Delay Analysis of the Stop-and-Wait ARQ Protocol over a Correlated Error Channel

Stijn De Vuyst
Ghent University

Sabine Wittevrongel
Ghent University

Herwig Bruneel
Ghent University

Delay Analysis of the Stop-and-Wait ARQ Protocol over a Correlated Error Channel

S. De Vuyst, S. Wittevrongel and H. Bruneel
SMACS* Research Group, Vakgroep Telecommunicatie en Informatieverwerking,
Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium
Tel. : +32-9-2648902 Fax : +32-9-2644295 e-mail : sdv@telin.ugent.be

Abstract

In this paper, we present the analysis of the Stop-and-Wait ARQ (Automatic Repeat reQuest) protocol with the notable complication that the transmission errors occur in a bursty, correlated manner. Fixed-length packets of data are sent from transmitter to receiver over an error-prone channel. The receiver notifies the transmitter whether a packet was received correctly or not by returning a feedback message over the backward channel. If necessary, the packet is retransmitted until it is received correctly, after which the transmission of another packet starts. For the Stop-and-Wait protocol, no other packets are transmitted while the transmitter waits for a feedback message.

We model the transmitter side as a discrete-time queue with infinite storage capacity and independent and identically distributed (*iid*) packet arrivals. Arriving packets are stored in the queue until they are successfully transmitted over the channel. The probability of an erroneous transmission is modulated by a two-state Markov Chain, rather than assuming stationary channel errors.

Our analysis is based on the use of probability generating functions (pgf) and is exact at the packet level. In previous work we have analysed the queue content distribution of the transmitter, while in the present paper, we give an intuitive derivation of the throughput of the system and the distribution of the packet delay. For the latter, we use the spectral decomposition theorem from linear algebra and give an accurate approximation for the asymptotic behaviour. Finally, we illustrate the importance of accounting for the error correlation in the analysis by means of some numerical examples.

1 Introduction

Whenever packets of data need to be transmitted from point A (the *transmitter*) to point B (the *receiver*), there is always a chance that something bad happens to them while they move through the medium between A and B (the *channel*): some packets may be corrupted or even lost entirely. To cope with this, ARQ (Automatic Repeat reQuest) protocols have been used to provide a more reliable way of communication between the transmitter and the receiver. In the present paper, we study the Stop-and-Wait ARQ protocol (SW-ARQ), and more specifically the packet delay in the transmitter queue. In SW-ARQ, the transmitter sends a packet available in its queue and then simply *waits* until it receives the corresponding feedback message. If the packet was transmitted correctly (ACK), the next packet waiting in the queue is transmitted. Otherwise, if an error occurred (NACK), the packet is retransmitted. Note that the transmitter is inactive during the *feedback delay*, i.e. the time a packet is travelling through the channel, is being processed by the receiver and the feedback message is travelling back. SW-ARQ is simple to implement and ensures that packets are received in the same order as they arrived to the transmitter such that no resequencing is needed.

The model we propose distinguishes itself from previous studies in that we allow the errors occurring in the channel to be *correlated in time*. Instead of assuming that the probability of an erroneous packet is static in time, we propose that this probability depends on what *state* the channel is in when the packet is transmitted. Specifically, the channel alternates between two states which could be termed the GOOD state and the BAD state, both of which reflect different conditions with regard to the error probability. The channel state process is modelled as a two-state Markov Chain with a fixed error probability in either state, resulting in what is also known as the Gilbert-Elliot model [1]. This complication is inspired by the observation that real-life communication channels rarely have the same error sensitivity during their

*SMACS: Stochastic Modeling and Analysis of Communication Systems whole time of operation. Factors such as electromagnetic interference, availability of intermediate network nodes and links, presence of data traffic with higher priority and so on, may all influence the behaviour of the channel and are mostly time-varying in nature. This holds especially when the *wireless* medium is considered where the conditions may change on an even smaller timescale than in a wired medium due to user mobility, interference and channel fading.

The performance of SW-ARQ has been studied before, both in terms of throughput and queueing behaviour, but almost always in case of static error probabilities. Several modifications have been proposed to enhance the performance of SW-ARQ, such as sending multiple copies of a packet during the time the transmitter is waiting for feedback ([2, 3, 4] or combining ARQ with improved error detection techniques (hybrid ARQ). Towsley ([5, 6]) has worked on the SW-ARQ model presented here before, but did not have results for the distribution of the packet delay as we do in Section 5. Also, our analysis is quite different on several accounts and we provide more explicit results in case of a two-state Markov modulated error channel.

The organisation of the paper is as follows. The mathematical model of the SW-ARQ transmitter queue is introduced in Section 2, along with some specific assumptions. In Section 3 we identify a sufficient description for the state of the system at an arbitrary slot and restate the joint distribution in equilibrium of the system state from previous work. In Section 4 we prove a simple expression for the throughput of the system by observing the system at departure slots rather than arbitrary slots. In Section 5 the concept of conditional service period is used to obtain the pgf of the packet delay, hereby referring to Appendix A for the spectral decomposition of the matrix with conditional service time distributions and Appendix B for the asymptotic analysis of the delay distribution. Some numerical examples are discussed in Section 6 and finally, conclusions are drawn in Section 7.

2 Model Description

We model the transmitter of a system operating under the Stop-and-Wait (SW) ARQ protocol as a discrete-time queue. We assume that time is divided in fixed-length intervals called *slots*, whereby one slot is the time required to transmit one packet from the queue into the channel. Let us assume the feedback delay is a fixed number of slots, denoted by s. The operation of SW-ARQ is illustrated in Fig. 1. A Packet P arrives at the system and is queued for some time until all preceding packets are transmitted correctly and the ACK of the previous packet is received. Then P is transmitted for the first time. If the transmission is erroneous, a NACK is returned and P is retransmitted $s+1$ slots after its previous transmission. The packet is retransmitted until finally, an ACK is returned. Then the transmitters knows P was transmitted correctly and there is no need to keep it in the queue any longer.

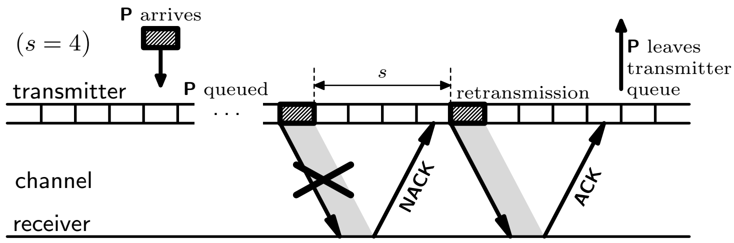

Figure 1: Operation of the transmitter queue under SW-ARQ, meaning of the feedback delay s.

Packets of information enter the system according to a *general independent arrival process*, i.e. the numbers of arrivals during consecutive slots forms a sequence of independent and identically distributed (*iid*) random variables with common mass function $a(n) = \text{Prob}[a = n]$ $(n \geqslant 0)$ and probability generating function (pgf) $A(z) = \sum_{n=0}^{\infty} a(n)z^n$. Furthermore, let a_k be the number of arriving packets during slot k, then we assume that these packets are not stored in the queue until the end of slot k. This way, an arriving packet can only be served (i.e. transmitted) for the first time during the *next* slot $(k+1)$ at the very earliest. When a packet is transmitted, its successful receipt depends on the channel state during the slot in which the feedback message is returned to the transmitter. The transitions between those states, the 0-state (GOOD) and the 1-state (BAD), are governed by a two-state Markov Chain as depicted in Fig. 2: if the feedback of a packet is returned during a slot with channel i, this means the packet was transmitted erroneously with probability e_i, $i = 0, 1$. If we adopt the notation $\bar{q} \triangleq 1-q$, then $\bar{e}_i$ is the probability of a correct transmission. Evidently, the designations GOOD and BAD are only

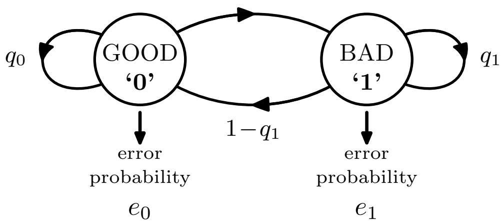

Figure 2: Markovian error model for the transmission channel.

meaningful if $e_0 < e_1$, although this condition is not a requirement for the analysis.

