Opportunistic File Transfer over a Fading Channel: A POMDP Search Theory Formulation with Optimal Threshold Policies

Leigh A. Johnston and Vikram Krishnamurthy, *Fellow, IEEE*

***Abstract*— We present a computationally efficient algorithm that minimizes the transmission energy and latency associated with transmitting a file across a Gilbert Elliott fading channel. We formulate the optimal tradeoff between transmission energy and latency as a partially observed Markov decision process problem (POMDP). The channel state is not directly observed and hence transmission decisions must be based on ACK/NAK information provided over a feedback channel. The key idea is to reformulate the resulting POMDP as a Markovian search problem, with optimal transmission control policies that are threshold in nature. Threshold policies are computationally inexpensive to implement. Our analysis shows that for different parameter values of the Gilbert Elliott fading channel, the optimal transmission policy, while threshold in structure, exhibits vastly different behaviour – from persistent retransmission to back-off and wait. Numerical examples demonstrate the performance improvements that can be obtained using the optimal threshold policies as compared to existing heuristic algorithms.**

***Index Terms*— Partially observed Markov decision process, threshold policy, Gilbert Elliott channel, ARQ protocol, optimal search theory.**

I. Introduction

ENERGY conservation and latency minimization are important issues in personal communication services such as mobile telephones and notebook computers, and also in wireless sensor networks. ARQ (automatic repeat request) protocols form the basis of transmission protocols for wireless networks due to their inherent simplicity and flexibility. Due to multipath fading and shadowing in wireless channels, transmission errors are statistically correlated (i.e., memory is introduced). This memory can be exploited to devise adaptive ARQ schemes that adjust the operation of mobile users based on the predicted state of the channel. Such adaptive ARQ transmission schemes are useful for delay tolerant traffic applications such as email, web-browsing and remote login. Indeed, protocols at the data link and MAC layers have flexibility to defer transmissions and voluntarily release the channel if the channel conditions deteriorate. Since each failed transmission is a waste of power consumption and limited network resources, there is strong motivation for devising novel transmission control algorithms for retransmission or suspension of transmission to enhance channel utilization as well as conserve battery power for the mobile users. Such transmission control algorithms that determine the optimal tradeoff between latency and transmission energy are also of interest in sensor networks where maximizing battery life is of key importance. Several heuristic ARQ protocols such as Stop-and-Wait, Back-Off-N and Selective-Repeat [1] have been developed.

In this paper we present computationally efficient packet transmission control algorithms that minimize a weighted sum of the transmission energy and latency for file transfer over an uplink fading channel. The evolution of the correlated fading channel is modelled as a two state Gilbert Elliott Markovian channel. We assume that a Stop-and-Wait ARQ protocol is operating between the transmitter (user) and receiver (base station). Thus a feedback channel relays positive acknowledgements (AKs) of successful transmissions and negative acknowledgements (NAKs) for failed transmissions to the transmitter. These AKs and NAKs are a probabilistic function of the underlying unobserved state of the Markovian channel. That is, the sequence of AKs and NAKs fed back to the transmitter form the realization of a Hidden Markov Model and provides indirect information about the channel state.

The problem of computing the optimal adaptive file transfer strategy for a Stop-and-Wait ARQ protocol over an uplink Gilbert Elliott fading channel can then be formulated as an instance of a stochastic control problem called a *partially observed Markov decision process* (POMDP). The solution of this POMDP yields the optimal tradeoff between transmission energy and latency (transmission delay). Although POMDPs have recently received much attention in sensor scheduling for network-centric warfare [2] and artificial intelligence research [3], in general they are PSPACE hard to solve, i.e., require exponential computational complexity and memory. The main contribution of this paper is to show that optimal file transfer strategies are given by threshold policies. Such threshold policies are computationally inexpensive to implement.

Main Results: The main ideas in this paper are as follows: (i) We present a novel formulation of the opportunistic file transfer problem using a Stop-and-Wait ARQ transmission protocol over a Gilbert Elliott fading channel as an *optimal*

Manuscript received December 1, 2002; revised November 5, 2004; accepted December 22, 2004. The associate editor coordinating the review of this paper and approving it for publication was Z. Haas. This work was supported by a NSERC Strategic Research Grant. Some of the results in this paper appear in the IEEE Communications Conference (ICC) 2003.

L. Johnston is with the Department of Electrical and Electronic Engineering, University of Melbourne, Victoria, 3010 Australia (e-mail: leighj@ee.mu.oz.au).

V. Krishnamurthy is with the Department of Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, V6T 1Z4, Canada (email: vikramk@ece.ubc.ca).

Digital Object Identifier 10.1109/TWC.2006.02019

search problem of a Markovian target. Optimal search theory for a Markovian target has a long history – we refer the reader to [4] and references therein for an excellent exposition. We show in Sec. III that the existence of a threshold policy for the optimal file transfer policy is a particular instance of a well known conjecture made by Ross [5] regarding the optimal search of a Markovian target. Until recently, Ross' conjecture has remained unproven. In 1995, MacPhee and Jordan proved Ross' conjecture for a large class of search problems [6]. By interpreting the results in [6], we show that optimal transmission policy for file transfer does indeed satisfy threshold policies.

(ii) We then interpret these optimal threshold policies by demonstrating that depending on channel conditions there are two types of optimal strategies: (a) Back-Off-N scheme, i.e., after a failed transmission the controller waits N intervals (where N is a variable interval whose value depends on the random evolution of the state of the channel and specified cost parameters) before attempting retransmission, or (b) alternate in a sequence between retransmission and suspension. The un-discounted cost formulation we consider in this paper allows us to consider more general problems such as determining the optimal strategy to minimize the expected time until a successful transmission (delay cost function). Also, under certain ergodicity assumptions un-discounted search problems are stochastic shortest path problems [7] that terminate in finite time with probability one.

Context: The Gilbert Elliott model used in this paper describes the evolution of a correlated fading channel as a two state Markov chain. Such models have been widely used to abstract the physical layer channel in the context of link and network layer algorithm design [8]–[10]. For example, [8]–[10] analyze the performance of various ARQ schemes in the presence of Rayleigh fading using a 2 state Gilbert Elliott model. In [9, pp.286] it is shown that for low correlation fading, persistent ARQ is most efficient in terms of transmission energy (since the channel is almost independent from slot to slot) and for high correlation fading a probing protocol is more efficient. We show in this paper that the threshold based optimal transmission controller exhibits similar exhibit similar properties.

At the physical layer, techniques for estimating/predicting wireless channel states include pilot-symbol aided channel estimation (see [11], [12], and references therein), blind estimation [13], and techniques involving periodically probing the channel to detect any changes in the channel state information [9]. However, traditional ARQ protocols such as Stop-and-Wait ARQ (used in IEEE802:11 standards, IEEE802:11a and 802:11b), Go-Back-N (GBN) ARQ and Selective- Repeat (SR) ARQ recover from errors by retransmitting the erroneously transmitted packets regardless of the channel state (channel unaware). If the channel happens to be in the $\texttt{bad}$ state, then such persistent retransmission schemes may lead to wasted energy consumption due to unnecessary packet losses and therefore additional retransmissions [9]. In this paper we use the probabilistic information contained in the ACKs and NAKs about the channel state to determine the transmission policy that optimizes an infinite horizon cost function comprising of a weighted sum of the latency and energy. In a more sophisticated setting, one could also consider a cross layer approach where the finite state Markov model of the channel dynamics at the link layer is combined with channel state information from a channel estimator at the physical layer – this is the subject of future work.

Related work: Several papers consider dynamic programming methods for solving wireless resource allocation problems with complete channel information. An excellent reference on opportunistic file transfer over fading channels is the recent preprint [14] – the paper also considers average delay constraints. Recently, Koole *et al* develop a POMDP formulation for the optimal transmission policies of multiple users sharing a single wireless link in [15]. They show that the policies are myopic (threshold) in nature. The problem formulation in [15] is different from ours – they assume the transmitter has complete knowledge of the channel state via the receipt of positive or negative acknowledgements. In [16], the problem of determining the optimal transmission policy over a fading channel is studied in the context of a Back-Off-N scheme. In Sec. VI, we compare the throughput, efficiency and cost of our optimal threshold policy with a Back-Off-N method.

