Selective-Repeat ARQ: The Joint Distribution of the Transmitter and the Receiver Resequencing Buffer Occupancies

ZVI ROSBERG AND MOSHE SIDI, SENIOR MEMBER, IEEE

Abstract—**Consider a communication network which regulates retransmissions of erroneous packets by a Selective-Repeat (SR) ARQ protocol. Packets are assigned consecutive integers, and are transmitted by the transmitter in order, until a NACK or a time out is observed. The receiver, upon receipt of a packet, checks for errors and returns ACK / NACK accordingly. Only packets for which either a NACK or a time out have been observed are retransmitted.**

The overall delay of a packet in a system that operates under the SR ARQ protocol consists of the queueing delay at the transmitter and the resequencing delay at the receiver. In this paper, we present a model for the SR ARQ, and derive the joint distribution of the buffer occupancies at the transmitter and at the receiver and compare the two types of delay.

I. INTRODUCTION

THE procedures whereby computer communication networks preserve the integrity of data sent from a transmitter to a receiver over a noisy path are known as Automatic Repeat Request (ARQ) protocols. In these protocols, the data are sent by packets, each of which is encoded for error detection by the receiver. The packets that arrive at the transmitter are assigned subsequent numbers that uniquely identify them and are referred to as *identifiers*. Furthermore, the first time transmissions of packets occur in an increasing order of the packet identifiers.

Based on the error detection results, positive/negative acknowledgments (ACK/NACK) are sent by the receiver over a feedback channel, arriving at the transmitter after a roundtrip propagation delay. The acknowledgments carry the identifiers of the packets they acknowledge. If no acknowledgment is received within a predetermined interval, the transmitter interprets this as a NACK and retransmits the packet. This event is referred to as a *time out*.

The retransmissions of erroneous packets depend on the particular ARQ protocol being used. There are three basic ARQ schemes: Stop-and-Wait, Go-Back-N, and Selective-Repeat [2], [12], [15], [6]. Under the Selective-Repeat protocol, on which we focus in this paper, the transmitter continuously sends new packets and the receiver accepts every packet that arrives error-free. Upon receipt of a NACK by the transmitter, only the corresponding packet is retransmitted.

To maintain integrity, it is a common requirement in computer networks that the receiver must *send out/release* packets (to the user, to the next node, or to the upper layer of the network architecture) *in their original order*. Under the Go-Back-N protocol, packets that arrive out of order are ignored by the receiver, which consequently does not have to allocate any buffers for them. Under the Selective-Repeat protocol, those packets must be stored in the receiver buffers until they can be sent out according to their original order. The buffers needed for this purpose are referred to as resequencing buffers, and the time that packets spend there is called the resequencing delay.

There are several important performance measures that are associated with the ARQ protocols: the throughput, the packet delay, buffer occupancy at the transmitter, and buffer occupancy at the receiver. In [4], the performance under Go-Back-1 has been analyzed, and the buffer occupancy distribution obtained when the buffer capacity at the transmitter is unlimited. For a finite buffer capacity, the overflow probability has been derived, from which optimal time-out values have been obtained. In [18], [9], the distribution of the packet delay and the buffer occupancy at the receiver under Go-Back-N and SR were derived. A modified version of the Go-Back-N, the Stutter Go-Back-N, has been presented in [17]. The protocol has been proposed for links with high error rate or long propagation delay, and its average buffer occupancy has been shown to approach the corresponding value of an "idealized" ARQ protocol. The throughput under several variants of the Selective-Repeat ARQ has been studied in [10] and [19]. In [3], it has been shown that the throughput under the Selective-Repeat ARQ protocol outperforms that of the Go-Back-N over a wide range of error rates. Other related studies are [14], [7], [16], [1].

Every ARQ protocol requires the transmitter to buffer packets since arrival instances are random and retransmissions might be needed. Under the SR ARQ protocol, the receiver is also required to buffer packets which arrived out of order. Thus, the overall delay of a packet under the SR ARQ protocol consists of two parts:

1) queueing delay at the transmitter (i.e., the time between the packet's arrival and its successful transmission)

2) resequencing delay at the receiver.

The queueing delay and the buffering requirement at the transmitter have been studied in [14], [18], [9], [1]. The resequencing delay and the buffering requirement at the receiver have been studied in [13] under a "heavy traffic" assumption, namely, the transmitter is assumed to always have a packet to transmit. Therefore, this model provides upper bounds on the expected resequencing delay and the buffer requirement at the receiver.

In this study, we relax the "heavy traffic" assumption and analyze a system where packets arrive at the transmitter according to a general renewal process. The analysis of both queues enables us to compare them and to explore the intriguing problem of which (if any) of them is dominant.

Another feature of our analysis technique is the ability to evaluate the transmitter queueing delay for large window sizes. Although the statistics of this metric have been analytically derived in [9], computational complexity inhibits from obtaining results for windows larger than four. The computational complexity has been eliminated in [1] by considering an approximate solution of the model.

In this study, we analyze both the queuing delay at the transmitter and the resequencing delay at the receiver. The queueing model that

Paper approved by the Editor for Communication Networks of the IEEE Communications Society. Manuscript received August 18, 1989; revised September 9, 1989.

Z. Rosberg was on leave at the IBM T.J. Watson Research Center, Yorktown Heights, NY 10598. He is with IBM Israel, Science and Technology Center, Technion City, Haifa 32000, Israel.

M. Sidi was on leave at the IBM T.J. Watson Research Center, Yorktown Heights, NY 10598. He is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel.

IEEE Log Number 9037858.

0090-6778/90/0900-1430$01.00

we consider is slightly different from the model in [9], [1] in that an arriving packet is prohibited from transmission during the window in which it has arrived. In the next section, we elaborate on that and discuss how our model relates to the one in [9], [1]. Also, in Section V, we present several numerical examples in which a comparison is made between the results of our model and those obtained from a simulation of the original one in [9].

In Section II, we formulate the system model and discuss how it corresponds to the model in [9], [1]. In Section III, we analyze the transmitter queueing statistics, and in Section IV the joint distribution of the transmitter and the receiver buffer occupancies. From the latter, we derive the packet resequencing waiting time at the receiver. Numerical results that compare the queueing delay at the transmitter and the resequencing delay at the receiver are presented in Section V.

II. Model Formulation

Consider a pair of nodes, a transmitter and a receiver, which communicate data packets of fixed lengths through a slotted channel, and where packet transmission times equal one slot. Moreover, the forward channel and the feedback channel (through which acknowledgments are transmitted) are both noisy. We assume that the nodes are synchronized with the channel and packet transmissions start at slot beginnings.

New packets that arrive at the transmitter are assigned consecutive integers that serve as their identifiers. We assume that the transmitter and the receiver have unlimited buffer capacities, and that they regulate the retransmissions of the erroneous packets according to the following Selective-Repeat ARQ protocol. The transmitter transmits new packets in increasing order, as long as ACK's are received for the transmitted packets and new packets are available. Upon receipt of a packet, the receiver checks for errors and returns an ACK/NACK accordingly. Given the transmission of a packet, say packet i, the transmitter waits for an acknowledgment of that packet, until after $M - 1$ slots, $1 \leq M < \infty$. (Observe that in a synchronized system, an acknowledgment is due to arrive exactly at that moment; otherwise, a time out is assumed.) If an ACK arrives, the corresponding packet is released from the transmitter buffer and the next transmission is of a new packet (if available). If a NACK or no acknowledgment has been observed (time out), the next transmission is again packet i.

The value M is referred to in the literature as the *window size*. In the receiver buffer, a packet is released if and only if *all* preceding packets have been released. Thus, at any given time, the receiver buffer holds all packets that have been received correctly, but for which at least one packet with a lower identifier has not yet been received. The set of packets held by the receiver is denoted as the *resequencing buffer occupancy*. This model, commonly used in the study of satellite channels, has been adopted in most of the previous studies of the SR ARQ protocol, e.g., [9], [1].