First of all, as a convention for the remainder of this paper, let the index i always be either 0 or 1. As indicated in Fig. 2, the probability of remaining in state i during a slot transition is given by q_i. We denote the channel state (0 or 1) in slot k by r_k and let $\omega_{i,k} \triangleq \text{Prob}[r_k = i]$. Then we have

$$\boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k \mathbf{q} \qquad \text{with } \mathbf{q} \triangleq \begin{bmatrix} q_0 & \bar{q}_0 \\ \bar{q}_1 & q_1 \end{bmatrix} , \tag{1}$$

where $\boldsymbol{\omega}_k$ is the row vector with elements $\omega_{0,k}$ and $\omega_{1,k}$ and $\mathbf{q}$ is the transition probability matrix of the channel state process. Since the channel states are Markovian, both the 0-periods and 1-periods have geometric sojourn times, i.e. Prob[i-period of length n] $= (1-q_i)q_i^{n-1}$ $(n \geqslant 1)$. Rather than using q_0 and q_1, we define the parameters

$$\sigma = \frac{1-q_0}{2-q_0-q_1} \quad \text{and} \quad K = \frac{1}{2-q_0-q_1} ,$$

to be understood as follows. Suppose the channel is in state 0 with probability $\bar{\sigma}$ and in state 1 with probability σ, independently from slot to slot, such that the mean sojourn times are $1/\sigma$ and $1/\bar{\sigma}$ respectively. It is clear that the overall fraction of 1-slots remains equal to σ if the mean lengths of 0- and 1-periods are both multiplied by the same factor K, i.e. if the geometric distributions are chosen such that the mean lengths are $1/(1-q_0) = K/\sigma$ and $1/(1-q_1) = K/\bar{\sigma}$ respectively. Therefore, the factor K can be seen as a measure for the *absolute* lengths of the 0- and 1-periods, while σ characterises their *relative* lengths. Indeed, σ is the relative fraction of 1-slots, since we have from (1) in equilibrium:

$$\boldsymbol{\omega} = \lim_{k\to\infty} \boldsymbol{\omega}_k = \begin{bmatrix} \bar{\sigma} & \sigma \end{bmatrix} . \tag{2}$$

Moreover, the correlation coefficient ϕ between the channel states in two consecutive slots is given by $1-K^{-1}$. Note that $\phi = 0$ and $K = 1$ for uncorrelated errors, whereas for positive correlation we have $0 < \phi < 1$ and $K > 1$. The more correlation present in the channel state process $\{r_k\}$, the higher K is and the fewer the channel changes state.

3 Preliminary Analysis: Joint Distribution of the System State

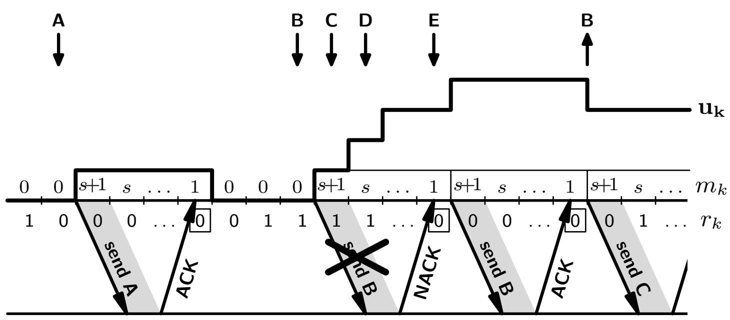

Figure 3: Evolution of the system state given by the channel state r_k, the residual roundtrip time m_k and the queue content u_k.

The analysis of this model with regard to the queue content was presented before in detail in [8]. Since the analysis of the packet delay relies on the equilibrium distribution of the system state, we repeat here the main steps and extend them where necessary. Our study of the transmitter queue described above is done by modelling the system as a (multi-dimensional) Markov Chain and calculating its equilibrium distribution, assuming such equilibrium exists. As the system state variables, we choose the set illustrated in Fig. 3, constructed as follows. Let u_k be the queue content at the beginning of slot k, which obviously needs to be included since $\{u_k\}$ is the process we are interested in. Next, we also need to know how far a packet has progressed through the channel during slot k and when we can expect its feedback message. For this purpose we define the supplementary variable m_k, in a similar way as was done in [9]. The *residual roundtrip time* m_k indicates the remaining number of slots at the beginning of slot k, needed to complete the roundtrip of the most recently transmitted packet if $u_k \geqslant 1$, and $m_k = 0$ if and only if $u_k = 0$. So $m_k = s+1$ when a packet is transmitted and then counts one down in each of the following slots. After s slots, when $m_k = 1$, we know that the feedback message for this packet is being returned and the packet will either leave the queue at the end of the slot (ACK) or be retransmitted in the next slot (NACK). Finally, as before, let the random variable r_k be the channel state during slot k. The channel state comes into play during slots with $m_k = 1$, where it determines the probability that either an ACK or a NACK is returned, or equivalently, that the packet departs from the queue or is to be retransmitted. Let d_k be equal to 1 if a packet departs at the end of slot k and equal to 0 otherwise. Then if $m_k = 1$, the probability of an error, and therefore of '$d_k = 0$', is e_i if $r_k = i$. Hence, the pgf of d_k is given by $d_i(z)=\bar{e}_i z + e_i$ if $r_k=i$. We also define

$$\bar{d}_i(z) \triangleq e_i z+\bar{e}_i\,, \tag{3}$$

which is the pgf of $\bar{d}_k=1-d_k$.

One verifies that the triple (r_k, m_k, u_k) is an adequate Markovian description of the system state at the beginning of slot k. The transitions from slot k to slot $k+1$ in this (three-dimensional) Markov Chain are described by the system equations given in [8] and are used to obtain separate expressions for the distribution of the system state during slots where the queue is *idle* ($m_k=0$) and during slots where the queue is *busy* ($m_k>0$). First, let us define $p_{i,k}$ as the probability that the queue is empty and that the channel is in state i in slot k,

$$p_{i,k} \triangleq \mathrm{Prob}[m_k=0, r_k=i]\,. \tag{4}$$

Secondly, let $yzH_{i,k}(y,z)$ be the joint partial pgf of the residual roundtrip time and the queue content in slot k for a busy queue and channel state i in that slot:

$$H_{i,k}(y,z) \triangleq \mathrm{E}[y^{m_k-1} z^{u_k-1}\{m_k>0, r_k=i\}]\,, \tag{5}$$

where we use the notation $\mathrm{E}[X\{Y\}]=\mathrm{E}[X|Y]\,\mathrm{Prob}[Y]$. Additionally, we define $zR_{i,k}(z)$ as the partial pgf of the queue content in slot k for the case it is the last slot of a roundtrip period ($m_k=1$) and the channel state is i:

$$R_{i,k}(z) \triangleq \mathrm{E}[z^{u_k-1}\{m_k=1, r_k=i\}] = H_{i,k}(0,z)\,. \tag{6}$$

We assume that for $k\to\infty$, the system reaches equilibrium, such that $p_{i,k}$ and the functions $R_{i,k}(z)$ and $H_{i,k}(y,z)$ converge to a limiting value which we indicate by dropping the index k. Let r, m, u and e denote the channel state, the remaining roundtrip time, the queue content and the number of arrivals respectively, in an arbitrary slot during equilibrium. We also define the row vectors $\mathbf{p}$ and $\mathbf{H}(y,z)$ as

$$\mathbf{p} = \begin{bmatrix} p_0 & p_1 \end{bmatrix}, \qquad \mathbf{H}(y,z) = \begin{bmatrix} H_0(y,z) & H_1(y,z) \end{bmatrix}. \tag{7}$$