Limitations of our analysis: A limitation of the threshold policy result presented in this paper is that it can be proved optimal only for a two-state Gilbert Elliott channel with fixed transmission power and with ACK/NAK feedback (i.e., simple ARQ protocol). For channels with a larger state space, or multiple power levels, or where further information about the channel is available (beyond ACKs and NAKs), the optimal policy is not necessarily threshold. For these cases the optimal policy for the equivalent search problem can be computed via stochastic dynamic programming but the computational cost is often prohibitive. In Sec.V, for completeness we outline how the general search problem can be formulated as a POMDP and solved using available software. However, we believe the threshold policy for the two-state Gilbert Elliott model is of interest as it allows us to gain deeper insight into how the channel parameters affect the nature of the re-transmission policies.

II. System Model

The time axis is divided into slots of equal duration which correspond to the round-trip time required for transmitting a single fixed length packet. Propagation and processing delays are assumed to be negligible.

Fading Channel: The uplink channel for transmission of packets is assumed to evolve as a two-state Gilbert Elliott model [17]. Under this model, the state of the channel evolves according to a discrete-time, two-state first order Markov chain $\{s_k\}$, where $k = 1, 2, \ldots$ denotes the transmission slot. The state of the channel at transmission slot k is $s_k \in \left\{\texttt{good}, \texttt{bad}\right\}$. Denote the transition matrix governing the channel model as

$$R = \begin{bmatrix} g & 1-g \\ 1-b & b \end{bmatrix} \quad (1)$$

Here g denotes the probability that the channel remains in the $\texttt{good}$ state between discrete time instants, etc. We assume $\{s_k\}$ is aperiodic and irreducible, i.e., $0 < g, b < 1$. The

Markovian channel memory is defined by $\mu = g + b - 1$. We assume throughout this paper that μ is non-negative, i.e., $0 \leq \mu < 1$. Boundary case $\mu = 0$ denotes a memoryless channel, while $\mu = 1$ denotes a channel with infinite memory. We refer the reader to [9, Eq(3)-(6)] for relating the transition probabilities g and b to physical layer fading channel parameters such as the fading margin and Doppler frequency.

The presence of memory allows the user to predict the state of the channel via a Hidden Markov Model predictor, see (8) below. Note that the optimal search formulation and threshold policies in [6] holds for negative μ, i.e., $-1 < \mu < 0$ as well – however negative μ is without practical meaning in Gilbert Elliott channels.

Transmission Controller and Feedback Information: At each transmission slot k, the transmission controller chooses an action $a_k \in \{\text{Tr} \stackrel{\triangle}{=} \text{Retransmit}, \text{Su} \stackrel{\triangle}{=} \text{Suspend}\}$. At the end of each transmission slot, the user receives error-free feedback from the receiver specifying whether a packet is successfully received. Let $y_k \in \{\text{ACK}, \text{NAK}\}$ denote the observation (feedback). $y_k = \text{ACK}$ denotes a successful transmission. $y_k = \text{NAK}$ occurs due to a failed transmission (when a packet is lost or corrupted) or when the controller chooses action $a_k = \text{Su}$ (conformation that no packet was sent). The observation probabilities are modelled as:

$$\begin{aligned} P(y_k = \text{ACK}|s_k = \texttt{good}, a_k = \text{Tr}) &= 1 - P_e, \\ P(y_k = \text{ACK}|s_k = \texttt{bad}, a_k = \text{Tr}) &= 0, \\ P(y_k = \text{ACK}|s_k = \texttt{good}, a_k = \text{Su}) &= 0, \\ P(y_k = \text{ACK}|s_k = \texttt{bad}, a_k = \text{Su}) &= 0 \end{aligned} \tag{2}$$

and $P(y_k = \text{NAK}|s_k, a_k) = 1 - P(y_k = \text{ACK}|s_k, a_k)$. In words (2) states: The action $a_k = \text{Tr}$ (transmit packet) while the channel in the $s_k = \texttt{bad}$ state results in either the loss or corruption beyond redemption of that packet, both of which result in a $y_k = \text{NAK}$ being returned. Conversely, $a_k = \text{Tr}$ over an $s_k = \texttt{good}$ channel results in $y_k = \text{NAK}$ with probability P_e. Thus a positive packet acknowledgement (ACK) occurs with probability $1 - P_e$ when the transmission is made over the channel in the good state. Notice that the state of the channel is not directly observed – thus the state of the channel is a Hidden Markov Model.

The value of P_e is determined by the system parameters including packet length and codeword length for a pure ARQ system, in which case P_e is a reflection of the error detection capability. For hybrid-ARQ, P_e also incorporates the error correction capability of the system (see [1, Sec.15.3] for explicit expressions of error probabilities in type-I hybrid-ARQ systems). Finally, a noisy feedback channel is easily incorporated in the above formulation via adaptation of the observation probabilities.

There are costs associated with the transmission control decisions. Upon receipt of a NAK, the control algorithm decides whether to retransmit the packet, or wait for a more opportune moment to attempt retransmission. A latency cost of c_0 is associated with suspension of a packet, while the energy required for retransmission incurs a cost of c_1 (see (5)).

III. Formulation as a Search Problem

The aim here is to formulate the optimal transmission control problem as a two-state Markovian search problem described in [5], [6]. The analogy between the transmission control problem and the optimal search problem is this: In this search problem a target (channel s_k) moves between two boxes (states) according to a Markov chain. At each time instant the searcher is to choose one of two possible actions: search a box (Tr) or wait (Su). The objective is to determine the optimal sequence of actions to be taken to find the moving target with minimum cost. Each search or suspension incurs a cost and an observation is received – target found (ACK) or target not found (NAK). The problem terminates when the target has been found (ACK received). We now formulate this Markovian search problem more precisely.

A. Dynamics and Cost Function

As long as the transmission controller receives ACKs, it is easily seen that the optimal policy is to continue transmission of packets, until a NAK is encountered. As soon as a NAK is received, indicating a transmission failure, the search problem commences ($k = 1$) and the transmission control algorithm takes over to implement the optimal policy of retransmissions and suspensions until an ACK is received, after which the control algorithm terminates. (A complete summary of the transmission control algorithm is given below in Sec.IV-B). Note that time indexing k restarts at the beginning of each control problem, i.e., upon receipt of the first NAK after a sequence of ACK symbols.

To allow for termination of the search problem when an ACK is received, it is necessary in our mathematical formulation to replace s_k with the observation dependent Markov chain x_k, $k = 1, 2, \ldots$, on the augmented state space $X = \{\texttt{good}, \texttt{bad}, T\}$. The state T is a fictitious "terminal" state that is added in order to terminate the search upon the reception of an ACK. The state of the target $x_k \in X$ evolves according to the observation dependent transition probability matrices $R^y(i,j) = P(x_{k+1} = X_j | x_k = X_i, y_k = y)$

$$R^{\text{ACK}} = \begin{bmatrix} \mathbf{0}_{3\times 2} & \mathbf{1}_{3\times 1} \end{bmatrix}, \quad R^{\text{NAK}} = \begin{bmatrix} R & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{1\times 2} & 1 \end{bmatrix}. \tag{3}$$

Here R defined in (1) is the transition probability matrix of the Gilbert Elliott channel.

Let $\Psi_1 \stackrel{\triangle}{=} \{y_0, a_0\}$. Define the initial distribution of x_k as

$$\begin{aligned} \pi_1(i) &= P(x_1 = X_i | \Psi_1), \quad i = 1, 2, \\ \pi_1(3) &= P(x_1 = T | \Psi_1) = 0 \end{aligned} \tag{4}$$

The observation probabilities $P(y_k = y | x_k = X_i, a_k = a)$ are defined identically to (2) for $i = 1, 2$ and $P(y_k = \text{ACK} | x_k = T, a_k = a) = 1$ regardless of the action, $a \in \{\text{Su}, \text{Tr}\}$, taken. In search theory, $P(y_k = \text{NAK} | x_k = \texttt{good}, a_k = \text{Tr}) = P_e$ is called the "overlook probability" [5], [6].

The search costs are the latency cost for suspension and energy cost for retransmission, respectively

$$\begin{aligned} c_0 &\stackrel{\triangle}{=} c(x_k = X_i, a_k = \text{Su}) > 0, \\ c_1 &\stackrel{\triangle}{=} c(x_k = X_i, a_k = \text{Tr}) \geq 0. \end{aligned} \tag{5}$$

The terminal state incurs no cost, i.e. $c(x_k = T, a_k = a) = 0$. The assumption c_0 strictly greater than zero is required to make the transmission control problem meaningful – otherwise if $c_0 = 0$, it would be optimal to suspend for all time.

Define the observation and action histories: $Y_k = [y_1, \dots, y_k]$, $A_k = [a_1, \dots, a_k]$. At slot $k = 2, 3, \dots$, the information vector available to the transmission controller is $\Psi_k = \{\pi_1, A_{k-1}, Y_{k-1}, \text{ACK} \notin Y_{k-1}\}$.