For analysis tractability, we impose the following restriction. *A newly arriving packet is inhibited from transmission during the window in which it has arrived.* It becomes available for transmission at the subsequent window. All packets that have been transmitted at least once are always available for transmission.

Let A_i, $i \geq 1$, be the number of new packets that arrive during slot number i. We assume that $(A_i,\ i \geq 1)$ are independent and identically distributed (i.i.d.) random variables, with finite first and second moments λ and ξ, respectively.

We consider the contents of the transmitter and the receiver buffers at window boundaries. The windows are numbered and denoted by t, $t = 1, 2, 3, \cdots$.

Let $N(t)$ be the number of packets in the transmitter buffer at the beginning of window t. To denote the packets that are being transmitted during window t, it is mathematically convenient to introduce *dummy* packets whenever $N(t) < M$. In that case, we add $M - N(t)$ dummy packets that assume the identifiers of the subsequent packets yet to arrive. Every new arrival replaces its corresponding dummy packet. Dummy packets are not transmitted, and therefore are not subject to errors. This convention facilitates the description of the resequence buffer content and the evolution of the underlying Markov chain.

Let $X(t) = (X_1(t), X_2(t), \cdots, X_M(t))$ be the M smallest identifiers of the packets that are present at the transmitter at the beginning of window t (including the dummy packets whenever $N(t) < M$). Without loss of generality, $X_1(t) < X_2(t) < \cdots < X_M(t)$. Note that since new arrivals during window t are inhibited from transmission during that window, only the real packets among $(X_1(t), X_2(t), \cdots, X_M(t))$ are being transmitted during window t. The process $\{(N(t), X(t)),\ t = 1, 2, \cdots\}$ governs the evolution of the transmitter and the receiver buffer occupancies at window boundaries. To illustrate our notation, suppose that $M = 7$, $N(1) = 5$, and $X(1) = (1, 2, 3, 4, 5, 6, 7)$. That is, packets 6 and 7 are dummy, and packets $1, 2, \cdots, 5$ are being actually transmitted during window 1. Suppose also that the transmissions of packets 2 and 4 fail, i.e., the transmitter observes for them a NACK or a time out. Under the SR ARQ protocols, packet 1 is released from the receiver buffer, while packets 3 and 5 are held there. Packets 2 and 4 are retransmitted in the next window. If four new packets arrive during window 1, then $N(2) = 6$ and $X(2) = (2, 4, 6, 7, 8, 9, 10)$ where packet 10 is a dummy packet.

Assume that the probability of a packet failure is p, $0 < p < 1$, and that failures are mutually independent.

Let

$$D_i(t) = X_{i+1}(t) - X_i(t), \qquad i = 1, 2, \cdots, M - 1;$$

$$D_M(t) = 1; \text{ and } W_k(t) = \sum_{i=k}^{M} D_i(t), \qquad k = 1, 2, \cdots, M.$$

Also let

$$P_k^t(w, n) = \Pr\left(W_k(t) = w,\ N(t) = n\right) \quad \text{and}$$

$$G_k^t(x, y) = E\left(x^{W_k(t)} \cdot y^{N(t)}\right), \qquad |x|, |y| \leq 1.$$

Note that $W_k(t) - (M - k + 1)$ is the number of packets that are held by the receiver at the beginning of window t due to the kth packet at the transmitter. When $t \to \infty$, the limiting random variables, the probabilities, and the probability generating functions are denoted by N, D_i, W_k, $P_k(w, n)$, and $G_k(x, y)$, respectively. From [9] and [13], it follows that if $\lambda < 1 - p$, the system is ergodic and the limits exist.

Let $X_{\min}(t)$ ($X_{\max}(t)$) be the minimum (maximum) value of the packet identifiers that have been transmitted up to and including window t, but have not been yet released by the receiver. This includes the dummy identifiers if $N(t) < M$.

From the protocol description, it immediately follows (as in [13]) that for every t,

$$X_{\min}(t) = X_1(t) \quad \text{and} \quad X_{\max}(t) = X_M(t). \tag{1}$$

These equations imply that during every window t, the packets (including dummy ones, if present) with the global maximum identifier and the global minimum identifier (not yet released by the receiver) are transmitted.

It is clear that the number of packets in the receiver buffer at the beginning of window t is $W_1(t) - M$. For example, suppose that $M = 7$, $N(t) = 5$, and $(X_1(t), \cdots, X_7(t)) = (7, 8, 11, 16, 17, 20, 21)$. That is, packets 20 and 21 are dummies. From (1), this state indicates that all packets whose identifiers are less than or equal to 6 have been released from the receiver buffer. Furthermore, packets 9, 10, 12, 13, 14, and 15 have arrived at the receiver and have been positively acknowledged by it. Moreover, packets 18 and 19 were also positively acknowledged by the receiver. Since packets are released according to their identifier's order, they have to be stored at the receiver at the beginning of window t. Hence, the receiver buffer occupancy is $W_1(t) - M = X_{\max}(t) + 1 - X_{\min}(t) - M = (22 - 7) - 7 = 8$ packets. We denote by $B(t)$ the resequencing buffer occupancy at the beginning of window t, i.e.,

$$B(t) = W_1(t) - M. \tag{2}$$

The evolution of the transmitter buffer occupancy $(N(t),\ t \geq 1)$ is given by

$$N(t+1) = N(t) + A(t) - V(t), \qquad t = 1, 2, \cdots \tag{3}$$

where $N(t)$ is the number of packets at the beginning of window t, $A(t)$ is the number of packets that arrive at the transmitter buffer during window t, and $V(t)$ is the number of packets that are successfully transmitted during window t.

As one can see from (3), the main advantage of considering the system at window boundaries is that the states at the transmitter buffer are described by a one-dimensional Markov chain, not as in [9], [1] where an M-dimensional Markov chain is required. This description is crucial when the joint distribution of the transmitter and the receiver queueing processes are to be analyzed.

Notice that our model is not an exact description of the real SR ARQ system as in [9], [1]. Therefore, it is necessary to compare the results derived from the analysis of our model and the performance of the real protocol. To that end, we carry out a simulation of the real protocol whose results are presented in Section V. Here, we discuss how the two models relate to each other.

Regarding the buffer occupancy at the transmitter, our model implies that $N(t)$ has the same evolution as if packets arrive and depart at window boundaries [see (3)], and all packets are transmitted during the same time and each of them requires M slots. Therefore, to evaluate the expected delay at the transmitter (excluding transmission time), we have to subtract M slots due to the transmission time. Note that similarly to [9], [1], a packet whose transmission fails stays in the transmitter queue until the end of the window.

Using the interpretation above, the computed delay at the transmitter is measured in whole units of window size. In particular, packets in the real SR ARQ that experience a delay which is less than a window size have a delay of zero in our model. Hence, our model predicts smaller delays at the transmitter. To compensate for that phenomenon, we add a corrective term to the expected delay at the transmitter that is derived by our analysis. This term is heuristic and approximates the average delay of packets whose delay is less than one window size times the proportion of those packets. For details, see Section III. Very good agreement with the real SR ARQ is observed in all of our numerical examples.

Regarding the expected resequencing delay, it appears that our model captures the correct behavior of the real system. This is indeed verified in our examples in Section V.

In the next section, we shall derive the distribution of $N(t)$ under stationary conditions, and in Section IV, we analyze the joint distribution of the transmitter and the receiver buffer occupancies under stationary conditions.

