian waveform channel of single-sided bandwidth W as a sequence of complex Gaussian *scalar* channels C_i, $i \in \mathbf{Z}$, by first bringing the waveform channel to baseband, and sampling the complex baseband waveform channel of two-sided bandwidth W at the Nyquist rate W. The channel C_i is the channel corresponding to the ith sample. Each channel C_i will be used only once. The noise for the channel C_i is a complex Gaussian random variable of uniform phase and power

$$\sigma_i^2 = N_0 + (u_k - 1)P/W, \qquad Wt_{k-1} \le i < Wt_k.$$

This expression indicates that the noise density seen by a particular transmitter–receiver pair when there are $u-1$ other active transmitters is Gaussian with intensity $N_0 + (u-1)P/W$. The (complex valued) input to this channel is limited in variance to P/W.

Note that the scalar channels are made available over time at the Nyquist rate of W per unit time. Let the number of codewords be M (i.e., the message is $\log_2 M$ b $= \ln M$ nats long).

A. Simple Decoding Rule

If we use the output of the first d channels to decode the transmitted message (i.e., decoding at time d/W),[5] we get the following random coding bound on the error probability [10, ch. 5, pp. 149–150]: for any $0 \le \rho \le 1$

$$\bar{P}_e \le \exp\Big[\rho \ln M - \sum_{i=1}^{d} E_0(\rho, \sigma_i)\Big].$$

For a complex additive Gaussian noise channel with noise variance σ^2 and Gaussian input ensemble with variance P/W

$$E_0(\rho, \sigma) = \rho \ln\Big(1 + \frac{P}{W\sigma^2(1+\rho)}\Big).$$

If we fix a $\rho \in (0, 1]$ and a tolerable error probability P_e, then we can view $-\ln P_e + \rho \ln M$ as the demand and $E_0(\rho, \sigma)$ as the service rate (per transmitter–receiver pair per degree of freedom). Note that to cast these parameters in the context of processor-sharing queues we need to express service rate in terms of total service per unit time. This leads to a service rate at time t as

$$\begin{aligned} &Wu(t)E_0(\rho, \sigma(t)) \\ &\quad = u(t)W\rho \ln\Big(1 + \frac{P}{(1+\rho)(N_0W + (u(t)-1)P)}\Big). \end{aligned}$$

Thus, we have a demand

$$S = -(\ln P_e) + \rho \ln M]$$

[5] The careful reader will notice that when the decoder instructs the transmitter to stop transmitting after the decoding time, all the transmitter can do is to force the Nyquist samples of its signal to zero. This is not the same as the transmitted signal being equal to zero after the decoding time. This arises simply from the bandlimitedness of the transmitted signal. One way to get around this is for the receiver to subtract the transmitter's signal from the received signal once it decodes it.

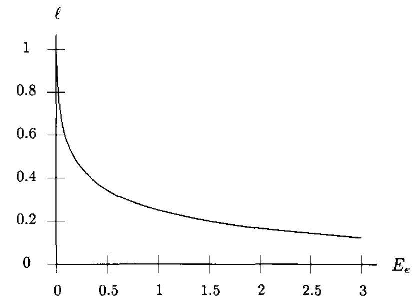

Fig. 3. Stability region of the multiaccess system with the simple decoding rule. Any (E_e, ℓ) pair under the curve belongs to the stability region. Note the sharp falloff in achievable ℓ at $E_e = 0$.

and a service rate ϕ as a function of the number of active users u given by

$$\phi(u) = W\rho u \ln\left(1 + \frac{P}{(1+\rho)(N_0 W + (u-1)P)}\right).$$

Let us first examine the stability of the multiaccess system as predicted by this analysis. For stability we need

$$\lim_{u\to\infty} \frac{\lambda E[S]}{\phi(u)} < 1.$$

Define

$$\ell \stackrel{\text{def}}{=} \lambda E[\ln M]/W$$

as the loading of the system (average nat arrival rate per unit bandwidth) and

$$E_e \stackrel{\text{def}}{=} -(\ln P_e)/E[\ln M]$$

as the error exponent.[6] Since $\phi(u) \to W\rho/(1+\rho)$ as u tends to infinity and