In [8] we then find the following relations between the probabilities p_0, p_1 and $R_0(0), R_1(0)$:

$$p_0+p_1 = \frac{A(0)}{1-A(0)}\big(\bar{e}_0R_0(0)+\bar{e}_1R_1(0)\big) \quad \text{and} \quad \sigma p_0-\bar{\sigma}p_1 = \frac{\phi A(0)}{1-\phi A(0)}\big(\sigma\bar{e}_0R_0(0)-\bar{\sigma}\bar{e}_1R_1(0)\big) \tag{8}$$

and explicit expressions for $H_i(y,z)$:

$$\begin{aligned}(y-A(z))(y-\phi A(z))H_i(y,z) &= \frac{A(z)}{z}\Big[y\bar{q}_{\bar{\imath}}\big(y^{s+1}\bar{d}_{\bar{\imath}}(z)-z\big)R_{\bar{\imath}}(z)+(yq_i-\phi A(z))\big(y^{s+1}\bar{d}_i(z)-z\big)R_i(z)\Big]\\ &+\frac{y^{s+1}}{(1-\phi)z}\Big[\bar{q}_{\bar{\imath}}(\phi A(z)-y)(1-A(z))(p_0+p_1)+(\phi A(z)-1)(y-A(z))(\bar{q}_{\bar{\imath}}p_{\bar{\imath}}-\bar{q}_ip_i)\Big]\,.\end{aligned} \tag{9}$$

Herein the functions $R_i(z)$ are determined as

$$R_i(z) = \frac{A(z)^s}{N(z)}\Big[\phi^s(\phi A(z)-1)\big(A(z)^{s+1}\bar{d}_{\bar{\imath}}(z)-z\big)(\bar{q}_{\bar{\imath}}p_{\bar{\imath}}-\bar{q}_ip_i)-\bar{q}_{\bar{\imath}}(A(z)-1)\big((\phi A(z))^{s+1}\bar{d}_{\bar{\imath}}(z)-z\big)(p_0+p_1)\Big]\,, \tag{10}$$

with the denominator $N(z)$ being a known function of the system parameters:

$$N(z) \triangleq \sum_{i=0}^{1}\bar{q}_i\big((\phi A(z))^{s+1}\bar{d}_i(z)-z\big)\big(A(z)^{s+1}\bar{d}_{\bar{\imath}}(z)-z\big)\,. \tag{11}$$

The only remaining unknowns are the probabilities p_0 and p_1. To determine those, we need two additional relations. A first relation can be found from the normalisation condition $1=p_0+p_1+H_0(1,1)+H_1(1,1)$, which after taking the limit $y\to1$ and $z\to1$ in (9), turns out to be equivalent to $A'(1)=\bar{e}_0R_0(1)+\bar{e}_1R_1(1)$. By taking the limit $z\to1$ in (10), we explicitly find

$$(1-\phi^{s+1})\big[(s+1)A'(1)+(p_0+p_1-1)(\bar{\sigma}\bar{e}_0+\sigma\bar{e}_1)\big] = \phi^s(s+1)(e_0-e_1)(\bar{q}_0p_0+\bar{q}_1p_1)\,. \tag{12}$$

A second relation for p_0 and p_1 can be obtained from (10) by exploiting the fact that $R_i(z)$ is a (partial) pgf and must therefore be analytic and therefore bounded in the unit disc $|z|<1$. It can be proven with Rouché's theorem that the denominator $N(z)$ has exactly one zero z^* inside the unit disc, i.e. $N(z^*)=0$. This zero can be calculated numerically and a relation between p_0 and p_1 is thus obtained by substituting z^* into the numerator of (10) and let it equal zero. Now, from (9) and (10), one finds the (unconditional) joint pgf of the system state in equilibrium as

$$P(x,y,z) \triangleq \mathrm{E}[x^ry^mz^u] = p_0+xp_1+yzH_0(y,z)+xyzH_1(y,z)\,. \tag{13}$$

4 Calculation of the Throughput

We now focus on the calculation of the throughput of the SW-ARQ system described above. What follows is a new and more intuitive proof for the throughput expression already stated in [6]. Let us call the *service period* of a packet the period it stays in the server of the queue, i.e. from the slot in which it is transmitted for the first time, up to and including the slot in which it departs from the queue. We define the *throughput* η of the system as the maximum number of packets *per slot* that can be correctly delivered to the receiver. Hence, η is a measure for the maximum output rate at which the system can transmit incoming packets and should therefore be compared to $A'(1)$, the mean arrival rate. Indeed, in order to have a stable queue that reaches equilibrium, we must have

$$A'(1) < \eta \,.$$

Obviously, the maximum output rate is only achieved if the system operates under overload conditions, i.e. if we assume there are *always* packets waiting in the queue for transmission. Under these conditions, we can also define the throughput η as the inverse of the mean service time of an arbitrary packet. Therefore, we proceed by deriving the service time distribution of the packets.

Unfortunately, as a consequence of the correlated nature of the channel, the service times are *not independent*. Specifically, a packet's service time distribution will generally be different when its initial transmission happens during a 0-slot than during a 1-slot. Let us say that the service time of a packet *starts* in channel state i if the channel is in state i in the slot *before* the initial transmission of the packet, as indicated in Fig. 4. Conversely, the service *ends* in channel state i' if the channel state is i' in the slot where the packet leaves the queue. Now, let us define $\gamma_{ii'}(n)$ $(n\geqslant 1)$ as the conditional probability that the service of a packet requires n roundtrip periods (or n transmission attempts) and that the channel state is i' at the end of the service *given* that the service time starts in state i $(i, i'=0,1)$, i.e.

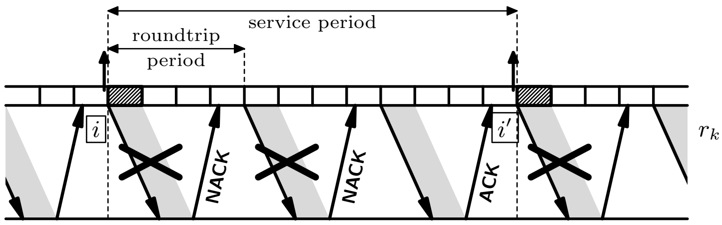

Figure 4: The service time of a packet of $n=3$ roundtrip periods starting in channel state i and ending in state i'.

$$\gamma_{ii'}(n) = \text{Prob}[\,\text{packet needs } n \text{ transmissions}, r_{k+n(s+1)}=i' \,|\, r_{k-1}=i\,]\,, \tag{14}$$

if the packet is first transmitted in slot k.

Since we assume that the service periods are not interrupted by idle periods, every service period starts in the same channel state that the previous one ended with, which clearly is also the channel state during the *departure slot* of the previous packet. The *embedded* channel state process *at departure slots* only is of particular importance to us, since we need to know the equilibrium probabilities of being in either channel state when a service period starts (or ends). Let r_k^* be the channel state during the k-th departure slot and $\pi_{i,k}$ the probability that $r_k^* = i$. Then we have for the row vector $\boldsymbol{\pi}_k$ with elements $\pi_{0,k}$ and $\pi_{1,k}$ respectively,

$$\boldsymbol{\pi}_{k+1} = \boldsymbol{\pi}_k \sum_{n=1}^{\infty} \boldsymbol{\gamma}(n) \qquad \text{with } \boldsymbol{\gamma}(n) = \begin{bmatrix} \gamma_{00}(n) & \gamma_{01}(n) \\ \gamma_{10}(n) & \gamma_{11}(n) \end{bmatrix}\,, \tag{15}$$

which is to be compared with (1) for the channel state probabilities $\boldsymbol{\omega}_k$ at consecutive slots. The row vector of equilibrium probabilities of $\{r_k^*\}$ is $\boldsymbol{\pi} = \lim_{k\to\infty} \boldsymbol{\pi}_k$ which is different from the probabilities $\boldsymbol{\omega}$ in (2), as will be shown.