Let $\mathcal{D}$ denote the set of admissible *stationary* search policies. A stationary search policy is a measurable function that maps the information vector Ψ_k to action $a_k \in \{\text{Tr}, \text{Su}\}$, i.e., $\mathbf{d} \in \mathcal{D}$ is such that $a_k = \mathbf{d}(\Psi_k)$.

Let $V(\pi_1; \mathbf{d})$ denote the expected cost to find the target (receive an ACK), given the initial state distribution π_1. That is,

$$V(\pi_1; \mathbf{d}) = \lim_{N \to \infty} \mathbf{E}_{\mathbf{d}} \left\{ \sum_{k=1}^{N} c(x_k, a_k) \;\middle|\; \pi_1 \right\} \tag{6}$$

where $\mathbf{E}_{\mathbf{d}}$ denotes expected value parametrized by the stationary policy $\mathbf{d}$. The aim is to determine the optimal stationary policy $\mathbf{d}^* \in \mathcal{D}$ and associated minimum expected cost $V(\pi_1)$ defined as

$$V(\pi_1, \mathbf{d}^*) = V(\pi_1) \triangleq \inf_{\mathbf{d} \in \mathcal{D}} V(\pi_1, \mathbf{d}) \tag{7}$$

Remarks: (i) The above problem is an instance of a partially observed stochastic shortest path problem. We refer the reader to [7] where the finiteness of this sum is proved for aperiodic irreducible R.
(ii) The costs c_0 and c_1 should be chosen to reflect the relative importance of latency and transmission energy, respectively. It is interesting to note that choosing the latency cost equal to the transmission energy cost, i.e., $c_0 = c_1$ in (7), minimizes the expected search delay to find the object (receive an ACK). This search delay cost was used in the classical search theory paper [18].

B. Information State Formulation

To compute the optimal transmission (search) policy in (7) via stochastic dynamic programming, it is necessary to express the above search problem in terms of the *information state*, see [19]. The conditional probability distribution of the channel state, given the available information, is denoted by the information state [19], $\pi_k(x) = P(x_k = x | \Psi_k)$, $x \in X$. Let $\pi_k = \begin{bmatrix} \pi_k(1) & \pi_k(2) & \pi_k(3) \end{bmatrix}'$. The information state vector is computed recursively via Bayes' rule according to the following Hidden Markov Model predictor [19]:

$$\pi_{k+1} = \frac{R^{y_k\prime} Q^{a_k}(y_k) \pi_k}{\mathbf{1}' R^{y_k\prime} Q^{a_k}(y_k) \pi_k} \tag{8}$$

where $Q^a(y) = \text{diag}[\Pr(y_k = y | x_k = \texttt{good}, a_k = a), \Pr(y_k = y | x_k = \texttt{bad}, a_k = a), \Pr(y_k = y | x_k = T, a_k = a)]$. This HMM predictor predicts the state of the channel at time $k+1$ given that a NAK was received at time k. Because of the structure of R^y and $Q^a(y)$ it is easily verified from (8) that π_{k+1} has the following structure:

$$\pi_{k+1} = \begin{cases} \begin{bmatrix} p_{k+1} & q_{k+1} & 0 \end{bmatrix}' & \text{if } y_k = \text{NAK} \\ \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}' & \text{if } y_k = \text{ACK} \end{cases} \tag{9}$$

where $p_{k+1} + q_{k+1} = 1$. Moreover, it is easily shown from (8) that if $y_k = \text{NAK}$ then the HMM predictor yields

$$p_{k+1} = \begin{cases} L_{\text{Tr}}(p_k) \triangleq \frac{P_e \, g \, p_k + (1-b) q_k}{P_e p_k + q_k} & \text{if } a_k = \text{Tr} \\ L_{\text{Su}}(p_k) \triangleq \mu p_k + 1 - b & \text{if } a_k = \text{Su} \end{cases} \tag{10}$$

Note that from (4) and (10) it follows that the initial state distribution is given by

$$p_1 = L_{\text{Tr}}(g) \;\Rightarrow\; \pi_1 = \begin{bmatrix} L_{\text{Tr}}(g) & (1 - L_{\text{Tr}}(g)) & 0 \end{bmatrix}' \tag{11}$$

It is well known [19] that π_k contains all the information in Ψ_k so that the action $a_k = \mathbf{d}(\Psi_k) = \mathbf{d}(\pi_k)$ where $\mathbf{d} \in \mathcal{D}$. The cost function (6) can be expressed in terms of the information state π_k as [19]

$$V(\pi_1; \mathbf{d}) = \lim_{N \to \infty} \mathbf{E}_{\mathbf{d}} \left\{ \sum_{k=1}^{N} \sum_{i=1}^{3} c(X_i, a_k) \pi_k(i) \;\middle|\; \pi_1 \right\} \tag{12}$$

Because $q_{k+1} = 1 - p_{k+1}$ when $y_k = \text{NAK}$ for any k, we can completely parameterize π_{k+1} by p_{k+1}, see (9). In particular we can write $V(p_1, \mathbf{d})$ for $V(\pi_1, \mathbf{d})$ and $V(p_1)$ for the optimal cost $V(\pi_1)$. The optimal cost function $V(p_1)$ is the solution of Bellman's dynamic programming functional equation [5]

$$V(p_1) = \min \Big\{ c_1 + V(L_{\text{Tr}}(p_1))(P_e p_1 + 1 - p_1), \\ c_0 + V(L_{\text{Su}}(p_1)) \Big\} \tag{13}$$

It is well known [6] that (13) has a unique bounded solution, $V(p_1)$, that is piecewise linear and concave in p_1.

At this point, the ARQ transmission control has been formulated as a Markovian search problem of the form described in [6].

IV. Threshold Policy and Optimal Transmission Protocols

In general the solution to the DP equation (13) can be computed via a value iteration algorithm [20] – this is often computationally intractable. We show that due to the special structure of our search problem, the optimal retransmission control takes the form of a threshold policy, whereby the packet is retransmitted if the probability that the channel is in the good state exceeds a certain threshold level, or a suspension occurs if it does not.

A. Threshold Policies

The existence of optimal threshold policies for two-state Markovian search problems was conjectured by Ross in [5]. Modified for the transmission control problem, the conjecture reads as follows.

Ross' conjecture applied to transmission control: *There exists a threshold value, P^*, at any time n given the information state, p_n, such that for $p_n \geq P^*$, the optimal policy is to transmit, while for $p_n < P^*$, the optimal policy is to suspend.*

MacPhee and Jordan [6] prove this conjecture for a class of search problems that include the search problem we have formulated above in Sec.III. The reader is referred to [6] for details of the proof.

The threshold value, P^*, depends entirely on the channel transition probabilities R and costs c_0, c_1. Solving for P^* requires determining the class to which the system under consideration belongs. The classes are determined by the fixed points of the HMM channel predictor (10). Denote by P_{Tr} the fixed point of L_{Tr}, and P_{Su} the fixed point of L_{Su} in (10) of the HMM predictor. That is:

$$P_{\text{Tr}} = \frac{2-(P_e\,g+b)-\sqrt{(P_e\,g+b)^2-4\,P_e\,\mu}}{2\,(1-P_e)} \tag{14}$$

$$P_{\text{Su}} = \frac{1-b}{1-\mu} \tag{15}$$

Define also the mapping of a retransmission followed by a suspension, and vice-versa, required in the class definitions below:

$$P_{\text{Tr,Su}}(p) \stackrel{\triangle}{=} L_{\text{Su}}(L_{\text{Tr}}(p)), \qquad P_{\text{Su,Tr}}(p) \stackrel{\triangle}{=} L_{\text{Tr}}(L_{\text{Su}}(p)) \tag{16}$$

Lemma 1: $P_{\text{Su}} > P_{\text{Tr}}$ and $P_{\text{Tr,Su}} > P_{\text{Su,Tr}}$ for all $p \in (0,1)$.

Proof: Under the assumption that $0<\mu<1$ (see Sec.II), $L_{\text{Su}}(p_k)$ defined in (10) is both an increasing function of p_k and a contraction mapping because $(L_{\text{Su}}(x) - L_{\text{Su}}(y))^2 \leq \mu(x-y)^2$, for $0<x,y<1$.