III. The Transmitter Buffer Statistics

The chain $(N(t), t \geq 1)$ whose evolution is given in (3) is a Markov chain. From our assumptions, $(A(t),\ t \geq 1)$ is a sequence of independent and identically distributed random variables with first and second moments of λM and $\xi M + \lambda^2 M(M-1)$, respectively. Let $F(y)$ be the generating function of $A(t)$, i.e., $F(y) = E[y^{A(t)}]$. Obviously, $F(y) = [f(y)]^M$ where $f(y)$ is the generating function of the number of packets arriving in a slot. We restrict our attention to generating functions $F(y)$ such that $E[(1+\epsilon)^{A(t)}] < \infty$ for an arbitrarily small $\epsilon > 0$. (This is not a hard restriction, and it is required for applying Rouche's theorem below.) Also, let V_i, $0 \leq i \leq M$, be independent Bernoulli numbers that assume the value 0 with probability p and the value 1 with probability $q = 1 - p$. With the notation $\bar{N}(t) = \min\{N(t), M\}$, $V(t)$ is clearly given by

$$V(t) = \sum_{i=0}^{\bar{N}(t)} V_i. \tag{4}$$

From [9], the Markov chain $(N(t),\ t \geq 1)$ is irreducible and aperiodic and is ergodic if and only if

$$\lambda < 1 - p. \tag{5}$$

We assume that condition (5) holds.

Let g_i be the probability of having i packets at the transmitter at the beginning of an arbitrary window under stationary conditions, and let $G(y) = \lim_{t\to\infty} E[y^{N(t)}] = \sum_{i=0}^{\infty} g_i y^i$. From (3) and (4), we obtain

$$G(y) = F(y)\left[\sum_{i=0}^{M-1} g_i y^i (qy^{-1} + p)^i + \left[G(y) - \sum_{i=0}^{M-1} g_i y^i\right](qy^{-1} + p)^M\right]. \tag{6}$$

By rearranging (6), we have

$$G(y) = F(y)\frac{\sum_{i=0}^{M-1}\left[(q+py)^i y^M - (q+py)^M y^i\right] g_i}{y^M - F(y)(q+py)^M}. \tag{7}$$

In (7), the probabilities g_i, $i = 0, 1, 2, \cdots, M-1$ are unknown. Yet, standard application of Rouche's theorem and exploiting the analyticity of $G(y)$ within the unit disk $|y| \leq 1$ (for details, see [5, pp. 121–124]) yields these probabilities. Specifically, using the above condition on $F(y)$, Rouche's theorem implies that the denominator of (7) has exactly M zeros, $\sigma_1, \sigma_2, \cdots, \sigma_M$, within the unit disk. Furthermore, when condition (5) holds, all the zeros are distinct. If they were not, then the first derivative of the denominator of (7) at one of the zeros would also vanish. It is easy to verify that this contradicts the ergodicity condition (5). One root is clearly 1, and without loss of generality, let $\sigma_M = 1$. Since $G(y)$ is analytic within the unit disk, its numerator vanishes whenever the denominator vanishes (within the unit disk). Thus,

$$\sum_{i=0}^{M-1}\left[(q+p\sigma_j)^i \sigma_j^M - (q+p\sigma_j)^M \sigma_j^i\right] g_i = 0, \qquad 1 \leq j \leq M-1. \tag{8}$$

Another equation is obtained from the normalization condition $G(1) = 1$. By applying L'hopital's rule in (7), we have

$$M(1-\lambda-p) = (1-p)\sum_{i=0}^{M-1}(M-i)g_i. \tag{9}$$

Equations (8) and (9) form a set of M independent linear equations whose solution yields the M unknowns probabilities g_i, $i = 0, 1, 2, \cdots, M-1$. The independence is verified by checking the positivity of the corresponding determinant as in [5, pp. 121–124]. An alternative method to determine the unknown probabilities is via the matrix-geometric approach [11].

Once the probabilities g_i, $i = 0, 1, 2, \cdots, M-1$ are determined, $G(y)$ is known, and the expected number of packets in the transmitter under stationary conditions $E[N]$ is given by $\dot{G}(y) = dG(y)/dy$ at $y = 1$. By taking the derivative in (7) and using L'hopital's rule twice, we obtain,

$$E[N] = \frac{(1-p^2)\left\{\sum_{i=0}^{M-1}\left[M(M-1) - i(i-1)\right]g_i - M(M-1)\right\} + M[\xi + \lambda(2M - \lambda M - \lambda - 1)]}{2M(1-p-\lambda)}. \tag{10}$$

The expected holding time (in slots) of a packet at the transmitter $E[H_T]$ is computed from (10) by Little's law:

$$E[H_T] = \frac{E[N]}{\lambda}. \tag{11}$$

In our model, each packet that is ready to be transmitted in some frame leaves the queue at the end of the frame. Hence, the expected delay time (not including service) of a packet at the transmitter $E[D_T]$ is

$$E[D_T] = E[H_T] - M. \tag{12}$$

From the discussion in Section II, $E[D_T]$ predicts an expected delay which is smaller than the one in the real SR ARQ. Therefore, we introduce the following heuristic corrective term.

It seems promising to add the term

$$E\left[D_{M/D/1} \mid D_{M/D/1} \le M\right] \cdot \Pr\left[D_{M/D/1} \le M\right]$$

where $D_{M/D/1}$ is the delay of a packet in an $M/D/1$ queue. This term attempts to take into account the events within a window that generate delays. The term above is more complex to compute than the corresponding one in the $M/M/1$ queue. Therefore, since $E[D_{M/D/1}] = \frac{1}{2}E[D_{M/M/1}]$, we propose the following corrective term:

$$\frac{1}{2}E\left[D_{M/M/1} \mid D_{M/M/1} \le M\right] \cdot \Pr\left[D_{M/M/1} \le M\right]$$
$$= \frac{\lambda}{2(1-\lambda)} - \frac{\lambda}{2}\left(M + \frac{1}{1-\lambda}\right)\exp\{-(1-\lambda)M\}. \tag{13}$$

This simple corrective term yields an excellent fit to the simulation results of the real SR ARQ.

Remark 3.1: The expected time that is required for a packet to reach the receiver is $E[D_T] + E[S]$ where S is the transmission time plus the propagation delay. In satellite communication systems, $E[S]$ is about half of the window size. In terrestrial networks, it is about one slot. This observation is crucial when one compares the resequencing delay to the delay at the transmitter.

IV. The Joint Transmitter and Receiver Buffer Statistics

In this section, we derive a finite recursive procedure for computing the joint probability generation functions (pgf) of (W_k, N) and $E[W_k]$ for $k = 1, 2, \cdots, M$ under stationary conditions.

A. The Joint Probability Generating Function – Recursive Procedure

Let $m(t)$ be the number of NACK's received at the transmitter during window t. Since NACK's can be obtained only for real packets, and no more than M packets can be transmitted during a window, $0 \le m(t) \le \bar{N}(t)$ where $\bar{N}(t) = \min\{N(t), M\}$.

From the definition, $W_M(t) = 1$. The evolution of $(W_{M-i}(t), t \ge 1)$, $1 \le i < M - 1$ is governed by the following events.

1) For $\bar{N}(t) < M - i$ (fewer than $M - i$ packets in the transmitter buffer at the beginning of window t), $W_{M-i}(t+1) = i + 1$, independently of the number of NACK's that have been obtained. The reason is that in this event, only new or dummy packets contribute to $W_{M-i}(t+1)$.

2) For $\bar{N}(t) \ge M - i$ (at least $M - i$ packets in the transmitter buffer at the beginning of window t), if $m(t) < M - i$ (there were fewer than $(M - i)$ NACK's), then $W_{M-i}(t+1) = i + 1$. Again, only new or dummy packets contribute to $W_{M-i}(t+1)$.

3) For $\bar{N}(t) \ge M - i$ (at least $(M - i)$ packets in the transmitter buffer at the beginning of window t), if $m(t) \ge M - i$ (there were at least $(M - i)$ NACK's) *and* the $(M - i)$th NACK was of packet $X_l(t)$, $M - i \le l \le M - i + \bar{N}(t) - m(t)$, then $W_{M-i}(t+1) = W_l(t) + (\bar{N}(t) - m(t))$. The reason is that in this event, the lth packet at the beginning of window t becomes the $(M - i)$th packet at the beginning of window $t + 1$. In addition, there are $\bar{N}(t) - m(t)$ packets with higher identifiers that were positively acknowledged, and are held at the receiver due to the $X_l(t)$ failure.