The probabilities $\boldsymbol{\gamma}(n)$ can be found as follows. According to (1), the $(s+1)$-*step transition probabilities* of the channel state process $\{r_k\}$ are given by the matrix $\mathbf{q}^{s+1}$, such that $[\mathbf{q}^{s+1}]_{ii'}$ is the probability that the channel state is i' at the end of a roundtrip period given that it is i at the end of the previous roundtrip. This matrix has eigenvalues 1 (since it is stochastic) and ϕ. The spectral decomposition representation of $\mathbf{q}^h$ is given by

$$\mathbf{q}^h = \begin{bmatrix} \bar{\sigma} & \sigma \\ \bar{\sigma} & \sigma \end{bmatrix} + \phi^h \begin{bmatrix} \sigma & -\sigma \\ -\bar{\sigma} & \bar{\sigma} \end{bmatrix} \qquad (h\geqslant 0)\,. \tag{16}$$

Now, the matrix $\boldsymbol{\gamma}(n)$ in (15) is found as

$$\boldsymbol{\gamma}(n) = \left(\mathbf{q}^{s+1}\mathbf{e}\right)^{n-1}\mathbf{q}^{s+1}\bar{\mathbf{e}}\,, \qquad n \geqslant 1\,, \tag{17}$$

where the channel error probabilities are arranged in the matrices $\mathbf{e} \triangleq \left[\begin{smallmatrix} e_0 & 0 \\ 0 & e_1 \end{smallmatrix}\right]$ and $\bar{\mathbf{e}} \triangleq \left[\begin{smallmatrix} \bar{e}_0 & 0 \\ 0 & \bar{e}_1 \end{smallmatrix}\right]$. The geometric-like expression (17) can easily be interpreted as follows. If the service of a packet requires n roundtrip periods, then there are first $n-1$ roundtrip periods – each of length $s+1$ – wherein an error occurred followed by one roundtrip without channel error.

The z-transform of the matrix $\boldsymbol{\gamma}(n)$ gives us the probability generating matrix (or *pgm*) $\mathbf{g}(z)$ of the number of roundtrip periods required for the successful transmission of a packet, accounting for the channel state at the start and end of the service. From (17) we find

$$\mathbf{g}(z) = \sum_{n=1}^{\infty} \boldsymbol{\gamma}(n)\, z^n = (\mathbf{I} - z\,\mathbf{q}^{s+1}\mathbf{e})^{-1}\mathbf{q}^{s+1}\bar{\mathbf{e}}z = \mathbf{g}(z) = \frac{z}{\nu(z)}\big(\mathbf{q}^{s+1}\bar{\mathbf{e}} - z\,\phi^{s+1}\begin{bmatrix} \bar{e}_0 e_1 & 0 \\ 0 & e_0\bar{e}_1 \end{bmatrix}\big)\,, \tag{18}$$

where $\mathbf{I}$ is the 2×2 identity matrix and with

$$\nu(z) = \det(\mathbf{I} - z\,\mathbf{q}^{s+1}\mathbf{e}) = 1 - z\big(e_0(\bar{\sigma}+\phi^{s+1}\sigma) + e_1(\sigma+\phi^{s+1}\bar{\sigma})\big) + z^2 e_0 e_1 \phi^{s+1}\,. \tag{19}$$

Note that the inverse matrix in (18) always exists for $|z|\leqslant 1$, because the elements of $\mathbf{g}(z)$ are (partial) pgfs and must be analytic in that region. For future purposes, we also introduce the pgm $\mathbf{S}(z) \triangleq \mathbf{g}(z^{s+1})$ which gives the distributions of the number of *slots* in a service period rather than the number of roundtrips. The nice thing about the matrix representation $\mathbf{S}(z)$ is that it allows us to handle the distribution of contiguous service times *as if they were independent*. Indeed, the conditional distribution of the length of n contiguous services (i.e. without idle periods in between) ending in state i', given that the first service starts in state i is simply given by $[\mathbf{S}(z)^n]_{ii'}$.

From (14) and (18), the transition probability matrix of the embedded process $\{r_k^*\}$ is seen to be given by $\mathbf{g}(1)$ and is easily obtained from (18) as

$$\mathbf{g}(1) = \begin{bmatrix} q_0^* & \bar{q}_0^* \\ \bar{q}_1^* & q_1^* \end{bmatrix} \qquad \text{with } q_i^* = \frac{\bar{e}_i}{\nu(1)}\big(\bar{\sigma} + \phi^{s+1}(\sigma - e_{\bar{i}})\big)\,.$$

Hence, the equilibrium probabilities $\boldsymbol{\pi}$ of the channel state r^* at an arbitrary departure slot must satisfy $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{g}(1)$, from which

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_0 & \pi_1 \end{bmatrix} = \begin{bmatrix} \dfrac{\bar{e}_0\bar{\sigma}}{\bar{e}_0\bar{\sigma} + \bar{e}_1\sigma} & \dfrac{\bar{e}_1\sigma}{\bar{e}_0\bar{\sigma} + \bar{e}_1\sigma} \end{bmatrix}\,. \tag{20}$$

Remember that for the queue working under overload conditions we have defined the throughput η as the inverse of the mean service time. The distributions of the service times conditioned on the channel state in which the service starts are given by $\mathbf{S}(z)$ whereas $\boldsymbol{\pi}$ in (20) are the equilibrium probabilities of being in either state 0 or 1 at the start of a service time. Therefore, using the moment-generating property of pgfs, we find the throughput as

$$\eta^{-1} = \boldsymbol{\pi}\,\mathbf{S}'(1)\,\mathbf{1} = (s+1)\boldsymbol{\pi}\,\mathbf{g}'(1)\,\mathbf{1} \quad \text{with } \mathbf{1} \triangleq \begin{bmatrix} 1 \\ 1 \end{bmatrix}\,, \tag{21}$$

where $\mathbf{1}$ is the 2×1 column vector with 1 on both entries and where $\mathbf{S}'(1)$ indicates the matrix $\mathbf{S}(z)$ with each element differentiated to z and evaluated for $z=1$ (and likewise for $\mathbf{g}'(1)$). After properly evaluating (21), we finally find the following expression for the throughput of the Stop-and-Wait ARQ protocol over the correlated error channel:

$$\eta = \bar{\sigma}\frac{\bar{e}_0}{s+1} + \sigma\frac{\bar{e}_1}{s+1}\,. \tag{22}$$

5 Analysis of the Packet Delay

In this Section, we derive an expression for the pgf $D(z)$ of the total delay d experienced by an arbitrary packet traversing the transmitter queue. Of all the packets arriving to the system, consider an arbitrary packet and tag it as packet $\mathcal{P}$. Also, let us mark the arrival slot of $\mathcal{P}$ as slot I. We define the delay d as the number of slots between the end of the slot in which $\mathcal{P}$ arrives (slot I) and the end of the slot in which $\mathcal{P}$ departs from the transmitter queue. The delay of $\mathcal{P}$ foremost depends on the system state (r_I, m_I, u_I) at the beginning of slot I. Hence, we first need an expression for the system state distribution $P_I(x,y,z)$ in slot I. Now, it has been argued before (see e.g. [7]) that due to the uncorrelated (iid) nature of the packet arrival process, the system as 'seen' by an *arbitrary arriving packet* has the same distribution as the system in an *arbitrary slot*, i.e. $P_I(x,y,z) = P(x,y,z)$ Apart from the system state at the beginning of slot I, the delay of $\mathcal{P}$ also depends on the number and order of arrivals *during* that slot. Let ℓ be the