The mapping $L_{\text{Tr}}(p_k)$ (10) is increasing and convex since

$$L'_{\text{Tr}}(p_k) = \frac{P_e\mu}{((P_e-1)p_k+1)^2} > 0,$$

$$L''_{\text{Tr}}(p_k) = \frac{2P_e\mu(1-P_e)}{((P_e-1)p_k+1)^3} > 0.$$

Note also that $L_{\text{Tr}}(0) = L_{\text{Su}}(0) = 1-b$ and $L_{\text{Tr}}(1) = L_{\text{Su}}(1) = g$. Therefore

$$L_{\text{Su}}(p) > L_{\text{Tr}}(p) \quad \forall p \in [0,1] \tag{17}$$

Due to (17), $L_{\text{Su}}(P_{\text{Tr}}) > L_{\text{Tr}}(P_{\text{Tr}})$. As $L_{\text{Su}}(p_k)$ is both increasing and contractive, $P_{\text{Su}} = L_{\text{Su}}(\ldots(L_{\text{Su}}(P_{\text{Tr}}))) > L_{\text{Tr}}(P_{\text{Tr}}) = P_{\text{Tr}}$. The proof of $P_{\text{Tr,Su}}(p) > P_{\text{Su,Tr}}(p)$ is given in [6, Lemma 4, pp.181]. ∎

The main result of [6, Sec.4 and 5] adapted to our transmission control problem reads as follows:

Result 1: Depending on the system parameters $\{b, g, c_0, c_1, P_e\}$, there are three classes into which the ARQ transmission control system can fall – Class 1, 2a and 2b. Each class has a different threshold value, P_1^*, P_{2a}^* and P_{2b}^* (19). Class membership rules are as follows:

Class 1: $P_{\text{Tr}} > P_1^*$ and $P_{\text{Su}} > P_1^*$

Class 2a: $P_{\text{Tr}} < P_{2a}^* < P_{\text{Su}}$ and $P_{\text{Su,Tr}}(P_{2a}^*) < P_{2a}^* < P_{\text{Tr,Su}}(P_{2a}^*)$ (18)

Class 2b: $P_{\text{Tr}} < P_{2b}^* < P_{\text{Su}}$ and $\left\{P_{2b}^* < P_{\text{Su,Tr}}(P_{2b}^*) \text{ or } P_{2b}^* > P_{\text{Tr,Su}}(P_{2b}^*)\right\}$

The fixed points, P_{Tr} and P_{Su}, are given in (14), while $P_{\text{Tr,Su}}(\cdot)$ and $P_{\text{Su,Tr}}(\cdot)$ are defined in (16). The three classes are mutually exclusive and exhaustive, i.e. the system belongs to one and only one of the classes. ∎

An important consequence of the above result is that class membership is a system property – the class is uniquely determined by the system parameters $\{b, g, c_0, c_1, P_e\}$.

The statement in Result 1 that there are only three classes can be argued as follows. There can be no class for which $P_{\text{Su}} < P_{\text{Tr}}$, as this violates the fixed point ordering in Lemma 1 above. Another potential case is that of $P_{\text{Tr}} < P_{\text{Su}} < P^*$, however this results in a threshold $P^* = 0$ given the parameters of the transmission system (see p.173 [6]), which implies $b=1$ (15), an illegal absorbing state in the channel – since we assume that channel is an aperiodic irreducible Markov chain.

The threshold values for Class 1 and 2a are explicitly computed in [6] as

$$P_1^* = \frac{(1-b)(c_1-c_0)}{(1-\mu)c_1}, \quad P_{2a}^* = \frac{(1-b)(c_1-\mu c_0)}{(1-\mu)(c_1+c_0)}. \tag{19}$$

The threshold for Class 2b, P_{2b}^*, cannot be obtained in closed form but is straightforwardly numerically computed as described in [6, Sec.4.2.4] by applying multiple compositions of L_{Tr} and L_{Su} (10).

B. Optimal Transmission Protocols

Based on Result 1, our optimal threshold policy based transmission controller operates as follows:

Optimal Transmission Control Algorithm

1) **System classification:** Given system parameters $\{b, g, c_0, c_1, P_e\}$ defined in (1) and (5),
   - Compute critical points, P_{Tr} and P_{Su}, according to (14)-(15).
   - Compute Class 1 and 2a thresholds, P_1^* and P_{2a}^*, according to (19).
   - Determine class membership, $m \in \{1, 2a, 2b\}$, according to (18). If the system is in Class 2b, determine the threshold numerically as described in [6, Sec. 4.4.2, p.176].
2) **Successful transmission until receipt of a NAK:**
   - Set time slot $k=0$, and
     - Transmit packet, $a_0 = \text{Tr}$.
     - Receive acknowledgement, $y_0 \in \{ACK, NAK\}$.
       * If $y_0 = ACK$, repeat Step 2.
       * Else if $y_0 = NAK$, then set $k=1$, initialise p_1 according to (11), and go to Step 3.
3) **Unsuccessful transmission until receipt of an ACK:** $(k = 1, 2, \ldots)$
   - If $p_k < P_m^*$, then $a_k = \text{Su}$.
   - Else if $p_k \geq P_m^*$, then
     - Retransmit packet, $a_k = \text{Tr}$.
     - Receive acknowledgement, $y_k \in \{ACK, NAK\}$.
       * If $y_k = ACK$, go to Step 2.
       * Else if $y_k = NAK$, then update p_k to p_{k+1} according to (10). Set $k = k+1$ and repeat Step 3.

In the rest of this section, we analyze the properties of the above optimal transmission control algorithm. We show that the class determines the structure of the optimal threshold policy. The following theorem formalizes the optimal control policies for the three different classes described in the above algorithm.

Theorem 1: (i) If the transmission system is in Class 1 and with initial condition $p_1 < P_1^*$, then the optimal policy is to first suspend transmission for $k^*(p_1)$ slots and then persistently retransmit until an ACK is received. Here $k^*(p_1)$ is the smallest non-negative integer satisfying

$$k^*(p_1) > \frac{1}{\log(\mu)} \log\left(\frac{c_0(1-b)}{c_1(1-b)-(1-\mu)c_1p_1}\right) \quad (20)$$

(ii) If the system is Class 1 and $p_1 > P_1^*$, the optimal policy is to persistently retransmit until an ACK is received.
(iii) If the system is in Class 2a or 2b, the optimal transmission policy is a threshold policy involving a combination of suspensions and retransmissions.

Proof: From the threshold Result 1, it follows that for Class 1, if $p_1 < P_1^*$ the optimal policy is to first suspend. As L_{Su} is increasing (10), p_1 will increase monotonically after each suspension until it exceeds P_1^* (which occurs with probability 1, as $P_{\text{Su}} > P_1^*$ for Class 1 systems), after which retransmission occurs until an ACK is received (with the information state moving towards P_{Tr}). Suspension cannot occur again once the threshold has been exceeded, as $P_{\text{Su}} > P_{\text{Tr}} > P_1^*$. Thus assertion (i) and (ii) follow. We now show that the suspension time k^* satisfies (20). Note that $k^* = \min\{k \geq 0 : p_k \geq P_1^*\}$. From (10),

$$p_{k+1} = \mu^k p_1 + (1-b)\frac{(1-\mu^k)}{1-\mu} \quad (21)$$

From the above equation and (19) we have $p_{k+1} \geq P_1^*$ implies that

$$\mu^k((1-\mu)p_1 - (1-b)) \geq \frac{c_0}{c_1}(b-1)$$

But $(1-\mu)p_1-(1-b) < 0$. (This holds because by assumption $p_1 < P_1^* = \frac{(1-b)}{(1-\mu)}(1-c_0/c_1)$. This implies $p_1 < \frac{(1-b)}{(1-\mu)}$ which in turn is equivalent to $(1-\mu)p_1 - (1-b) < 0$).
Thus from (21) we have

$$\mu^k \leq \frac{c_0(1-b)}{c_1(1-b)-(1-\mu)c_1p_1}$$

which in turn implies (20).

For Class 2 systems, due to the updates' fixed points, P_{Tr} and P_{Su}, lying either side of the threshold (see (18)), the optimal policy is a mixture of suspensions and retransmissions. ∎

Remarks: (i) From (20) it follows that as the suspension cost c_0 decreases, $k^*(p_1)$ (the number of time slots for suspension before persistent transmission) increases – which makes intuitive sense. However, since c_0 is strictly positive by assumption (5), $k^*(p_1)$ is always finite. If the energy cost for retransmission $c_1 = 0$, then $k^*(p_1) = 1$, i.e., for $c_1 = 0$ it is optimal to persistently retransmit – this also makes intuitive sense.