$$\begin{aligned} E[S] &= E[\rho \ln M - (\ln P_e)] \\ &= E[\ln M](\rho + E_e) \end{aligned}$$

we get

$$\lim_{u\to\infty} \frac{\lambda E[S]}{\phi(u)} = \ell(1 + E_e/\rho)(1+\rho)$$

and we can rewrite the stability condition as

$$\ell(1+\rho)(1+E_e/\rho) < 1 \qquad \text{for some } \rho \in (0,1].$$

This is equivalent to the statement

$$\ell \inf_{0<\rho\le 1}(1+\rho)(1+E_e/\rho) < 1.$$

This minimization occurs at $\rho = \sqrt{E_e}$ for $0 < E_e \le 1$ and at $\rho = 1$ for $E_e > 1$. The corresponding value of the infimum is

[6]This definition of the error exponent is different than the usual definition, which would define it as minus the logarithm of the error probability divided by the transmission duration. Here we normalize by the average message length rather than by the transmission duration. This is both convenient and sensible.

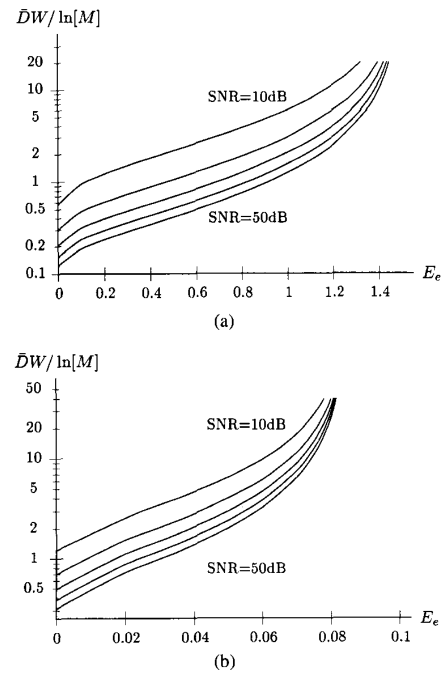

Fig. 4. Delay versus error exponent with the simple decoding rule. (a) $\ell = 0.2$. (b) $\ell = 0.6$.

$(1+\sqrt{E_e})^2$ for $0 \le E_e \le 1$ and $2(1+E_e)$ for $E_e > 1$. In sum, the stability region of the system is

$$\left\{(E_e, \ell) : E_e \in [0,1], 0 \le \ell(1+\sqrt{E_e})^2 \le 1\right\} \cup \left\{(E_e, \ell) : E_e > 1, 0 \le \ell(2+2E_e) \le 1\right\}.$$

Fig. 3 shows this stability region.

Whether one is interested in small or large values of E_e depends on the message length and the desired error probability. If one has long messages then a small value of E_e may be sufficient to drive the error probability down to acceptable levels. For short messages, however, E_e will need to be large to achieve the same error performance. Note that E_e can be made arbitrarily large if one is willing to sacrifice throughput.

Given a stable (E_e, ℓ) pair, an SNR $\stackrel{\text{def}}{=} P/(N_0 W)$, and $\rho \in (0,1]$ we can use Theorem 1 to compute the steady-state distribution of the number of active transmitters and use Little's law to compute the average delay. One can further choose the value of ρ to minimize this average delay. Fig. 4 shows the value of this optimized average delay as a function of E_e for various values of the loading ℓ and SNR.

Note that we can interpret the results of Section II as a limiting case of the results of this section by letting E_e approach 0.

B. Improved Decoding Rule

Let us examine the operation of the decoding rule we are employing above: When a transmitter becomes active, the receiver accumulates the values of $E_0(\rho, \sigma(t))$ for this transmitter, and when this sum exceeds the service demand, the receiver decodes the message. Observe that $\sigma(t)$ only depends on the number of active transmitters at time t , and thus the receiver chooses the time of decoding only on the basis of the number of interferers, and ignores what the received signal actually is in making the decision about decoding time. It can turn out that at this time for decoding the receiver finds that the received signal has been corrupted by noise and thus making correct decoding unlikely. For this reason we will consider a modification of the previous decoding rule as follows:

In this modified rule, the decoder proceeds in stages. The duration of each stage is determined by the number of interferers in that stage and at the end of each stage the decoder decides whether to decode or to proceed to the next stage. The decision to decode or to proceed is made on the basis of the received signal in the current stage. If the decoder chooses to proceed to the next stage, it disregards the signal it has received in the previous stages, and starts anew. Note that the transmitter does not need to know about the stages, it just keeps transmitting the signal corresponding to the message throughout the stages. Let P_X be the probability that at the end of a stage the receiver will not decode but will proceed to the next stage. We will call P_X the erasure probability. Then, the number of stages the decoder will take to decode the message is a geometrically distributed random variable with mean $(1-P_X)^{-1}$. If the expected service demand for the individual stages is $E[S]$, then the overall service demand will have expected value $E[S]/(1-P_X)$. To analyze this decoding rule we shall make use of a result of Forney [11]. Forney proves his results for discrete channels, but it is easy to generalize them to channels with continuous input and output alphabets as we have here: if a stage uses the output of d channels (i.e., the stage lasts d/W units of time) we have the following random coding bound error probability P_e and the erasure probability P_X: for any $0 \le s \le \rho \le 1$ and $T > 0$

$$\bar{P}_e \le \exp\left[\rho \ln M + (s-1)T - \sum_{i=1}^{d} E_0(\rho, s, \sigma_i)\right]$$

and

$$\bar{P}_X \le \exp\left[\rho \ln M + sT - \sum_{i=1}^{d} E_0(\rho, s, \sigma_i)\right].$$

Note that e^T is the ratio of the two upper bounds; the parameter T controls the trade-off between P_e and P_X. The function E_0 for a complex additive Gaussian noise channel with noise variance σ^2 and Gaussian input ensemble with variance P/W is given by

$$E_0(\rho, s, \sigma) = \rho \ln\left[1 + (s/\rho)\frac{P}{W\sigma^2}\right] + \ln\left[1 + \frac{(\rho - s - s\rho)(s/\rho)P/(W\sigma^2)}{1 + (s/\rho)P/(W\sigma^2)}\right].$$

If we fix $\rho \in (0, 1]$, $s \in (0, \rho]$ and tolerable values of P_X and P_e, we may identify $\ln(P_X/P_e)$ as T, $-\ln P_X + s \ln T + \rho \ln M$ as the demand and $E_0(\rho, s, \sigma_i)$ as service rate. We may now use Theorem 1. We first eliminate the parameter T from our definition of demand and write it only in terms of the error and erasure probabilities

$$S' = -(1-s)\ln P_X - s \ln P_e + \rho \ln M.$$

The demand thus defined does not take into account the fact that the transmission consists of multiple stages. This can be remedied by recalling that the number of transmission stages is a geometrically distributed random variable with mean $1/(1-P_X)$. Thus the expected value for the overall demand is

$$E[S] = E[S']/(1-P_X) = \lambda^{-1}\ell W(\rho + sE_e + (1-s)E_x)$$

where

$$\ell = \lambda E[\ln M]/(W(1-P_X)) \qquad \text{(loading)} \qquad (2)$$

$$E_e = -(\ln P_e)/E[\ln M] \qquad \text{(error exponent)} \qquad (3)$$

and

$$E_X = -(\ln P_X)/E[\ln M] \qquad \text{(erasure exponent).} \qquad (4)$$

As before, the service rate function $\phi(u)$ will be given by $WuE_0(\rho, s, \sigma)$ evaluated at the σ corresponding to $u-1$ interferers

$$\phi(u) = \rho W \ln\left[1 + (s/\rho)\frac{P}{N_0 W + (u-1)P}\right] + W \ln\left[1 + \frac{(\rho - s - s\rho)(s/\rho)(P/(N_0 W + (u-1)P))}{1 + (s/\rho)(P/(N_0 W + (u-1)P))}\right].$$

We proceed, just as before, by first examining the stability of the system. Noting that $\lim_{u\to\infty} \phi(u) = Ws(2 - s(1+\rho)/\rho)$ the stability condition

$$\lim_{u\to\infty} \frac{\lambda E[S]}{\phi(u)} < 1$$

reduces to

$$\ell \frac{\rho + sE_e + (1-s)E_X}{s(2 - s(1+\rho)/\rho)} < 1. \qquad \text{for some } 0 \le s \le \rho \le 1.$$