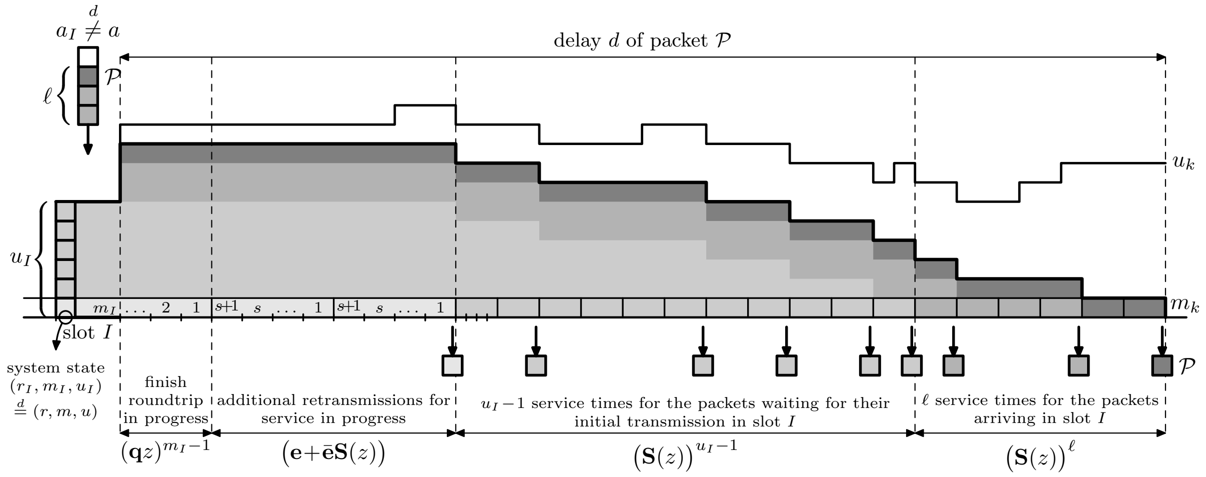

Figure 5: The delay d of the arbitrary packet $\mathcal{P}$ that arrives in slot I with system state (r_I, m_I, u_I).

number of packets arriving in the queue in slot I that will be served no later than (but including) $\mathcal{P}$, as indicated in Fig. 5. The pgf $L(z)$ of ℓ is found as (see [7])

$$L(z) = \frac{z(1 - A(z))}{A'(1)(1 - z)}\,. \tag{23}$$

In the following, we derive the pgf $D(z)$ of the delay of $\mathcal{P}$ by conditioning on the system state at the beginning of slot I. Specifically, as in Section 3, we make the distinction between the cases where the system is *idle* in slot I (i.e. $m_I = 0$) or *busy* ($m_I > 0$). In both cases, we can refer to Fig. 5 for a visual representation of the time periods that constitute the total delay of $\mathcal{P}$.

Let us first consider the case where $\mathcal{P}$ arrives when the queue is idle. This means that the first of the ℓ packets will immediately be transmitted in the next slot. If the service period of the first packet is finished, the next of the ℓ packets is served, without interruption in between, and so forth until finally, the packet $\mathcal{P}$ is served. Therefore, we know from the previous Section that the pgfs of the length of these ℓ contiguous services ending in channel state i' given that they start in state i ($i, i' = 0, 1$) are the entries of the pgm

$$\mathbf{D}_{m_I=0}(z) = \sum_{v=1}^{\infty} \mathbf{S}(z)^v \,\mathrm{Prob}[\ell = v] = L(\mathbf{S}(z))\,. \tag{24}$$

The probabilities that the queue is empty and in either channel state during slot I, are given by the vector $\mathbf{p}$ in (7). Hence,

$$\mathrm{E}[z^d\{m_I = 0\}] \;=\; \mathbf{p}\,\mathbf{D}_{m_I=0}(z)\,\mathbf{1} \;=\; \mathbf{p}\,L(\mathbf{S}(z))\mathbf{1}\,. \tag{25}$$

In Appendix A, we derive the spectral decomposition (41) of the matrix $\mathbf{S}(z)$ with eigenvalue functions $\lambda_1(z)$ and $\lambda_2(z)$. Let us agree that the index j is always used to indicate one of the two eigenvalues (i.e. $j = 1, 2$) and that we may write λ_j when in fact, we mean $\lambda_j(z)$. We find

$$\mathrm{E}[z^d\{m_I = 0\}] \;=\; \sum_{j=1}^{2} L(\lambda_j)\,\mathbf{p}\,\mathbf{S}_j(z)\mathbf{1}\,. \tag{26}$$

Secondly, we consider the more complicated case when $\mathcal{P}$ arrives when the queue is busy serving another packet. Suppose the system state at the beginning of slot I is (r_I, m_I, u_I). From Fig. 5 it is seen that the pgfs of the delay period ending in channel state i' given that it starts in state i (i.e. $r_I = i$) ($i, i' = 0, 1$) are the entries of the pgm

$$\mathbf{D}_{m_I>0}(m_I, u_I, z) = (\mathbf{q}z)^{m_I-1} \cdot \big(\bar{\mathbf{e}} + \mathbf{e}\,\mathbf{S}(z)\big) \cdot \mathbf{S}(z)^{u_I-1} \cdot L(\mathbf{S}(z)) \tag{27}$$

where each factor corresponds to a certain part of the delay. The first factor indicates the number of slots needed to finish the roundtrip period of the packet being served during slot I. If this roundtrip period is finished, either an ACK or NACK was returned to the transmitter. In case of an ACK, no channel error occurred (probabilities $\bar{\mathbf{e}}$) which means the service is finished and the packet departs from the queue. In case of a NACK, a channel error occurred (probabilities $\mathbf{e}$) and the packet is retransmitted such that an additional *remaining* service time must be accounted for. Note that the conditional length of this remaining service time has a distribution that is also given by $\mathbf{S}(z)$. Hence the second factor $\bar{\mathbf{e}} + \mathbf{e}\,\mathbf{S}(z)$. The third factor is due to the service times of the u_I-1 packets waiting in the queue at the beginning of slot I but that were not yet transmitted then. Finally, the fourth factor accounts for the packets arriving during slot I that are transmitted *before* $\mathcal{P}$ and for $\mathcal{P}$ itself, similar to (24). Now, for $1<h\leqslant s+1$ and $n\geqslant 1$ let

$$\boldsymbol{\chi}(h,n) \triangleq \big[\mathrm{Prob}[r_I=0, m_I=h, u_I=n] \quad \mathrm{Prob}[r_I=1, m_I=h, u_I=n]\big]\,, \tag{28}$$

then by using (27) we find for the distribution of the delay d in case the queue is busy during slot I:

$$\begin{aligned}
\mathrm{E}[z^d\{m_I>0\}] &= \sum_{h=1}^{s+1}\sum_{n=1}^{\infty}\boldsymbol{\chi}(h,n)\,\mathbf{D}_{m_I>0}(h,n,z)\mathbf{1} = \sum_{h=1}^{s+1}\sum_{n=1}^{\infty}\boldsymbol{\chi}(h,n)(\mathbf{q}z)^{h-1}\sum_{j=1}^{2}(\bar{\mathbf{e}}+\mathbf{e}\lambda_j)\lambda_j^{n-1}L(\lambda_j)\mathbf{S}_j(z)\mathbf{1}\\
&= \sum_{j=1}^{2}\sum_{h=1}^{s+1}\sum_{n=1}^{\infty}L(\lambda_j)\Big[z^{h-1}\lambda_j^{n-1}\boldsymbol{\chi}(h,n)\begin{bmatrix}\bar{\sigma} & \sigma\\ \bar{\sigma} & \sigma\end{bmatrix}(\bar{\mathbf{e}}+\mathbf{e}\lambda_j)\mathbf{S}_j(z)\mathbf{1}\\
&\qquad + (\phi z)^{h-1}\lambda_j^{n-1}\boldsymbol{\chi}(h,n)\begin{bmatrix}\sigma & -\sigma\\ -\bar{\sigma} & \bar{\sigma}\end{bmatrix}(\bar{\mathbf{e}}+\mathbf{e}\lambda_j)\mathbf{S}_j(z)\mathbf{1}\Big]\,.
\end{aligned} \tag{29}$$