(ii) As mentioned in Sec.I, there are many ARQ protocols governing the retransmission of packets in wireless networks. Persistent Retransmission is the simplest, the idea being to retransmit a packet at each time instant until it transmits successfully, at which point the next packet is transmitted and so on. Persistent Retransmission is a Class 1 optimal policy that arises only if $p_1 > P_1^*$ (as the information state starts and remains above the threshold). In the numerical examples of Sec. VI this is shown to occur for low channel memory values.

C. Optimal Transmission Cost for Class 1 System

As described in Result 1 and Theorem 1 above, for Class 2a and 2b, the optimal transmission policies are threshold policies involving mixtures of suspensions and retransmissions. Therefore, apart from specifying the thresholds of the optimal transmission policy, it is difficult to say anything more specific about the policy of a Class 2 system since the switching between suspensions and transmissions depends on the specific parameters (g, b, P_e, c_0, c_1). However, for a Class 1 system, much more can be said about the performance of the optimal transmission control algorithm. We now give an explicit upper bound to the optimal cost incurred for a Class 1 system.

Recall from (9) that when the channel is in the terminal state, $x_k = T$, then the information state is

$$\pi_k = e_T \triangleq [0, 0, 1]'$$

First we demonstrate in the following theorem that if persistent transmission is applied, then after any two time transitions the probability of successfully receiving an ACK is strictly greater than zero. Recall from Theorem 1, that the optimal transmission policy for a Class 1 system is to first suspend transmission and then persistently retransmit.

Theorem 2: Consider a Class 1 system with information state $\pi = [\pi(1), \pi(2), \pi(3)]'$ in which persistent retransmission is applied starting at or before time $2k-1$, $k = 1, 2, \ldots$. Then, the conditional probability of a successful transmission is lower bounded as follows:

$$P(\pi_{2k+1} = e_T | \pi_{2k-1} = \pi) \geq \alpha \quad (22)$$

where the coefficient of contraction, $\alpha \in (0, 1]$, is

$$\alpha \triangleq (1-P_e)(1-b) \quad (23)$$

Proof: See Appendix.

Note that (22) is equivalent to $P(\pi_{2k+1} \neq e_T | \pi_{2k-1} = \pi) < 1 - \alpha$, meaning that the probability of not reaching the terminal state in any two slot time interval is strictly less than one. Thus the constant α in the above theorem acts as a discount factor (coefficient of contraction) that ensures that the infinite sum (6) and hence (25) below (which represents the expected number of retransmissions) is finite. In fact the above search problem can be viewed as an instance of a stochastic shortest path problem, see [19, pp.367]. It is important to note that the above contraction property requires at least two retransmission time steps. For one time step, it is easily shown that $P(\pi_{2k+1} = e_T | \pi_{2k} = \pi)$ is not always positive for arbitrary π. For example $\pi = [0, 1, 0]'$ implies $P(\pi_{2k+1} = e_T | \pi_{2k} = \pi) = 0$.

Suppose for any arbitrary time n, the information state π_n satisfies $p_n > P_1^*$. From Theorem 1 above, the optimal policy is to persistently retransmit. Starting at time n, let $M(\pi_n)$

denote the expected number of retransmissions required until an is ACK received, i.e.,

$$M(\pi_n) \triangleq \mathbf{E_d} \left\{ \sum_{k=n}^{\infty} I(x_k \neq T) \,\Big|\, \pi_n \right\} \quad (24)$$

where the policy $\mathbf{d} = (\mathrm{Tr}, \mathrm{Tr}, \mathrm{Tr}, \ldots)$ (i.e., persistently re-transmit) and $I(\cdot)$ denotes the indicator function. Note that by construction, if $\pi_n = e_T$, then $M(\pi_n) = 0$ (since T is an absorbing state).

The following theorem gives both an explicit expression for the expected number of retransmissions, $M(\pi_n)$, for a Class 1 system, and a conservative upper bound on the expected number of retransmissions.

Theorem 3: For a Class 1 system with $p_n > P_1^*$ at any arbitrary time n, the expected number of retransmissions is

$$M(\pi_n) = \sum_{k=1}^{\infty} k \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} Q^{\mathrm{Tr}}(\mathrm{ACK}) \left(R^{\mathrm{NAK}\prime} Q^{\mathrm{Tr}}(\mathrm{NAK}) \right)^{k-1} \pi_n \quad (25)$$

Furthermore, $M(\pi_n)$ is finite and upper bounded as

$$M(\pi_n) \leq \frac{2}{\alpha}, \quad (26)$$

where the coefficient of contraction, α, is defined in (23).

Proof: See Appendix

Remark: The proof of (26) is similar to the proof of finite expected cost for a stochastic shortest path problem, see for example [19, Sec.7.2]. However, we have the added complexity that the state of the Markov chain is not observed directly. Although (26) is conservative, it serves as an easily computed upper bound and also shows that the expected cost using persistent re-transmission is finite. Furthermore it is a tighter bound than that obtained using norms as we now show.

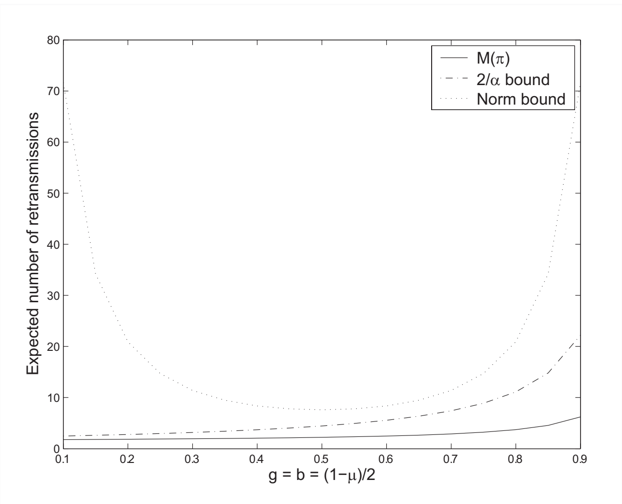

Fig. 1. A comparison of the expected number of retransmissions, $M(\pi_n)$, determined numerically according to (25), the upper bound (26), and the norm bound (27).

An upper bound on the expected number of retransmissions can be derived using a l_2 norm bound (see Appendix for the derivation):

$$M(\pi_n) \leq \frac{(1 - P_e)\sqrt{\pi(1)^2 + \pi(2)^2}}{(1 - \sigma_{\max})^2} \quad (27)$$

Here $\sigma_{\max}$ denotes the maximum singular value of the matrix $A_{\mathrm{NAK}} = R^{\mathrm{NAK}\prime} Q^{\mathrm{Tr}}(\mathrm{NAK})$. Fig. 1 compares the expected number of retransmissions (25) (determined numerically), the $2/\alpha$ bound (26), and the norm bound (27), for the parameter set (see Sec. VI) $c_0 = 1, c_1 = 3, P_e = 0.1, \pi(1) = \pi(2) = 0.5$. Note that the norm bound is at it's minimum for this choice of equiprobable initial channel state probabilities. As is evidenced in the plot, the $2/\alpha$ bound is tighter and follows the shape of $M(\pi_n)$ across varying channel memory, while the norm bound suffers from being symmetric.

The combination of Theorems 1 and 3 yields the following corollary which gives an explicit upper bound for the expected cost of the optimal transmission policy for a Class 1 system.

Corollary 1: For a Class 1 system, the expected cost of the optimal policy is given by

$$\begin{aligned} c_0 k^*(p_1) + c_1 M(\pi_{k^*(p_1)+1}) \quad & \text{if } p_1 < P^*, \\ c_1 M(\pi_1) \quad & \text{if } p_1 \geq P^* \end{aligned} \quad (28)$$

where $k^*(p_1)$ is defined in (20) and $M(\pi)$ is upper bounded by $2/\alpha$.

Thus we have shown in this section that for any Class 1 system, the optimal policy is to first suspend and then persistent retransmit once the information state has crossed the threshold P_1^*. Furthermore, the expected number of trans-missions is finite due to the contraction coefficient α. Finally, we have derived explicit expressions for the expected number of retransmissions and the cost.

V. Transformation to a POMDP

In this section we reformulate the search problem in Sec.III as a POMDP. This allows us to numerically solve the trans-mission control problem when the channel has more than two states or there are more than two observation symbols – in such cases the threshold policy derived earlier in this paper is no longer optimal. We can use available sophisticated software and various efficient exact POMDP algorithms such as the Witness algorithm, Incremental Pruning, etc which have recently been developed in the artificial intelligence community, see [21] for an excellent tutorial exposition with graphics of these various algorithms. Such software for solving POMDPs is typically designed assuming observation indepen-dent transition probabilities – whereas the search problem has observation dependent probabilities.