This in turn, is equivalent to

$$\ell \min_{0\le\rho\le 1} \min_{0\le s\le\rho} \frac{\rho + sE_e + (1-s)E_X}{s(2 - s(1+\rho)/\rho)} < 1.$$

The minimization over s can be done via differentiation and we obtain

$$s^* = \frac{\rho + E_X}{E_e - E_X}(\sqrt{1 + 2\beta} - 1)$$

as the minimizing value of s where $\beta = \beta(\rho) \stackrel{\text{def}}{=} \frac{\rho}{1+\rho}\frac{E_e - E_X}{\rho + E_X}$. Using the inequality $\sqrt{1+2x} \le 1 + x$, we see that $s^* \le \rho/(1+\rho)$ and thus $s^* \le \rho$ as desired. Substituting this value of s we get the following condition for stability

$$\ell \min_{0\le\rho\le 1} \frac{1}{2}(E_e - E_X)\frac{\sqrt{1+2\beta}}{(\sqrt{1+2\beta} - 1)(1 + \beta^{-1}) - 1} < 1.$$

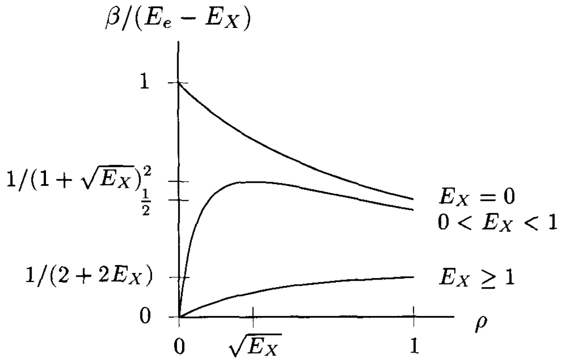

Fig. 5. β as a function of ρ for various cases of E_X .

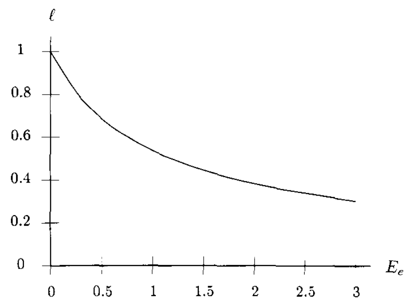

Fig. 6. Stability region of the multiaccess system as $E_X \to 0$. Any (E_e, ℓ) pair under the curve belongs to the stability region. Compare with Fig. 3 to see the enlargement of the stability region.

The quantity to be minimized is a decreasing function of β , and thus the value of ρ that minimizes the above expression is the one that maximizes β.

The nature of the mapping $\rho \mapsto \beta(\rho)$ depends on the value of E_X. There are three cases as illustrated in Fig. 5:

1) $E_X = 0$. The range of the mapping is $[\frac{1}{2}E_e, E_e]$. The maximum is achieved at $\rho = 0$.
2) $0 < E_X < 1$. The range is $[0, (E_e - E_X)/(1+\sqrt{E_X})^2]$. The maximum is achieved at $\rho = \sqrt{E_X}$.
3) $E_X \geq 1$. The range is $[0, (E_e - E_X)/(2 + 2E_X)]$. The maximum is achieved at $\rho = 1$.

Putting the above together we have the following stability conditions:

1) If $0 \leq E_X \leq 1$

$$\frac{1}{4}\ell\left(1 + \sqrt{E_X} + \sqrt{(1+\sqrt{E_X})^2 + 2(E_e - E_X)}\right)^2 < 1$$

2) if $E_X \geq 1$

$$\frac{1}{2}\ell\left(\sqrt{1+E_e} + \sqrt{1+E_X}\right)^2 < 1.$$

If we are interested in maximizing E_e for a given value of ℓ irrespective of E_X, we see that we should let E_X approach 0. The stability condition is then

$$\frac{1}{2}\ell(1 + E_e + \sqrt{1+2E_e}) < 1.$$

This region is shown in Fig. 6.