Note that we have used the spectral decomposition (41) again applied to $\mathbf{D}_{m_I>0}(h,n,z)$ in the second line and the representation (16) for $\mathbf{q}^{h-1}$ in the third. From (5) and (7) it is clear that the pgfs of the entries of $\boldsymbol{\chi}(h,n)$ are given by $H_0(y,z)$ and $H_1(y,z)$ respectively, which were determined Section 3. Hence, we find for the delay if the packet $\mathcal{P}$ arrives in a busy slot:

$$\mathrm{E}[z^d\{m_I>0\}] = \sum_{j=1}^{2}L(\lambda_j)\,\mathbf{C}_j(z)(\bar{\mathbf{e}}+\mathbf{e}\lambda_j)\mathbf{S}_j(z)\mathbf{1}\,, \tag{30}$$

where we have used the row vectors $\mathbf{C}_j(z)$ defined as

$$\mathbf{C}_j(z) \triangleq \mathbf{H}(z,\lambda_j)\begin{bmatrix}\bar{\sigma} & \sigma\\ \bar{\sigma} & \sigma\end{bmatrix} + \mathbf{H}(\phi z,\lambda_j)\begin{bmatrix}\sigma & -\sigma\\ -\bar{\sigma} & \bar{\sigma}\end{bmatrix}\,. \tag{31}$$

The entries $C_{0j}(z)$ and $C_{1j}(z)$ follow from (9) by applying the appropriate substitutions for (y,z) and eliminating the functions $R_i(z)$ given by (10):

$$\begin{aligned}
C_{ij}(z) = &\frac{A^{s+1}(\lambda_j)-z^{s+1}}{\bar{\phi}\lambda_j(z-A(\lambda_j))N(\lambda_j)}\Bigg[\big[N(\lambda_j) - A^{s+1}(\lambda_j)\bar{d}_i(\lambda_j)\big(\bar{d}_{\bar{i}}(\lambda_j)\phi^{s+1}\bar{\phi}A^{s+1}(\lambda_j) - \lambda_j(\bar{q}_{\bar{i}}+\phi^{s+1}\bar{q}_i)\big)\big]\\
&\qquad\cdot\big[\bar{q}_{\bar{i}}(1-A(\lambda_j))(p_0+p_1) + \phi^s(1-\phi A(\lambda_j))(q_i p_i - q_{\bar{i}}p_{\bar{i}})\big]\\
&+ \bar{q}_{\bar{i}}A^{s+1}(\lambda_j)\bar{d}_{\bar{i}}(\lambda_j)\lambda_j(1-\phi^{s+1})\big[\bar{q}_i(1-A(\lambda_j))(p_0+p_1) - \phi^s(1-\phi A(\lambda_j))(q_i p_i - q_{\bar{i}}p_{\bar{i}})\big]\Bigg]\,.
\end{aligned} \tag{32}$$

Finally, we can bring together (26) and (30) to obtain the *unconditional* pgf $D(z)$ of the packet delay d:

$$D(z) = \mathrm{E}[z^d] = \mathrm{E}[z^d\{m_I=0\}] + \mathrm{E}[z^d\{m_I>0\}] = \sum_{j=1}^{2}L(\lambda_j)\big[\mathbf{p} + \mathbf{C}_j(z)(\bar{\mathbf{e}}+\mathbf{e}\lambda_j)\big]\mathbf{S}_j(z)\mathbf{1}\,, \tag{33}$$

This can be simplified further by observing that $\bar{e}_i+e_i\lambda_j$ is in fact $\bar{d}_i(\lambda_j)$ and by using the expression (42) for the vector $\mathbf{S}_j(z)\mathbf{1}$. We find

$$D(z) = \sum_{j=1}^{2}\sum_{i=0}^{1}\frac{1}{2}L(\lambda_j)\big[p_i + \bar{d}_i(\lambda_j)C_{ij}(z)\big]\left[1 \pm \frac{(-1)^i\bar{\phi}\,\mu(z^{s+1}) + 2\bar{e}_{\bar{i}}\bar{q}_i(1-\phi^{s+1})}{\bar{\phi}\sqrt{\psi(z^{s+1})}}\right]\,, \tag{34}$$

again with $\pm$ being $+$ if $j=1$ and $-$ if $j=2$. To summarise, in (34) one has to substitute (32) for $C_{ij}(z)$, (38) for the eigenvalue functions $\lambda_j(z)$,(3) for $\bar{d}_i(z)$, (37) for $\mu(z)$, (36) for $\psi(z)$ and (23) for $L(z)$. It is possible to derive the moments of the packet delay from (34) by using the moment-generating property of pgfs. For example, we have verified numerically that the mean packet delay $\mathrm{E}[d]$ found as $D'(1)$ is exactly equal to $U'(1)/A'(1)$, with $U'(1)$ the mean queue content found in [8], as required by Little's theorem. We refer to Appendix B for the tail distribution of the packet delay where an elegant and accurate approximation is presented.

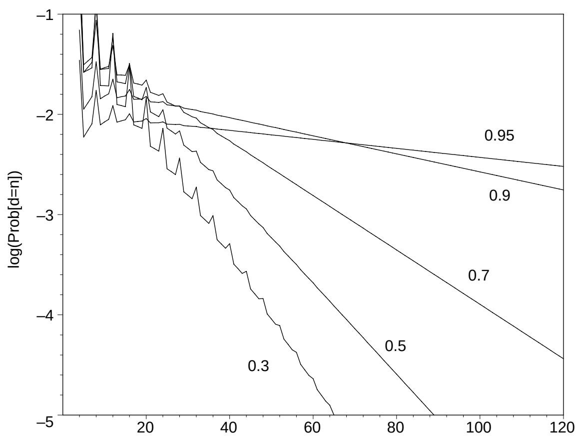

Figure 6: Logarithmic plot of Prob[$d = n$] in case $e_0 = 0.1$, $e_1 = 0.5$, $\sigma = 0.5$ and $K = 10$ for the channel model, feedback delay $s = 3$ and arrivals with geometric distribution and for various values of the load $\rho = 0.3, 0.5, 0.7, 0.9, 0.95$.

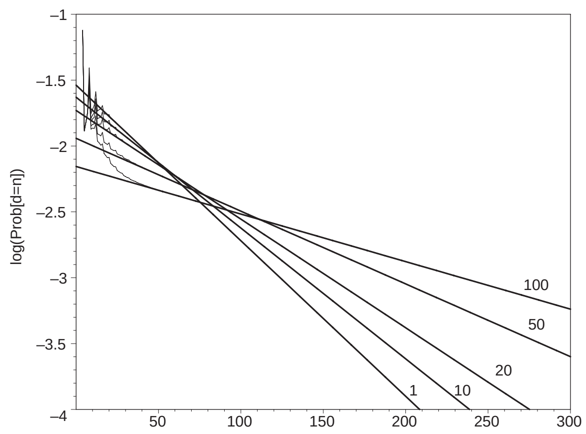

Figure 7: Logarithmic plot of Prob[$d = n$] in case $e_0 = 0.1$, $e_1 = 0.5$, $\sigma = 0.5$ with feedback delay $s = 3$ and Poisson arrivals with load $\rho = 0.9$, for various values of $K = 1, 10, 20, 50, 100$.