The search problem can be reformulated as a POMDP by a suitable coordinate transformation as outlined be-low, see [7] for details. Consider the augmented Markov process $\bar{x}_k \triangleq (x_k, y_{k-1})$ with state space $\bar{X} = \{(\mathtt{good}, \mathrm{NAK}), (\mathtt{bad}, \mathrm{NAK}), (T, \mathrm{ACK})\}$.

The action and observation sets remain the same. The transition probabilities of $\bar{x}_k$ are

$$R^a = \begin{bmatrix} Q^a(\mathrm{NAK})R & Q^a(\mathrm{ACK})\mathbf{1} \\ \mathbf{0}' & 1 \end{bmatrix} \quad (29)$$

which are no longer observation dependent. The observation probabilities $P(y_k = y|\bar{x}_k, a_k)$ are

$$O^a(y) = \begin{cases} \mathrm{diag}[1\ 1\ 0], & y = \mathrm{NAK} \\ \mathrm{diag}[0\ 0\ 1], & y = \mathrm{ACK} \end{cases} \quad (30)$$

The information state ($\bar{\pi}_k(i) = P(\bar{x}_k = \bar{X}_i|\Psi_k)$ is recursively computed via a Hidden Markov model filter update:

$$\bar{\pi}_{k+1} = \frac{O^{a_k} R^{a'_k} \bar{\pi}_k}{\mathbf{1}' O^{a_k} R^{a'_k} \bar{\pi}_k} \tag{31}$$

The costs are $g(\bar{x}_k, a_k) \triangleq c(x_k, a_k)$ – so that $g(T, a_k) = 0$, i.e., there is no cost for taking $a_k = \text{Tr}$ and receiving an ACK.

Solving the search problem (7) as formulated in Sec. III is equivalent to solving the undiscounted, infinite horizon POMDP:

$$V(p_1) = \sup_{\delta \in \Delta} V_\delta(\bar{\pi}_1), \qquad \text{where}$$

$$V_\delta(\bar{\pi}_1) = \lim_{N \to \infty} \mathbf{E}\Big\{ \sum_{k=1}^{N} g(\bar{x}_k, a_k)) \Big| \bar{\pi}_1 \Big\} \tag{32}$$

Here Δ denotes the class of stationary policies and $\delta : \bar{\pi}_k \to a_k$ denotes a stationary policy.

Remark: The formulation is straightfowardly extended to a $S > 2$ state channel with states $x_k \in \{X_1, X_2, \ldots, X_S, T\}$ (T again denotes the terminal state) and

$$\bar{x}_k \in \{(X_1, \text{NAK}), (X_2, \text{NAK}), \ldots, (X_S, \text{NAK}), (T, \text{ACK})\}$$

Let R^a now denote the $S \times S$ transition probability matrix of the channel and $Q^a(y) = \text{diag}[\Pr(y_k = \text{NAK}|s_k = 1, a_k = a), \ldots, \Pr(y_k = \text{NAK}|s_k = L, a_k = a)]$. Then (31), (32) hold. The POMDP solution software in [21] can solve POMDPs with arbitrary finite state, finite observation symbols and finite number of actions. For small size problems of up to ten states, Lovejoy's sub-optimal algorithm [22] can be used. Lovejoy's algorithm also includes a constructive procedure for computing upper and lower bounds on the approximate transmission policy. We refer the reader to [22] and [2] for details of Lovejoy's algorithm and applications to sensor scheduling in network centric warfare. It is important to note however, that if the channel has more than two states, the optimal policy is not necessarily threshold.

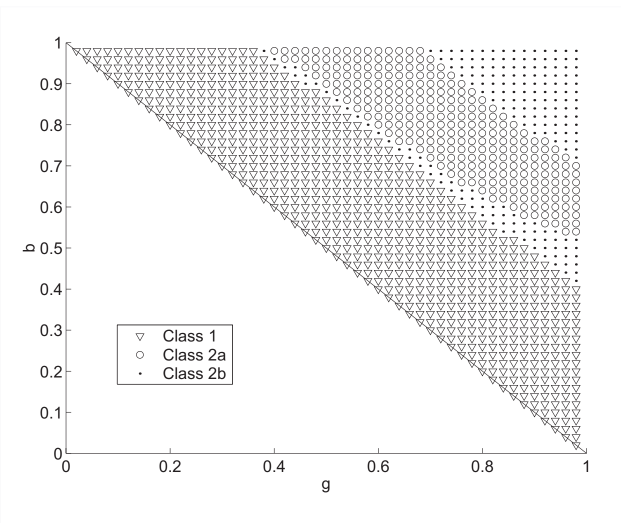

Fig. 2. Class boundaries as a function of channel parameters g and b, such that $\mu = g + b - 1 > 0$. Costs are $c_0 = 1, c_1 = 3$, and error probability, $P_e = 0.1$.

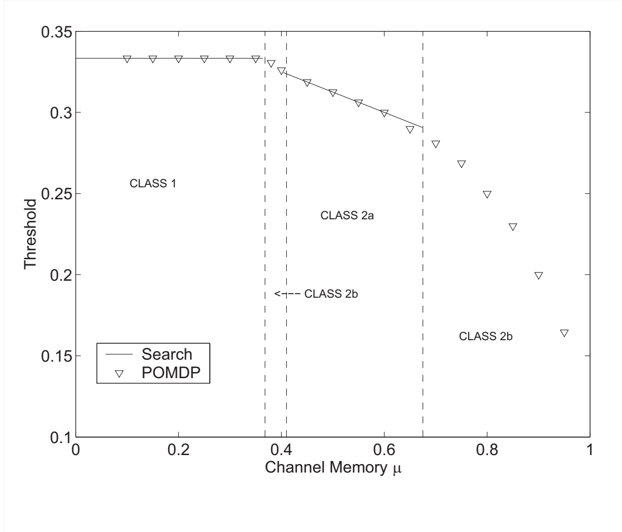

Fig. 3. Threshold values as specified by search formulation and POMDP solution. $c_0 = 1$, $c_1 = 3$, $P_e = 0.1$, $g = b$.

VI. Numerical Examples

Throughout this section we consider a two-state Gilbert Elliott channel with transition probabilities specified below, error probability $P_e = 0.10$, and latency (suspension) and transmission energy (failure) costs fixed, respectively, at $c_0 = 1$ and $c_1 = 3$. These parameter values were chosen since they permit full exploration of the range of classes (1,2a and 2b) through varying channel memory, $\mu = g + b - 1$, as will be shown in the following results.

a) Class Boundaries.: Here we elucidate the effect of the channel parameters on the threshold values determining class membership. Fig. 2 shows the class boundaries over the range of transition probabilities, g, b, for channel memory $\mu \geq 0$. i.e. the region for which $g + b \geq 1$. These were computed using the system classification step of the transmission control algorithm summarized in Sec.IV-B. Each of the three classes, Class 1, Class 2a and Class 2b, is observed to be active across the range of channel memory values $0 \leq \mu < 1$.

b) Equivalence of Search and POMDP Formulations.: Recall that in Sec.V, we analytically demonstrated the equivalence between the search and POMDP formulations. Here we illustrate their equivalence numerically. To visualize the thresholds on a 2-d plot we chose $g = b$. Thus we are moving across the various classes in Fig. 2, along a diagonal line starting at $(g, b) = (0, 0)$ and ending in $(g, b) = (1, 1)$.

Fig. 3 depicts the threshold values obtained using the search formulation. For Class 1 and 2a these threshold were computed using (19). For Class 2b the thresholds were numerically computed as in [6, Sec.4.2.2]. Note that for low correlation fading (small μ), the system is in Class 1 and suspensions followed by persistent transmission is the optimal strategy. This is similar to the finding of [9] where it was shown that for low correlation fading persistent ARQ is more energy efficient. For higher correlation fading (larger μ), the system is in Class 2a and 2b where the optimal policy is a mixture of suspensions and transmissions.