To illustrate the evolution described above, consider the following example. Let $M = 10$, $N(t) = 7$ (hence $\bar{N}(t) = 7$), and $X(t) = (6, 8, 11, 16, 17, 20, 21, 24, 25, 26)$. Therefore, $W(t) = (21, 19, 16, 11, 10, 7, 6, 3, 2, 1)$, and we know that packets $7, 9, 10, 12, 13, 14, 15, 18, 19, 22, 23$ are waiting at the receiver (note that packets 24, 25, 26 in the transmitter are dummy packets). Assume that the transmissions of packets 8, 16, 17 fail, and that five new packets arrive at the transmitter. Thus, $N(t+1) = 8$, $\bar{N}(t+1) = 8$, $X(t+1) = (8, 16, 17, 24, 25, 26, 27, 28, 29, 30)$, and $W(t+1) = (23, 15, 14, 7, 6, 5, 4, 3, 2, 1)$. Note that packet 7 is released from the receiver (since packet 6 arrived correctly), and packets 9–15 and 18–23 are held at the receiver at the beginning of window $t + 1$. Observe that $W_{10}(t) = 1$, $W_9(t+1) = 2$, and $W_8(t+1) = 3$ due to event 1) above. Also, $W_7(t+1) = 4$, $W_6(t+1) = 5$, $W_5(t+1) = 6$, and $W_4(t+1) = 7$ due to even 2) above; and $W_3(t+1) = 14$, $W_2(t+1) = 15$, and $W_1(t+1) = 23$ due to event 3) above. For instance, for $W_2(t+1)$, we note that the second $(M - i = 2)$ failure was of the fourth packet $(l = 4)$; hence, $W_2(t+1) = W_4(t) + \bar{N}(t) - m(t) = 11 + 7 - 3 = 15$.

From the above, we have

$$W_M(t+1) = W_M(t) = 1, \tag{14}$$

and by using (3), we have for $1 \le i \le M - 1$,

$$\begin{aligned} G_{M-i}^{t+1}(x, y) = &\sum_{n=0}^{M-i-1} \sum_{m=0}^{n} \sum_{w=i+1}^{\infty} P_{M-i}^{t}(w, n) \\ &\cdot x^w y^n b(n, m) x^{-(w-i-1)} y^{-(n-m)} F(y) \\ &+ \sum_{n=M-i}^{\infty} \sum_{m=0}^{M-i-1} \sum_{w=i+1}^{\infty} P_{M-i}^{t}(w, n) \\ &\cdot x^w y^n b(\bar{n}, m) x^{-(w-i-1)} y^{-(\bar{n}-m)} F(y) \\ &+ \sum_{n=M-i}^{\infty} \sum_{m=M-i}^{\bar{n}} \sum_{l=M-i}^{M-i+\bar{n}-m} \sum_{w=M-l+1}^{\infty} P_l^t(w, n) \\ &\cdot x^w y^n c_{M-i}(l, m, \bar{n}) x^{\bar{n}-m} y^{-(\bar{n}-m)} F(y) \end{aligned} \tag{15}$$

where $\bar{n} = \min\{n, M\}$, $b(n, m)$ is the probability of obtaining m NACK's when n packets are transmitted (i.e., $b(n, m) = \binom{n}{m} p^m (1-p)^{n-m}$) and $c_{M-i}(l, m, n)$ is the probability of obtaining m NACK's when n packets are transmitted *and* the $(M - i)$th NACK is of the lth packet [i.e., $c_{M-i}(l, m, n) = b(l - 1, M - i - 1) \cdot p \cdot b(n - l, m - M + i)$].

The first term on the right-hand side of (15) corresponds to event 1) above. In this event, if $W_{M-i}(t) = w$, then it is reduced by $w - i - 1$; hence, we multiply x^w by $x^{-(w-i-1)}$. Also, with probability $b(n, m)$, there are $n - m$ successful transmissions ($n - m$ packets leave the transmitter buffer); hence, we multiply y^n by $y^{-(n-m)}$. The second term on the right-hand side of (15) corresponds to event 2) above, and it is derived similarly, with the notification that the number of successful transmissions cannot exceed the window size M. The third term on the right-hand side of (15) corresponds to event 3) above. Here, $W_{M-i}(t+1)$ is increased [with respect to $W_l(t)$] by $\bar{n} - m$; hence, we multiply x^w by $x^{\bar{n}-m}$. The term $F(y)$ in (15) corresponds to new packets that arrive at the transmitter during window t. Notice that the summations over w start from $w = M - l + 1$ since the minimum value of W_l is $M - l + 1$.

When $\lambda < 1 - p$, the limiting distribution of $(W_1(t), W_2(t), \cdots, W_M(t), N(t))$ exists (see [9] and [13]), and by letting $t \to \infty$, it follows from (15) that for $1 \le i \le M-1$,

$$\begin{aligned} G_{M-i}(x,y) = &\sum_{n=0}^{M-i-1}\sum_{m=0}^{n}\sum_{w=i+1}^{\infty} P_{M-i}(w,n) \\ &\cdot x^{i+1} y^m b(n,m) F(y) \\ &+ \sum_{n=M-i}^{\infty}\sum_{m=0}^{M-i-1}\sum_{w=i+1}^{\infty} P_{M-i}(w,n) \\ &\cdot x^{i+1} y^{n-\bar{n}+m} b(\bar{n},m) F(y) \\ &+ \sum_{n=M-i}^{\infty}\sum_{m=M-i}^{\bar{n}}\sum_{l=M-i}^{M-i+\bar{n}-m}\sum_{w=M-l+1}^{\infty} P_l(w,n) \\ &\cdot x^{w+\bar{n}-m} y^{n-\bar{n}+m} c_{M-i}(l,m,\bar{n}) F(y). \end{aligned} \tag{16}$$

From (14), we immediately obtain

$$G_M(x,y) = xG(y), \tag{17}$$

and after tedious algebraic manipulations, we obtain from (16) (for $1 \le i \le M-1$),

$$G_{M-i}(x,y) = F(y)\frac{H_{M-i}(x,y) + \sum_{n=M-i}^{M-1} G_{M-i}^{(n)}(x)\left[y^i \sum_{m=M-i}^{n} y^m x^{n-m} c_{M-i}(M-i,m,n) - p^{M-i}(py+qx)^i y^n\right]}{y^i - F(y)p^{M-i}(py+qx)^i} \tag{18}$$

where $G_{M-i}^{(n)}(x) = \sum_{w=i+1}^{\infty} P_{M-i}(w,n)x^w$ and $H_{M-i}(x,y)$ is given by

$$\begin{aligned} H_{M-i}&(x,y) \\ &= y^i x^{i+1}\Bigg\{\sum_{n=0}^{M-i}(q+py)^n g_n \\ &+ \sum_{n=M-i+1}^{M-1} g_n \sum_{m=0}^{M-i-1} b(n,m) y^m \\ &+ \left[G(y) - \sum_{n=0}^{M-1} g_n y^n\right]\sum_{m=0}^{M-i-1} b(M,m) y^{m-M}\Bigg\} \\ &+ \sum_{m=M-i}^{M-1} y^{i+m} \sum_{l=M-i+1}^{M-i+M-1-m}\ \sum_{n=m+l-M+i}^{M-1} x^{n-m} c_{M-i} \\ &\cdot (l,m,n) G_l^{(n)}(x) + \sum_{m=M-i}^{M} y^{m-M+i} x^{M-m} \\ &\cdot \sum_{l=M-i+1}^{M-i+M-m} c_{M-i}(l,m,M) \\ &\cdot \left[G_l(x,y) - \sum_{n=l}^{M-1} y^n G_l^{(n)}(x) - x^{M-l+1}\sum_{n=0}^{l-1} g_n y^n\right] \\ &- x^{i+1} p^{M-i}(py+qx)^i \\ &\cdot \sum_{n=0}^{M-i-1} g_n y^n - y^M x^{i+1} p^{M-i} g_{M-i}. \end{aligned} \tag{19}$$

Note that the probabilities g_i, $i = 0, 1, 2, \cdots, M-1$ were determined in Section III and by definition $G_{M-i}^{(n)}(1) = g_n$.