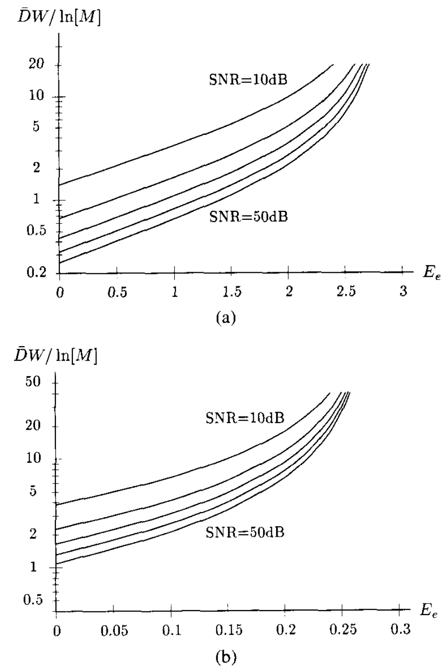

Fig. 7. Delay versus error exponent with the improved decoding rule. (a) $\ell = 0.2$, $E_x = 0.5$. (b) $\ell = 0.6$, $E_x = 0.04$.

For a comparison with the previous decoding rule consider values of ℓ close to 1. In this range E_e needs to be small, and we can approximate the stability condition to $E_e < 1 - \ell$. Compare that with the previous decoding rule: the largest error exponent the previous rule could support for large ℓ is approximately $(1-\ell)^2/4$. One should note that the comparison is meaningful only when P_X is small, since the definition of ℓ has a factor $(1 - P_X)^{-1}$ in the second case. We are thus assuming that although E_X is close to zero, the average message length is large enough to make P_X close to zero as well.

As before, we can compute the the average delay for any given E_e, E_X, SNR and ℓ. In Fig. 7 we show the normalized average delay as a function of E_e for various values of the other parameters. Points on each curve are the result of an optimization over ρ and s to yield minimum delay.

IV. Conclusion

We developed and analyzed a multiaccess communication model over the additive Gaussian noise channel. Unlike previous approaches to multiaccess we seek to combine queueing theory and information theory to arrive at our results.

The results presented here are not to be taken as a proposal to build a system which operates as described. Indeed, joint decoding of the transmitters and the feedback from the receiver can be used to greatly improve the throughput, which in our model is 1nat/s/Hz . Rather, the paper aims to demonstrate that it is possible to combine the methods of information theory with those of queueing theory to simultaneously address the two defining characteristics of multiaccess systems. We have analyzed a simplified model with a simple transmission strategy to investigate the trade-offs, such as delay versus probability of error, that have not been addressed before.

REFERENCES

[1] R. G. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 2, pp. 124–142, Mar. 1985.

[2] N. Abramson, "The ALOHA system—Another alternative for computer communications," in *1970 Fall Joint Comput. Conf., AFIPS Conf. Proc.*, vol. 37, pp. 281–285.

[3] J. I. Capetanakis, "Tree algorithms for packet broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 5, pp. 505–515, Sept. 1979.

[4] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Inform. Theory, 1971*, Tsahkadsor, Armenian S.S.R., 1973, pp. 23–52.

[5] H. Liao, "Multiple access channels," Ph.D. dissertation, Univ. Hawaii, Dept. Elec. Eng., 1972.

[6] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inform. Theory*, vol. IT-23, no. 1, pp. 1–37, Jan. 1977.

[7] A. El Gamal and T. M. Cover, "Multiple user information theory," *Proc. IEEE*, vol. 68, no. 12, pp. 1466–1483, Dec. 1980.

[8] A. D. Wyner, "The capacity of the band-limited Gaussian channel," *Bell Syst. Tech. J.*, vol. 45, pp. 359–395, Mar. 1966.

[9] F. P. Kelly, *Reversibility and Stochastic Networks, Wiley Series in Probability and Mathematical Statistics.* New York: Wiley, 1979.

[10] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[11] G. D. Forney, "Exponential error bounds for erasure, list and decision feedback schemes," *IEEE Trans. Inform. Theory*, vol. IT-14, no. 2, pp. 206–220, Mar. 1968.

İ. Emre Telatar (S'88–M'92) received the B.S. degree in electrical engineering from the Middle East Technical University, Ankara, Turkey, in 1986, the S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1988 and 1992 respectively.

Since 1992, he has been with the Communications Analysis Research Department at AT&T Bell Laboratories, Murray Hill, NJ. His research interests are in communication and information theories.

Robert G. Gallager (S'58–M'61–F'68), for a photograph and biography, please see page 935 of this issue.