Also shown in Fig. 3 are the threshold values obtained via solving the POMDP formulated in Sec.V. As might be expected, there is an exact match. The POMDP was solved using the "incremental-prune" algorithm recently developed in the artificial intelligence community [3]. The software is freely downloadable from [21].

c) Transmission Protocol Comparison.: Here we compare the performance of our threshold based optimal transmission control algorithm with two other schemes, namely, Persistent Retransmission, and the Back-Off-N scheme. As mentioned in Sec. I, Back-Off-N schemes are proposed in [16] for energy efficient transmission over fading channels. We therefore implement a Back-Off-N scheme with suspension length, N, designed as a function of channel memory μ as follows (the values of N chosen below are optimal for the ARQ system in [16]):

μ	0 – .3	.35 – .6	.65 – .75	.8	.85	.9	.95
N	0	1	2	3	4	5	7

To visualize the performance of the three transmission control algorithms on a 2-d plot, we consider $g = b$. Figs. 4-6 compare performance statistics of three algorithms, for a simulation of 1 million successful packet transmissions per channel memory value. Note that the jagged nature of the plots for the optimal transmission scheme is due to the fact that with increasing channel memory μ, the system class changes and thus the thresholds change. Similarly, for the Back-Off-N scheme, the value of N jump changes according to the above table thus exhibiting the jagged behaviour. It is important to note that the jagged behaviour is not due to the statistical variance of the simulation.

The performance statistics of the simulation depicted in Figs. 4-6 are throughput, efficiency and average cost, defined as follows:

$$\text{Throughput} = \frac{\text{number of packets successfully transmitted}}{\text{total number of transmission slots}}$$
$$\text{Efficiency} = \frac{\text{number of packets successfully transmitted}}{\text{total number of packets transmitted}}$$
$$\begin{aligned}\text{Cost} &= c_0 \times \text{number of packet suspensions}\\ &\quad + c_1 \times \text{number of transmission failures}\end{aligned}$$
$$\text{Average Cost} = \frac{\text{Cost}}{\text{total number of transmission slots}}$$

The Persistent Retransmission scheme shows constant performance in throughput, efficiency and cost versus channel memory μ. This is because choosing $g = b$ renders the steady state channel distribution $[0.5\ 0.5]$. Persistent Retransmission provides an upper limit on the throughput since a packet is transmitted every time slot, irrespective of the channel state.

Fig. 5 shows that the efficiency of the Back-Off-N scheme increases monotonically as μ increases. This is due to the fact that the number of suspensions N increases monotonically with channel memory μ in the above table. Fig. 5 shows that the efficiency of the optimal threshold based transmission controller also increases with μ. This can be intuitively explained as follows: As μ increases, the system moves out of Class 1 and into either Class 2a or 2b (see Fig. 2). In Class 2a and 2b, the fixed point P_{Tr} lies below the threshold (18). Therefore, in Class 2a and b, every transmission yields an updated information state π_k which results in a higher probability of suspension. A higher probability of suspension implies that more time is allowed between transmissions for the channel to recover to the good state.

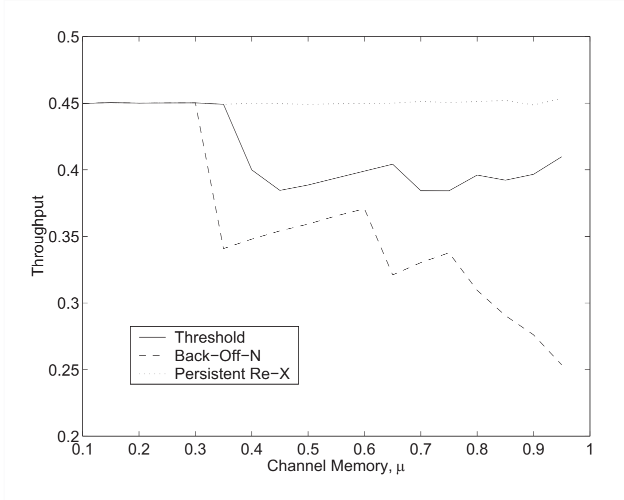

Fig. 4. Throughput vs. channel memory. For the Class 1 system ($\mu < 0.36$, refer to Fig. 3), throughput = 0.45 (see (33)).

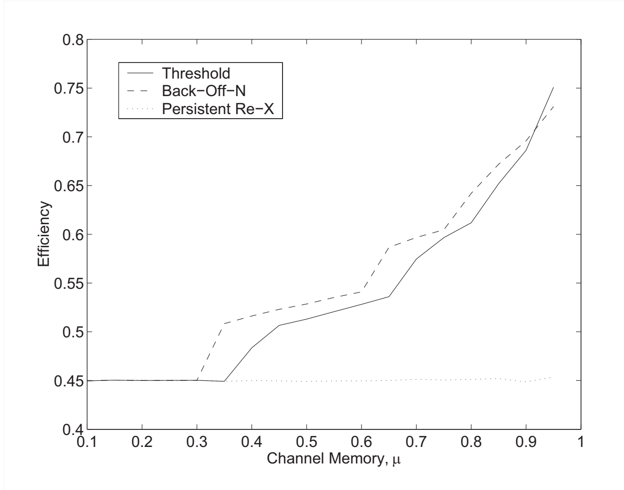

Fig. 5. Efficiency vs. channel memory. For the Class 1 system ($\mu < 0.36$, refer to Fig. 3), efficiency = 0.45 (see (33)).

The most important point to note from this example is the comparison between the optimal controller and the Back-Off-N scheme. For low channel memory both algorithms behave like the Persistent Retransmission as theory predicts. When packet suspensions start playing a role, ($\mu = 0.36$), the cost increases due to the impact of c_0. For higher channel memory values, the cost decreases as less retransmissions are attempted. The non-optimized Back-Off-N scheme does not match the decreasing cost curve of the search algorithm, clearly supporting the optimality of this latter method. Similarly for the throughput statistic, the search algorithm is superior to Back-Off-N across the entire memory range. Efficiency is slightly better for the Back-Off-N scheme however, as it is prone to suspend packets more than the search algorithm.

d) Numerical Evaluation of Class 1 System Performance.: For a Class 1 system, the values in Figs. 4-6 can be explicitly computed using (25) and (28) as follows: Since we have chosen $g = b = (\mu + 1)/2$, for $\mu < 0.36$ the system

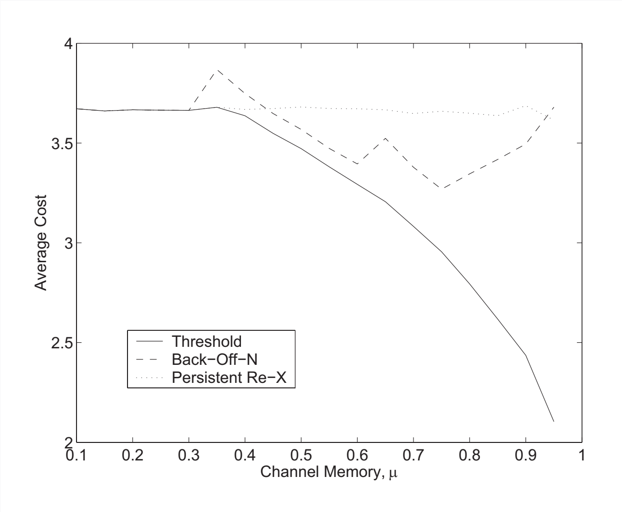

Fig. 6. Cost vs. channel memory. For the Class 1 system ($\mu < 0.36$, refer to Fig. 3), average cost = 3.67 (see (34)).

is in Class 1 (see Fig. 3) and $p_1 > P_1^*$. Thus, from Corollary 1, the optimal policy is persistent retransmission. From (28), $M(\pi_1) = 1.22$. Considering also the failed transmission at $k = 0$ which initiates the controller at each new session, the expected number of total transmissions required for a successful transmission of a packet in any one session is $1 + M(\pi) = 2.22$. Therefore, as is verified in Figs. 4-5,

$$\text{Class 1 efficiency} = \text{Class 1 throughput} = 1/2.22 = 0.45 \quad (33)$$

The average cost incurred $= P(\text{controller starts to operate})(1 + M(\pi_1))c_1$. In steady state, $P(\text{channel} = \texttt{good}) = P(\text{channel} = \texttt{bad}) = 0.5$, since $g = b$. Thus,

$$\begin{aligned} &P(\text{controller starts to operate}) \\ &\quad = P(y_k = \text{NAK}|s_k = \texttt{good})P(s_k = \texttt{good}) \\ &\quad + P(y_k = \text{NAK}|s_k = \texttt{bad})P(s_k = \texttt{bad}) \\ &\quad = P_e \times 0.5 + 1 \times 0.5 = 0.55 \end{aligned}$$

Hence the average cost for Class 1 ($\mu < 0.36$) persistent retransmission, as verified in Fig. 6, is

$$P(\text{controller starts to operate}) \times (1 + M(\pi_1)) \times c_1 = 3.67 \quad (34)$$

VII. Conclusion

The ARQ transmission control problem was formulated as an optimal search problem for a Markovian target which is an instance of a partially observed Markov decision process (POMDP). While in general POMDPs are computationally intractable, due to the special structure of ARQ transmission problem, we derived optimal *threshold* policies for transmission and suspension depending on the predicted state of a two-state Gilbert Elliott channel. We showed that depending on the system parameters $\{b, g, c_0, c_1, P_e\}$, there are only three classes of systems – each exhibiting a threshold policy. The resulting policies are computationally efficient to implement and are shown in Sec.VI to outperform non-optimal heuristic control algorithms.