From (18) and (19), we see that $G_{M-i}(x,y)$ is expressed in terms of $G(y)$ [which is known], $G_l(x,y)$ and $G_l^{(n)}(x)$ for $M-i+1 \le l \le M$, $l \le n \le M-1$, and of $G_{M-i}^{(n)}(x)$ for $M-i \le n \le M-1$. We now show how to compute $G_{M-i}(x,y)$, $i = 1, 2, \cdots, M-1$ in a recursive manner. Suppose, for the moment, that $G_l(x,y)$ and $G_l^{(n)}(x)$ for $M-i+1 \le l \le M$, $l \le n \le M-1$ were determined. Then from (19), $H_{M-i}(x,y)$ is known. To compute $G_{M-i}(x,y)$, we only need to determine $G_{M-i}^{(n)}(x)$ for $M-i \le n \le M-1$ [see (18)]. This is done by applying Rouche's theorem again in a similar manner as in Section III. That is, we first consider the denominator of (18):

$$y^i - F(y)p^{M-i}(py+qx)^i = 0. \tag{20}$$

By Rouche's theorem, for any $|x| \le 1$, (20) has exactly i roots within the unit disk $|y| \le 1$. Denote these roots by $\sigma_{M-i,1}(x), \sigma_{M-i,2}(x), \cdots, \sigma_{M-i,i}(x)$. Since $G_{M-i}(x,y)$ is an analytic function within the polydisk ($|x| \le 1$, $|y| \le 1$), its numerator vanishes whenever the denominator vanishes (within this region). Thus,

$$\begin{aligned} \sum_{n=M-i}^{M-1} G_{M-i}^{(n)}(x)\Bigg[&\sigma_{M-i,j}^i(x) \sum_{m=M-i}^{n} \sigma_{M-i,j}^m(x) \\ &\cdot x^{n-m} c_{M-i}(M-i,m,n) \\ &- p^{M-i}\left[p\sigma_{M-i,j}(x) + qx\right]^i \sigma_{M-i,j}^n(x)\Bigg] \\ &+ H_{M-i}(x, \sigma_{M-i,j}(x)) = 0, \qquad 1 \le j \le i. \end{aligned} \tag{21}$$

In (21), we have a set of i linear equations for the i unknowns $G_{M-i}^{(n)}(x)$, $M-i \le n \le M-1$. By solving this set of equations, we obtain $G_{M-i}^{(n)}(x)$ for $M-i \le n \le M-1$ and thus $G_{M-i}(x,y)$ is determined.

Starting with (17), the procedure described above yields $G_{M-1}(x,y), G_{M-2}(x,y), \cdots, G_1(x,y)$ in a recursive manner. Note that along with this procedure, we also compute $G_{M-i}^{(n)}(x)$, $M-i \le n \le M-1$ for $i = 1, 2, \cdots, M-1$.

B. The Expected Number of Packets in the Resequencing Buffer

The expected number of packets in the resequencing buffer under stationary conditions is given by [see (2)]

$$E[B] = E[W_1] - M \tag{22}$$

and by Little's law, the expected delay of a packet at the receiver $E[D_R]$ is

$$E[D_R] = \frac{E[B]}{\lambda}. \tag{23}$$

To derive $E[W_1]$ recursively, we compute $E[W_{M-i}]$, $1 \le i \le M-1$. First note that $E[W_M] = 1$. Then, substituting $y = 1$ in (18) and computing $\dot{G}_{M-i}(1,1) = dG_{M-i}(x,1)/dx$ at $x = 1$, we obtain

$$\begin{aligned} E[W_{M-i}] = &\frac{\dot{H}_{M-i}(1,1) + \sum_{n=M-i}^{M-1} \dot{G}_{M-i}^{(n)}(1)\left[\sum_{m=M-i}^{n} c_{M-i}(M-i,m,n) - p^{M-i}\right]}{1 - p^{M-i}} \\ &+ \frac{\sum_{n=M-i}^{M-1} g_n\left[\sum_{m=M-i}^{n}(n-m)c_{M-i}(M-i,m,n) - ip^{M-i}q\right] + ip^{M-i}q}{1 - p^{M-i}} \end{aligned} \tag{24}$$

where $\dot{G}^{(n)}_{M-i}(1) = dG^{(n)}_{M-i}(x)/dx\,|_{x=1}$ and $\dot{H}_{M-i}(1,1) = dH_{M-i}(x,1)/dx\,|_{x=1}$. Note that in the derivation of (24), we used the fact that at $x = y = 1$, the denominator and the numerator in (18) are equal. The computation of $\dot{H}_{M-i}(1,1)$ is straightforward from (19):

$$\begin{aligned}\dot{H}_{M-i}(1,1) = (i+1)\Bigg\{&\sum_{n=0}^{M-i} g_n + \sum_{n=M-i+1}^{M-1} g_n \sum_{m=0}^{M-i-1} b(n,m)\\&+\left[1-\sum_{n=0}^{M-1} g_n\right]\sum_{m=0}^{M-i-1} b(M,m)\Bigg\}\\&+\sum_{m=M-i}^{M-1}\sum_{l=M-i+1}^{M-i+M-1-m}\\&\cdot\sum_{n=m+l-M+i}^{M-1}\left[(n-m)c_{M-i}(l,m,n)g_n\right.\\&\left.+c_{M-i}(l,m,n)\dot{G}_l^{(n)}(1)\right]\\&+\sum_{m=M-i}^{M}\sum_{l=M-i+1}^{M-i+M-m} c_{M-i}(l,m,M)\\&\cdot\Bigg[(M-m)\left[1-\sum_{n=0}^{M-1} g_n\right]+E[W_l]\\&-\sum_{n=l}^{M-1}\dot{G}_l^{(n)}(1)-(M-l+1)\sum_{n=0}^{l-1} g_n\Bigg]\\&-\sum_{m=M-i}^{M}(M-m+i+1)\\&\cdot c_{M-i}(M-i,m,M)\sum_{n=0}^{M-i-1} g_n\\&-(i+1)p^{M-i}g_{M-i}.\end{aligned}\tag{25}$$

From (24) and (25), it is seen that if $\dot{G}^{(n)}_{M-i}(1)$, $1 \le i \le M-1$, $M-i \le n \le M-1$ were known, then $E[W_{M-i}]$ could have been recursively computed for $i = 1, 2, \cdots, M-1$. The procedure for computing $\dot{G}^{(n)}_{M-i}(1)$, $1 \le i \le M-1$, $M-i \le n \le M-1$ is also recursive, and it is given in the Appendix. The procedure for computing $\dot{G}^{(n)}_{M-i}(1)$ is simple and straightforward in principle. However, it is numerically unstable. The source of the instability is that for small error probabilities (which are of interest in practice), the roots of (20) when $x = 1$ are very small (their order of magnitude is p^{M-i}), and as a result, the solution of (34) in the Appendix is unstable. To overcome this computational difficulty, we develop upper and lower bounds for $\dot{G}^{(n)}_{M-i}(1)$ that are easily computed. By replacing $\dot{G}^{(n)}_{M-i}(1)$ in (24) with these bounds, one obtains upper and lower bounds for $E[W_{M-i}]$. As will be seen in our numerical examples, for error probabilities that are not too large, the upper and lower bounds are indistinguishable.