6 Numerical Examples

In order to illustrate how the equilibrium distribution of the packet delay d is influenced by the parameters of the model we now consider some practical examples. The Figs. 6 and 7 are plots of the mass function $d(n)$ of the packet delay. These probabilities were obtained by numerical inversion of the pgf $D(z)$ given by (34) using the algorithm presented in [12].

In Fig. 6 we show a logarithmic plot of the mass function $d(n)$ for increasing values of the load $\rho = A'(1)/\eta$. The fraction of slots in either channel state is the same ($\sigma = 0.5$), the correlation factor $K = 10$ and the error probabilities for in the GOOD and BAD state are 0.1 and 0.5 respectively. The roundtrip delay is $s+1 = 4$ slots and the numbers of arrivals per slot have a geometric distribution. As expected, we observe that the probability mass shifts towards higher values as the load increases. If more packets enter the queue in the same time period, they will have to wait longer before they can be transmitted. Secondly, as with all of the following plots, we see that $d(n) = 0$ for $n < s+1$ since the minimal delay of a packet is one roundtrip period of $s+1$ slots. A packet experiences this minimal delay when it is the first packet arriving during an idle slot and its first transmission is successful. Notice also the peaks in the mass function for n equal to multiples of the roundtrip delay $s+1$. The peaks are more pronounced and last into higher values of n when the load is lower, which can be explained as follows. For low load, there is a high probability that packets arrive during a slot when the system is idle. Hence, the first of the packets will be transmitted immediately in the next slot and the delay of all those packets will be a multiple of $s+1$. Note that the probability mass $d(n)$ for n *not* a multiple of $s+1$ is entirely due to arrivals in slots when $m_k = 2, \ldots, s+1$, which is more likely to happen when the load increases.

To illustrate the impact of the *correlated nature* of the transmission errors occurring in the channel, we have plotted in Fig. 7 the delay distribution for increasing values of the correlation factor $K = 1$, 10, 20, 50, 100. The arrivals are Poisson with load 0.9 for all curves and the other parameters are the same as in Fig. 6. Observe that although the number of slots in the BAD state is the *same* for all curves, the delay increases drastically with the factor K. The simple fact that *both* the BAD and the GOOD periods *last longer* for higher K, results in a higher delay for the arbitrary packet. This result strongly shows the importance of accounting for possible correlation in the transmission channel when estimating the packet delay (and the queue content, see [8]). Indeed, the delay may be severely underestimated when assuming only *static* errors ($K = 1$). Additionally, we demonstrate the effectiveness of the dominant pole approximation for the tail distribution of the delay as discussed in Appendix B. The bold lines correspond to the geometric decay of the tail distribution calculated from (44), which gives good results for high n.

7 Conclusion

We have analysed the transmitter queue of the SW-ARQ protocol in case of a two-state Markovian error channel. Closed-form expressions are derived for the pgf of the equilibrium distribution of the system state. By using the eigenvalues and the spectral decomposition of the matrix with the conditional lengths of the service time, we obtained the pgf of the delay experienced by an arbitrary packet. We also gave an accurate approximation for the tail of the delay distribution From the examples, we observe that the delay increases drastically as the correlation in the channel is higher (i.e. if K increases), although the overall fraction of BAD slots remains the same. This result emphasizes the importance of taking into account the correlation of the errors when dimensioning the buffer space or calculating the packet loss.

Acknowledgements

This work has been supported by the Interuniversity Attraction Poles Programme – Belgian Science Policy.

A Spectral Decomposition of $\mathbf{S}(z)$

For the analysis of the delay in Section 5, we need a suitable representation for a function f of the pgm $\mathbf{S}(z)$. The spectral decomposition theorem allows us to determine $f(\mathbf{S}(z))$ by evaluating f for *scalars* rather than matrices and can be found in any textbook on matrix algebra (such as e.g. [10, 11]). First, let us denote the two eigenvalues of $\mathbf{g}(z)$ as $\lambda_1^*(z)$ and $\lambda_2^*(z)$ respectively, which are the solutions to the characteristic equation $\det(\lambda\mathbf{I} - \mathbf{g}(z)) = 0$ equivalent to

$$\lambda^2\nu(z) + \lambda z[z\phi^{s+1}(e_0\bar{e}_1 + \bar{e}_0 e_1) - \bar{e}_0(\bar{\sigma} + \phi^{s+1}\sigma) - \bar{e}_1(\sigma + \phi^{s+1}\bar{\sigma})] + z^2\bar{e}_0\bar{e}_1\phi^{s+1} = 0\,. \tag{35}$$

Let $z\psi(z)$ be the discriminant of this ordinary square root equation in λ, i.e.

$$\psi(z) \triangleq \mu^2(z) + 4\bar{e}_0\bar{e}_1\sigma\bar{\sigma}(1-\phi^{s+1})^2\,. \tag{36}$$

with

$$\mu(z) \triangleq \bar{e}_0(\bar{\sigma}+\phi^{s+1}\sigma) - \bar{e}_1(\sigma+\phi^{s+1}\bar{\sigma}) - z\,(e_1 - e_0)\phi^{s+1}\,. \tag{37}$$

As in Section 5, we agree that the index j is always either 1 or 2. Solving the characteristic equation then yields the eigenvalue functions

$$\lambda_j^*(z) = \frac{z}{2\nu(z)}\Big[\bar{e}_0(\bar{\sigma}+\phi^{s+1}\sigma) + \bar{e}_1(\sigma+\phi^{s+1}\bar{\sigma}) - z\,\phi^{s+1}(\bar{e}_0 e_1 + e_0\bar{e}_1) \pm \sqrt{\psi(z)}\Big]\,, \tag{38}$$

where $\pm$ means $+$ for $j=1$ and $-$ for $j=2$. It can be verified that if $|z|<1$, the eigenvalue functions $\lambda_j^*(z)$ are situated inside the unit disc too. Therefore, in case $f(z)$ is a pgf and thus analytic for $|z|<1$, the spectral decomposition is always well defined. Also, we have that $|\lambda_1^*(z)| \geqslant |\lambda_2^*(z)|$, i.e. $\lambda_1^*(z)$ is the 'larger' of the two eigenvalue functions and yields the PF-eigenvalue $\lambda_1^*(1) = 1$ of the stochastic matrix $\mathbf{g}(1)$ with eigenvector $\boldsymbol{\pi}$ as given by (20). However, although $\lambda_1^*(1) = 1$, the function $\lambda_1^*(z)$ is in general *not* a pgf, since its singularities with smallest modulus are not real, as would be required by the Vivanti theorem. These singularities are the branch points given by the two (complex conjugate) roots of $\psi(z)$. The spectral projectors $\mathbf{G}_j$ of $\mathbf{g}(z)$ are found from as

$$\mathbf{G}_j(z) = \pm\frac{1}{2\sqrt{\psi(z)}} \cdot \begin{bmatrix} \mu(z) \pm \sqrt{\psi(z)} & 2\bar{e}_1\sigma(1-\phi^{s+1}) \\ 2\bar{e}_0\bar{\sigma}(1-\phi^{s+1}) & -\mu(z) \pm \sqrt{\psi(z)} \end{bmatrix}\,, \tag{39}$$

with the same convention for $\pm$ as before. Since $\mathbf{S}(z) = \mathbf{g}(z^{s+1})$, we now also have the spectral decomposition of $\mathbf{S}(z)$, with eigenvalues $\lambda_j(z)$ and spectral projectors $\mathbf{S}_j(z)$ given as

$$\lambda_j(z) = \lambda_j^*(z^{s+1})\,, \qquad \mathbf{S}_j(z) = \mathbf{G}_j(z^{s+1}). \tag{40}$$

Hence, we can write for the pgm $\mathbf{S}(z)$:

$$f(\mathbf{S}(z)) = f(\lambda_1(z))\,\mathbf{S}_1(z) + f(\lambda_2(z))\,\mathbf{S}_2(z). \tag{41}$$