A lower bound for $\dot{G}^{(n)}_{M-i}(1)$ is obtained by noting that the minimum value that W_{M-i} may assume is $i+1$. Therefore,

$$\begin{aligned}\dot{G}^{(n)}_{M-i}(1) &= \sum_{w=i+1}^{\infty} wP_{M-i}(w,n)\\&\ge (i+1)\sum_{w=i+1}^{\infty} P_{M-i}(w,n) = (i+1)g_n.\end{aligned}\tag{26}$$

This lower bound can be tightened by considering a system in which at the end of every window, the packets that were present at the beginning of the window are released; only packets that arrived at the receiver out of order during the current window are kept at the receiver. Since in our numerical examples we use the bound from (26), we present the result without its derivation. (The derivation goes along the same lines as the derivation of Section IV-A.)

$$\begin{aligned}\dot{G}^{(n)}_{M-i}(1) \ge (i+1)g_n + &\sum_{n_1=M-i+1}^{M-1} g_{n_1}\\&\cdot\operatorname{Prob}\{N(t)=n/N(t-1)=n_1\}f(n_1)\\&+\Bigg[g_n-\sum_{n_1=0}^{M-1} g_{n_1}\\&\cdot\operatorname{Prob}\{N(t)=n/N(t-1)=n_1\}\Bigg]f(M)\end{aligned}\tag{27}$$

where $f(n_1) = \sum_{k=M-i}^{n_1}[(n_1+M-i-k)b(n_1,k) - p\cdot k\cdot b(k-1, M-i-1)]$ and $\operatorname{Prob}\{N(t) = n/N(t-1) = n_1\}$ is easily computed from the arrival process.

An upper bound for $\dot{G}^{(n)}_{M-i}(1)$ is obtained by considering a system in which dummy packets are perceived as real packets, i.e., a system in which the transmitter always will have at least M real packets. In this system, let W^U_{M-i} be the random variable that corresponds to W_{M-i}. For every window t, $W_{M-i}(t) \le W^U_{M-i}(t)$. Therefore, under stationary conditions,

$$\begin{aligned}\dot{G}^{(n)}_{M-i}(1) &= \sum_{w=i+1}^{\infty} wP_{M-i}(w,n)\\&= \sum_{w=i+1}^{\infty} w\operatorname{Prob}\{W_{M-i}(t)=w,\ N(t)=n\}\\&= \operatorname{Prob}\{N(t)=n\}\\&\quad\cdot\sum_{w=i+1}^{\infty} w\operatorname{Prob}\{W_{M-i}(t)=w/N(t)=n\}\\&\le \operatorname{Prob}\{N(t)=n\}E[W^U_{M-i}(t)/N(t)=M]\\&= g_nE[W^U_{M-i}(t)].\end{aligned}\tag{28}$$

The last equality follows from the independence between $N(t)$ and $W(t)$ in the new system. The quantities $E[W^U_{M-i}(t)]$ have been derived in [13]. Quoting the results from there, we have

$$E[W^U_{M-i}(t)] = \sum_{m=M-i}^{M}\mu_m \tag{29}$$

where

$$\mu_M = 1;$$

$$\mu_i = \frac{B(M,i)+\sum_{m=i+1}^{M-1} b(m,i)\mu_m}{1-p^i}; \tag{30}$$

and $B(M,i) = \sum_{m=0}^{i} b(M,m)$.

Note that $E[W^U_{M-i}]$ can be derived from (24) and (25) by letting $\lambda \to 1-p$, in which case $g_n \to 0$, $n < \infty$.

To conclude, we have

$$(i+1)g_n \le \dot{G}^{(n)}_{M-i}(1) \le E[W^U_{M-i}(t)]g_n. \tag{31}$$

From the derivation of the bounds, it is clear that by replacing them in (24), we obtain upper and lower bounds for $E[W_{M-i}]$. From (23), we therefore obtain upper and lower bounds for the expected waiting time at the receiver.

Remark 4.1: We have observed from our extensive computations for practical error probabilities ($p \le 0.1$) that the lower bound in (26) and the upper bound in (29) are indistinguishable up to the eighth decimal point.

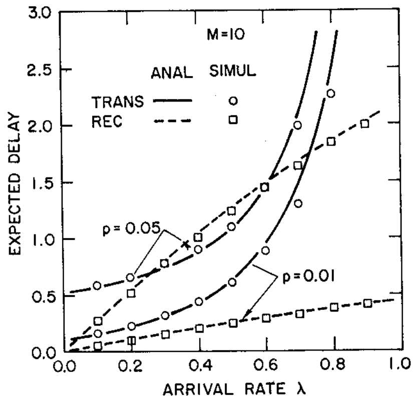

Fig. 1. Expected delays versus arrival rate.

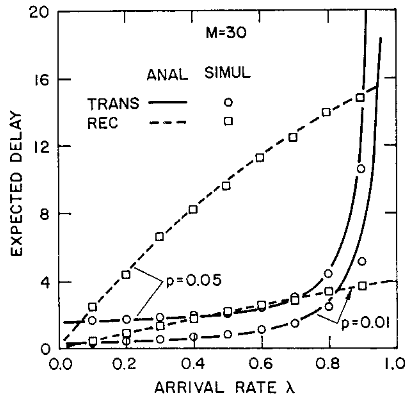

Fig. 2. Expected delays versus arrival rate.

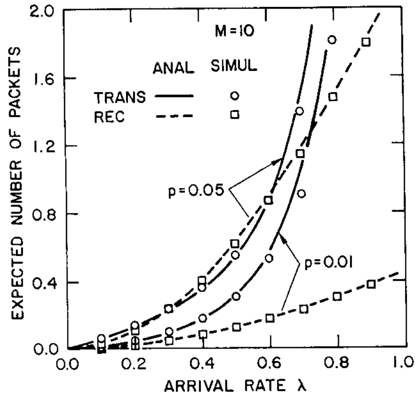

Fig. 3. Expected number of packets versus arrival rate (successfully transmitted packets are not included in the transmitter buffer).

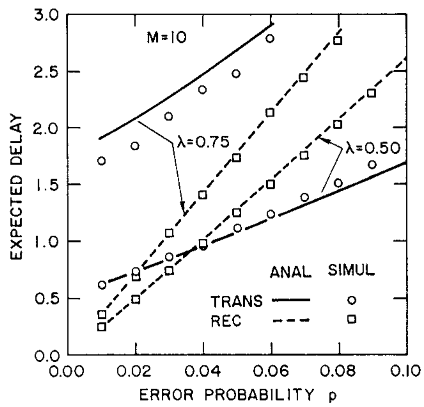

Fig. 4. Expected delays versus error probability.

V. Numerical Examples

In the numerical examples that follow, we assume that the number of arrivals during a single slot is Poissonian distributed (i.e., $F(y) = e^{M\lambda(y-1)}$), and we consider error probabilities that vary from 0.01 to 0.1. Within this range of parameters, the differences between the lower and the upper bounds of the resequencing delay are extremely small (the order of magnitude of these differences is 10^{-8}–10^{-10}).

Our analytical results of the approximate model are compared to simulation results of the exact model of an SR ARQ protocol as defined in [9]. The simulation results (carried out by using the *batch means* method [8, p. 296]) that are depicted in the figures are, with probability of 0.95, at most 2% away from the real values. As observed, the agreement between the analtyical results and the simulation results is very good.

In Figs. 1 and 2, we depict the expected delays at the transmitter [(12) plus the corrective term from (13)] and at the receiver (23) for different windows and error probabilities, as a function of the arrival rate λ. The main conclusion from these figures is that the resequencing delay is not negligible in comparison to the expected delay at the transmitter. When the window size and the error probability are small ($M = 10$, $p = 0.01$), the expected delay at the transmitter is dominant. For larger windows or larger error probabilities, one may observe the following interesting phenomena. For low and high arrival rates, the expected delay at the transmitter is dominant, whereas for a medium range of arrival rates, the expected delay at the receiver is dominant.

The intuitive explanation of this phenomenon is as follows. For low arrival rates, there is mainly at most one packet at the transmitter, and therefore the expected delay there is of order $M(p + p^2 + p^3 + \cdots)$. Furthermore, there is almost no delay at the receiver. As the arrival rate λ approaches $1 - p$, the expected delay at the transmitter explodes to infinity, while the expected delay at the receiver approaches a constant. The latter behavior is due to the fact that the process $(W(t), t \geq 1)$ is stable for every λ (see [13]). Note that the limiting value of the resequencing delay when $\lambda \to 1 - p$ is the upper bound $E[W_1^U]$.

Observe that in satellite systems where the window size represents half of the transmission delay, for a small error probability ($p = 0.01$), the expected delay at the transmitter and the resequencing delay at the receiver are both small in comparison to the transmission delay for a wide range of arrival rates λ. When the transmitter delay becomes comparable to the transmission delay, the resequencing delay is small with respect to both. When the error probability is not too small ($p = 0.05$), the resequencing delay is no longer small in comparison to the transmission delay.

In Fig. 3, we draw the expected number of packets in the transmitter and in the receiver as a function of the arrival rate. Note that for comparison purposes, we depict the expected number of all packets in the transmitter buffer *except* those that were successfully transmitted. The expected number of the latter packets is obviously λM (Little's law).

In Fig. 4, we depict the expected delays at the transmitter and at the receiver for two different arrival rates, as a function of the error probability p. Here, one may observe that the dominant component of the expected delay depends not only on the arrival rate, but also on the error probability. In this figure, we present two different types of behavior. One is where the expected delay at the transmitter is larger than that at the receiver ($\lambda = 0.75$), and another is where the expected delay at the receiver is smaller for low error rates and larger for high error rates ($\lambda = 0.50$).

Finally, in Fig. 5, we depict the expected delays as a function of the window size M. It is interesting to note here that the resequencing delay grows much faster than the expected delay at the transmitter. This holds for other values of error probabilities as well.

Appendix

The procedure for computing $G_{M-i}^{(n)}(1)$, $1 \leq i \leq M - 1$, $M - i \leq n \leq M - 1$ is as follows. One recalls that $\sigma_{M-i,1}(x)$,

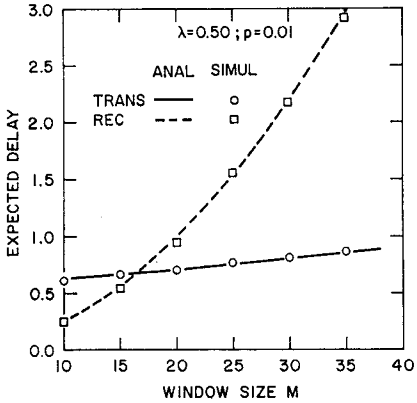

Fig. 5. Expected delays versus window size.

$\sigma_{M-i,2}(x),\cdots,\sigma_{M-i,i}(x)$ are the roots (within the unit disk) of (20), i.e., they satisfy the following equation:

$$\sigma^i_{M-i,j}(x) - F(\sigma_{M-i,j}(x))p^{M-i} \cdot [p\sigma_{M-i,j}(x) + qx]^i = 0, \qquad 1 \le j \le i. \tag{32}$$

For every i $(1 \le i \le M-1)$, one can compute $\sigma_{M-i,1}(1)$, $\sigma_{M-i,2}(1),\cdots,\sigma_{M-i,i}(1)$, and take the derivatives with respect to x in (32) at $x = 1$. The latter are denoted by $\dot\sigma_{M-i,j}(1)$ and can be expressed in terms of $\sigma_{M-i,j}(1)$:

$$\dot\sigma_{M-i,j}(1) = \frac{F(\sigma_{M-i,j}(1))\,i\sigma_{M-i,j}(1)(1-p)}{F(\sigma_{M-i,j}(1))\,i(1-p) - \dot F(\sigma_{M-i,j}(1))\sigma_{M-i,j}(1)[p\sigma_{M-i,j}(1)+q]}, \qquad 1 \le j \le i \tag{33}$$

where $\dot F(\sigma_{M-i,j}(1)) = dF(y)/dy\,|_{\sigma_{M-i,j}(1)}$. By taking the derivative with respect to x in (21) at $x = 1$, we obtain

$$\begin{aligned}
&\sum_{n=M-i}^{M-1} \dot G^{(n)}_{M-i}(1)\Bigg[\sum_{m=M-i}^{n} \sigma^{m+i}_{M-i,j}(1)\,c_{M-i}(M-i,m,n) - p^{M-i}\big(p\sigma_{M-i,j}(1)+q\big)^i\Bigg]\\
&+\sum_{n=M-i}^{M-1} g_n\Bigg\{\sum_{m=M-i}^{n} \sigma^{m+i}_{M-i,j}(1)\\
&\cdot(n-m)\,c_{M-i}(M-i,m,n)\\
&+\sum_{m=M-i}^{n}(m+i)\,\sigma^{m+i-1}_{M-i,j}(1)\\
&\cdot(n-m)\,\dot\sigma_{M-i,j}(1)\,c_{M-i}(M-i,m,n)\\
&-p^{M-i}i\big(p\sigma_{M-i,j}(1)+q\big)^{i-1}\\
&\cdot\big(p\dot\sigma_{M-i,j}(1)+q\big)\sigma^n_{M-i,j}(1)\\
&-p^{M-i}\big(p\sigma_{M-i,j}(1)+q\big)^i n\dot\sigma_{M-i,j}(1)+q\big)\sigma^{n-1}_{M-i,j}(1)\Bigg\}\\
&+\frac{d}{dx}\big[H_{M-i}(x,\sigma_{M-i,j}(x))\big]\,|_{x=1} = 0, \qquad 1 \le j \le i.
\end{aligned} \tag{34}$$

The explicit expression for $d/dx\,[H_{M-i}(x,\sigma_{M-i,j}(x))]\,|_{x=1}$ is somewhat complicated and is given at the end of the Appendix. The crucial point that can also be seen from (19) is that it depends only on $\dot G^{(n)}_l(1)$, $M-i+1 \le l \le M-1$, $l \le n \le M-1$. Therefore, when these quantities are known, (34) forms a set of i equations whose solution yields the i unknowns $\dot G^{(n)}_{M-i}(1)$, $M-i \le n \le M-1$. Consequently, the quantities $\dot G^{(n)}_{M-i}(1)$, $M-i \le n \le M-1$ can be recursively computed for $i = 1,2,\cdots,M-1$.

To simplify the expression for $d/dx\,[H_{M-i}(x,\sigma_{M-i,j}(x))]\,|_{x=1}$, we denote $u = \sigma_{M-i,j}(1)$ and $\dot u = \dot\sigma_{M-i,j}(1)$. From (19), one can compute