From (39) and (40) one also finds for the projectors $\mathbf{S}_j(z)$ postmultiplied by $\mathbf{1}$:

$$\mathbf{S}_j(z)\mathbf{1} = \frac{1}{2}\mathbf{1} \pm \frac{1}{2\sqrt{\psi(z^{s+1})}}\begin{bmatrix} \mu(z^{s+1}) + 2\bar{e}_1\sigma(1-\phi^{s+1}) \\ -\mu(z^{s+1}) + 2\bar{e}_0\bar{\sigma}(1-\phi^{s+1}) \end{bmatrix}\,. \tag{42}$$

Additionally, it follows directly from (38), (40) and (22) $\lambda_1(z)$ is related to the throughput as $\lambda_1'(1) = \eta^{-1}$. We also prove for $j=1,2$, and the function $N(z)$ given by (11):

$$z - A(\lambda_j(z)) = 0 \;\Rightarrow\; N(\lambda_j(z)) = 0\,. \tag{43}$$

The eigenvalues $\lambda_j(z)$ are functions for which the characteristic polynomial $\det(\lambda_j(z)\mathbf{I} - \mathbf{S}(z))$, given by (35) with all occurrences of z replaced by z^{s+1}, is zero. Now, substitution of $\lambda_j(z)$ into (11) yields an expression for $N(\lambda_j(z))$ which after some careful rearranging of the terms and substitution of $A(\lambda_j(z)) = z$ reduces exactly to the characteristic polynomial of $\mathbf{S}(z)$ in $\lambda_j(z)$. Hence, $N(\lambda_j(z)) = 0$.

B Asymptotic Analysis of the Packet Delay Distribution

We now use the pgf $D(z)$ obtained in (34) to assess the asymptotic behaviour of the delay distribution, i.e. the mass function $d(n)$ when large n. We use a well-known approximation technique (see e.g. [13]) to find $d(n)$ based on the *dominant* singularity of the pgf $D(z)$. In the case at hand, the dominant singularity is a single pole z_d, which must necessarily be real and positive in order to ensure a nonnegative mass function $d(n)$. After close inspection of the factors in (34), we found that the dominant pole z_d is always the smallest zero larger than one of the factor $z - A(\lambda_1(z))$ in the denominator of the functions $C_{i1}(z)$ given by (32). It can be shown that the other poles all have a larger modulus. From (43) we see that z_d is also a zero of the factor $N(\lambda_1(z))$ in the denominator of $C_{i1}(z)$, which would lead us to believe that z_d is a pole of $C_{i1}(z)$ with multiplicity 2. However, the numerator is *also* zero for $z = z_d$ due to the factor $A^{s+1}(\lambda_1(z)) - z^{s+1}$, such that z_d is a a pole with single multiplicity. Now, from the partial fractions expansion of $D(z)$ around z_d, it follows that the (dominant) contribution to $d(n)$ of the contour around z_d can be expressed by the following geometric form:

$$\mathrm{Prob}[d=n] \cong -\theta\, z_d^{-n-1}\,, \tag{44}$$

where θ is the residue of $D(z)$ in the point $z = z_d$, i.e. $\theta = \mathrm{Res}_{z_d} D(z) = \lim_{z\to z_d}(z-z_d)D(z)$. For convenience, let us define $y_d \triangleq \lambda_1(z_d)$ and $y'_d \triangleq \lambda'_1(z_d)$, then we have from (34):

$$\theta = \sum_{i=0}^{1} \frac{1}{2} L(y_d)\bar{d}_i(y_d)\mathrm{Res}_{z_d} C_{i1}(z)\left[1 + \frac{(-1)^i \bar{\phi}\,\mu(z_d^{s+1}) + 2\bar{e}_{\bar{i}}\bar{q}_i(1-\phi^{s+1})}{\bar{\phi}\sqrt{\psi(z_d^{s+1})}}\right], \tag{45}$$

where the residues of $C_{i1}(z)$ in z_d follow from (32) as

$$\begin{aligned}\mathrm{Res}_{z_d} C_{i1}(z) = \frac{(s+1)z_d^{2s+1}}{\bar{\phi}N'(y_d)y'_d y_d}\Big[&\bar{q}_{\bar{i}}\bar{d}_{\bar{i}}(y_d)y_d(1-\phi^{s+1})\big[-\bar{q}_i(1-z_d)(p_0+p_1) + \phi^s(1-\phi z_d)(q_i p_i - q_{\bar{i}}p_{\bar{i}})\big]\\ &+ \bar{d}_i(y_d)\big(\bar{d}_{\bar{i}}(y_d)\phi^{s+1}\bar{\phi}z_d^{s+1} - (\bar{q}_{\bar{i}} + \phi^{s+1}\bar{q}_i)y_d\big)\big[\bar{q}_{\bar{i}}(1-z_d)(p_0+p_1) + \phi^s(1-\phi z_d)(q_i p_i - q_{\bar{i}}p_{\bar{i}})\big]\Big].\end{aligned} \tag{46}$$

Note that the second eigenvalue function $\lambda_2(z)$ is of no importance to the tail distribution of the packet delay.

References

[1] E.N. Gilbert, *Capacity of a Burst-noise Channel*, *Bell Systems Tech. Journal*, Vol. 39, No. 9, 1960, pp. 1253–1265.

[2] M. Moeneclaey, H. Bruneel, I. Bruyland and D.Y. Chung, *Throughput Optimization for a generalized Stop-and-Wait ARQ Scheme*, *IEEE Transactions on Communications*, Vol. 34, No. 2, February 1986, pp. 205–207.

[3] R. Fantacci, *Performance Evaluation of Some Efficient Stop-and-Wait Techniques*, *IEEE Transactions on Communications*, Vol. 40, No. 11, November 1992, pp. 1665–1669.

[4] M. De Munnynck, A. Lootens, S. Wittevrongel, H. Bruneel *Transmitter Buffer Behaviour of Stop-and-Wait ARQ Schemes with repeated transmissions*, *Electronics Letters*, Vol. 38, No. 21, October 2002.

[5] D. Towsley, J.K. Wolf, "*On the Statistical Analysis of Queue Lengths and Waiting Times for Statistical Multiplexers with ARQ Retransmission Schemes*", *IEEE Transactions on Communications*, Vol. 27, No. 4, April 1979, pp. 693-702.

[6] D. Towsley, "*A statistical analysis of ARQ protocols operating in a nonindependent error environment*", *IEEE Transactions on Communications*, Vol. 27, No. 7, July 1981, pp. 971–981.

[7] H. Bruneel and B.G. Kim, *Discrete-time models for communication systems including ATM* (Kluwer Academic Publishers, Boston, 1993).

[8] K. Tworus, S. De Vuyst, S. Wittevrongel, H.Bruneel, "*Transmitter Buffer Behavior of the Stop-and-Wait ARQ Scheme Under Correlated Errors*", *Proceedings of the Conference on Design, Analysis and Simulation of Distributed Systems*, DASD 2004, (April 18–22 2004, Washington D.C.).

[9] H. Bruneel, *Performance of Discrete-time Queueing Systems*, *Computers & Operations Research*, Vol. 20, No. 3, 1993, pp. 303–320.

[10] F.R. Gantmacher, *The Theory of Matrices, Volume One* (AMS Chelsea Publishing, Providence, Rhode Island, 1959).

[11] C.D. Meyer, *Matrix Analysis and Applied Linear Algebra*, (SIAM, Philadelphia, 2000).

[12] J. Abate, W. Whitt, *Numerical Inversion of Probability Generating Functions*, *Operations Research Letters*, Vol. 12, No. 4, 1992, pp. 245–251.

[13] , H. Bruneel, B. Steyaert, E. Desmet, G. Petit, *Analytic Derivation of Tail Probabilities for Queue Lengths and Waiting Times in ATM Multiserver Queues*, *European Journal of Operational Research*, Vol. 76, 1994, pp. 563-572.