$$\begin{aligned}
&\frac{d}{dx}\big[H_{M-i}(x,\sigma_{M-i,j}(x))\big]\,|_{x=1}\\
&= \big[iu^{i-1}\dot u + (i+1)u^i\big]\\
&\cdot\Bigg\{\sum_{n=0}^{M-i}(q+pu)^n g_n\\
&+\sum_{n=M-i+1}^{M-1} g_n \sum_{m=0}^{M-i-1} b(n,m)u^m\\
&+\Bigg[G(u) - \sum_{n=0}^{M-1} g^n u^n\Bigg]\sum_{m=0}^{M-i-1} b(M,m)u^{m-M}\Bigg\}\\
&+u^i\Bigg\{p\dot u\sum_{n=0}^{M-i} n(q+pu)^n g_{n-1}\\
&+\dot u\sum_{n=M-i+1}^{M-1} g_n \sum_{m=0}^{M-i-1} m\,b(n,m)u^{m-1}\\
&+\Bigg[\dot G(u) - \dot u\sum_{n=0}^{M-1} n g^n u^{n-1}\Bigg]\sum_{m=0}^{M-i-1} b(M,m)u^{m-M}\\
&+\Bigg[G(u) - \sum_{n=0}^{M-1} g^n u^n\Bigg]\dot u\\
&\cdot\sum_{m=0}^{M-i-1}(m-M)\,b(M,m)u^{m-M-1}\Bigg\}\\
&+\dot u\sum_{m=M-i}^{M-1}(i+m)u^{i+m-1}\\
&\cdot\sum_{l=M-i+1}^{M-i+M-1-m}\ \sum_{n=m+l-M+i}^{M-1} c_{M-i}(l,m,n)\,G^{(n)}_l(1)\\
&+\sum_{m=M-i}^{M-1} u^{i+m}\sum_{l=M-i+1}^{M-i+M-1-m}\\
&\cdot\sum_{n=m+l-M+i}^{M-1} c_{M-i}(l,m,n)\\
&\cdot\big[(n-m)G^{(n)}_l(1) + \dot G^{(n)}_l(1)\big]\\
&+\sum_{m=M-i}^{M}\big[(m-M+i)\dot u u^{m-M+i-1}\\
&+(M-m)u^{m-M+i}\big]\sum_{l=M-i+1}^{M-i+M-m} c_{M-i}(l,m,M)\\
&\cdot\Bigg[G_l(1,u) - \sum_{n=l}^{M-1} u^n G^{(n)}_l(1) - \sum_{n=0}^{l-1} g_n u^n\Bigg]\\
&+\sum_{m=M-i}^{M} u^{m-M+i}\sum_{l=M-i+1}^{M-i+M-m} c_{M-i}(l,m,M)\\
&\cdot\Bigg[\dot G_l(1,u) - \sum_{n=l}^{M-1}\big[n\dot u u^{n-1}G^{(n)}_l(1) + u^n\dot G^{(n)}_l(1)\big]
\end{aligned}$$

$$
\begin{aligned}
&\left. - (M - l + 1) \sum_{n=0}^{l-1} g_n u^n \sum_{n=0}^{l-1} n g_n \dot{u} u^{n-1} \right] \\
&- \sum_{m=M-i}^{M} \big[(m - M + i) \dot{u} u^{m-M+i-1} \\
&+ (M - m + i) u^{m-M+i} \big] \\
&\cdot c_{M-i}(M - i, m, M) \sum_{n=0}^{M-i-1} g_n y^n \\
&- \sum_{m=M-i}^{M} u^{m-M+i} c_{M-i}(M - i, m, M) \sum_{n=0}^{M-i-1} n g_n \dot{u} u^{n-1} \\
&- \left[M \dot{u} u^{M-1} + (i + 1) u^M \right] p^{M-i} g_{M-i}. \qquad (35)
\end{aligned}
$$

In (35), $G(u)$ and $\dot{G}(u) = dG(\sigma_{M-i,j}(x))/dx\,|_{x=1}$ are computed directly from (7), and $G_l(1, u)$ and $\dot{G}_l(1, u) = dG_l(x, \sigma_{M-i,j}(x))/dx\,|_{x=1}$ are computed from (17) and (18).

References

[1] M. E. Anagnostou and E. N. Protonotarios, "Performance analysis of the selective repeat ARQ protocol," *IEEE Trans. Commun.*, vol. COM-34, pp. 127-135, Feb. 1986.

[2] H. Burton and D. Sullivan, "Error and error control," *Proc. IEEE*, vol. 60, pp. 1293-1301, Nov. 1972.

[3] D. M. Chitre, "A selective-repeat ARQ scheme and its throughput analysis," in *Proc. IEEE ICC*, vol. 3, June 1982, pp. 6G.4.1-6G.4.6.

[4] G. Fayolle, E. Gelenbe, and G. Pujolle, "An analytic evaluation of the performance of the send and wait protocol," *IEEE Trans. Commun.*, vol. COM-26, pp. 313-319, Mar. 1978.

[5] J. F. Hayes, *Modeling and Analysis of Computer Communications Networks.* New York: Plenum, 1984.

[6] T. Housley, *Data Communications and Teleprocessing Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[7] A. K. Kaul, "A model for a packet switching node including packet retransmission effects," *COMSAT Tech. Rev.*, vol. 9, pp. 669-688, 1979.

[8] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis.* New York: McGraw-Hill, 1982.

[9] A. Konheim, "A queueing analysis of two ARQ protocols," *IEEE Trans. Commun.*, vol. COM-28, pp. 1004-1014, July 1980.

[10] M. Miller and S. Lin, "The analysis of some selective repeat ARQ schemes with finite receiver buffer," *IEEE Trans. Commun.*, vol. COM-29, pp. 1307-1315, Sept. 1981.

[11] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach.* Baltimore, MD: John Hopkins Univ. Press, 1981.

[12] B. Reiffen, W. G. Schmidt, and H. L. Yudkin, "The design of an error free data transmission system for telephone circuits," *AIEE Trans. Commun. Electron.*, vol. 79, pp. 224-231, 1961.

[13] Z. Rosberg and N. Shacham, "Resequencing delay and buffer occupancy under the selective-repeat ARQ," *IEEE Trans. Inform. Theory*, vol. 35, pp. 166-173, Jan. 1989.

[14] B. H. Saeki and I. Rubin, "An analysis of a TDMA channel using stop-and-wait, block and select-and-repeat ARQ error control," *IEEE Trans. Commun.*, vol. COM-30, pp. 1162-1173, 1982.

[15] R. D. Stuart, "An insert system for use with feedback communication links," *IEEE Trans. Commun. Syst.*, vol. CS-11, pp. 142-143, 1963.

[16] D. Towsley, "A statistical analysis of ARQ protocols operating in a nonindependent error environment," *IEEE Trans. Commun.*, vol. COM-29, pp. 971-981, July 1981.

[17] —, "The stutter go back-*N* ARQ protocol," *IEEE Trans. Commun.*, vol. COM-27, pp. 869-875, June 1979.

[18] D. Towsley and J. K. Wolf, "On the statistical analysis of queue lengths and waiting times for statistical multiplexers with ARQ retransmission strategies," *IEEE Trans. Commun.*, vol. COM-27, pp. 693-702, Apr. 1979.

[19] E. J. Weldon, "An improved selective-repeat ARQ strategy," *IEEE Trans. Commun.*, vol. COM-30, pp. 480-486, Mar. 1982.

★

Zvi Rosberg received the B.Sc., M.A., and Ph.D. degrees from the Hebrew University, Jerusalem, Israel, in 1971, 1974, and 1978, respectively.

From 1972 to 1978 he was a Senior System Analyst in the General Computer Bureau of the Israeli Government. From 1978 to 1979 he had a Research Fellowship at the Center of Operation Research and Econometrics (CORE), University of Louvain, Belgium. From 1979 to 1980 he was a Visiting Assistant Professor at the University of Illinois. From 1980 to 1989, he was with the Department of Computer Science, Technion, Israel. Since 1990, he has been with IBM Israel, the Science and Technology Center. During 1985-1987 he was on leave at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY. His main research interests include probabilistic models of communication networks and computer systems, performance evaluation, queueing theory, and applied probability.

★

Moshe Sidi (S'77-M'82-SM'87) received the B.Sc., M.Sc., and D.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, in 1975, 1979, and 1982, respectively, all in electrical engineering.

From 1975 to 1981 he was a Teaching Assistant and a Teaching Instructor at the Technion in communication and data networks courses. In 1982 he joined the faculty of the Department of Electrical Engineering, Technion. During the academic year 1983-1984 he was a Postdoctoral Associate at the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge. During 1986-1987 he was a Visiting Scientist at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY. His current research interests are in queueing systems and in the area of computer communication networks.

Dr. Sidi received the New England Academic Award in 1989. He currently serves as Editor for Communication Networks for the IEEE Transactions on Communications.