Apart from the threshold policies derived in this paper, there is one other instance of a POMDP that leads to computationally tractable optimal policies – namely, multiarmed bandit POMDPs. We refer the reader to [23] for details of multiarmed POMDPs. It is of interest to exploit the results in [23] to develop efficient cross-layer scheduling algorithms in wireless networks.

References

[1] S. Wicker, *Error Control Systems for Digital Communication and Storage*. Upper Saddle River, NJ: Prentice Hall, 1995.

[2] V. Krishnamurthy, "Algorithms for optimal scheduling and management of hidden markov model sensors," *IEEE Trans. Signal Processing*, vol. 50, pp. 1382–1397, June 2002.

[3] A. Cassandra, M. Littman, and N. Zhang, "Incremental pruning: A simple fast exact method for partially observed Markov decision processes," *Proc. 13th Annual Conference on Uncertainty in Arficial Intelligence*, 1997.

[4] L. Stone, "What's happened in search theory since the 1975 Lanchester prize," *Operations Research*, vol. 37, pp. 501-506, May-June 1989.

[5] S. Ross, *Introduction to Stochastic Dynamic Programming*. San Diego: Academic Press, 1983.

[6] I. MacPhee and B. Jordan, "Optimal search for a moving target," *Probability in the Engineering and Information Sciences*, vol. 9, pp. 159-182, 1995.

[7] S. Singh and V. Krishnamurthy, "The optimal search for a Markovian target when the search path is constrained: the infinite horizon case," *IEEE Trans. Automat. Contr.*, vol. 48, pp. 487-492, Mar. 2003.

[8] M. Zorzi and R. Rao, "On the statistics of block errors in data transmission over fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 660-667, June 1997.

[9] M. Zorzi and R. Rao, "Error control and energy consumption in communications for nomadic computing," *IEEE Trans. Comput.*, vol. 46, pp. 279-289, 1997.

[10] M. Zorzi, R. Rao, and L. Milstein, "Error statistics in data transmission over fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 1468-1477, Nov. 1998.

[11] J. Cavers, "An analysis of pilot-symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686-693, Nov. 1991.

[12] P. Spasojevic and C. Georghiades, "Complementary sequences for ISI channel estimation," *IEEE Trans. Inform. Theory*, vol. 3, pp. 1145-1152, Mar. 2001.

[13] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, pp. 1951-1968, 1998.

[14] H. Wang and N. Mandayam, "Opportunistic file tranfer over a fading channel under energy and delay constraints," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 632-644, Apr. 2005.

[15] G. Koole, Z. Liu, and R. Righter, "Optimal transmission policies for noisy channels," *Operations Research*, vol. 49, pp. 892-899, Nov.-Dec. 2001.

[16] D. Zhang and K. M. Wasserman, "Energy efficient data communication over fading channels," *Proc. IEEE Wireless Communcations and Networking Conference*, pp. 986-991, 2000.

[17] M. Zorzi, R. R. Rao, and L. B. Milstein, "On the accuracy of a first-order Markov model for data transmission on fading channels," *Prof. IEEE Universal Pers. Commun. Conf.*, pp. 211-215, 1995.

[18] S. Pollock, "A simple model of search for a moving target," *Operations Research*, vol. 18, pp. 893-903, 1970.

[19] D. Bertsekas, *Dynamic Programming and Optimal Control*, vols. 1 and 2. Belmont, MA: Athena Scientific, 1995.

[20] W. Lovejoy, "A survey of algorithmic methods for partially observed Markov decision processes," *Annals of Operations Research*, vol. 28, pp. 47-66, 1991.

[21] A. Cassandra, "Tony's POMDP page." Available: http://www.cs.brown.edu/research/ai/pomdp/index.html.

[22] W. Lovejoy, "Computationally feasible bounds for partially observed Markov decision processes," *Operations Research*, vol. 39, pp. 162-175, Jan.-Feb. 1991.

[23] V. Krishnamurthy, “Emission management for low probability intercept sensors in network centric warfare,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, issue 1, pp. 133-151, Jan. 2005.

APPENDIX

PROOF OF THEOREM 2

For notational convenience denote $\mathbf{e}_1 = [1, 1, 0]$. Then,

$$\begin{aligned} P(\pi_{2k+1} = e_T|\pi_{2k-1} = \pi) &= P(\pi_{2k-1} = e_T|\pi_{2k-1} = \pi) \\ &+ P(\pi_{2k} = e_T, \pi_{2k-1} \neq e_T|\pi_{2k-1} = \pi) \\ &+ P(\pi_{2k+1} = e_T, \pi_{2k} \neq e_T, \pi_{2k-1} \neq e_T|\pi_{2k-1} = \pi) \\ &= \pi(3) + \mathbf{e}_1 Q^{\text{Tr}}(\text{ACK})\pi \\ &+ \mathbf{e}_1 Q^{\text{Tr}}(\text{ACK}) R^{\text{NAK}\prime} Q^{\text{Tr}}(\text{NAK})\pi \end{aligned} \tag{35}$$

Expanding (35) results in

$$\begin{aligned} P(\pi_{2k+1} &= e_T|\pi_{2k-1} = \pi) \\ &= \pi(3) + (1 - P_e)\Big((1 + P_e g)\pi(1) + (1 - b)\pi(2)\Big) \\ &\geq \pi(3) + (1 - P_e)(1 - b)\Big(\pi(1) + \pi(2)\Big) \\ &\geq \alpha \end{aligned}$$

where the first inequality follows from the fact that $(1+P_e g) > (1 - b)$, and the second from the definition of α (23), and (9), since either $(\pi(1)+\pi(2)) = 1$ and $\pi(3) = 0$, or $\pi = [0, 0, 1]'$. ■

PROOF OF THEOREM 3

To show (25) it is convenient to work with the following equivalent definition of $M(\pi_n)$:

$$M(\pi_n) = \sum_{k=n}^{\infty}(k + 1 - n)P(y_k = \text{ACK}, y_{k-1} = \text{NAK}, \ldots, y_n = \text{NAK}|\pi_n) \tag{36}$$

The equality is explained as follows. With the search problem commencing at time n, $P(y_k = \text{ACK}, y_{k-1} = \text{NAK}, \ldots, y_n = \text{NAK}|\pi_n)$ represents the probability that the object is found in the $(k + 1 - n)$th search, i.e., an ACK is received for the first time. Similar to the derivation of (8) via Bayes' rule it readily follows by summing over all possible realizations of $(x_{k-1}, \ldots, x_n)$ that

$$\begin{aligned} P(y_k = \text{ACK}, x_k, y_{k-1} &= \text{NAK}, \ldots, y_n = \text{NAK}|\pi_n) \\ &= Q^{\text{Tr}}(\text{ACK})(R^{\text{NAK}\prime} Q^{\text{Tr}}(\text{NAK}))^{k-1}\pi_n. \end{aligned}$$

Summing over the possible outcomes $x_k = \{\texttt{good}, \texttt{bad}\}$ (clearly $x_k = T$ is impossible since $y_{k-1} = \text{NAK}$) is equivalent to pre-multiplying the LHS by $\begin{bmatrix}1 & 1 & 0\end{bmatrix}$. Then subsitituting this expression into (36) yields (25).

To simplify notation, in the remainder of the proof we consider $n = 1$. (Dealing with $n > 1$ merely involves a time shift.) Note that from (24), because the expected value of an indicator variable is a probability, $M(\pi_1) = \sum_{k=1}^{\infty} P(\pi_k \neq e_T|\pi_1)$. Because of the Markovian transition structure of π_k in (8),

$$\begin{aligned} &P(\pi_{2k+1} \neq e_T|\pi_1) \\ &= \int_{\pi \neq e_T} P(\pi_{2k+1} \neq e_T|\pi_{2k-1} = \pi) \\ &p_{\pi_{2k-1}|\pi_1}(\pi_{2k-1} = \pi|\pi_1) d\pi \end{aligned}$$

where $p_{\pi_{2k-1}|\pi_1}(\cdot|\pi_1)$ denotes the conditional density of π_{2k-1} given π_1. As shown in Lemma 2, for any k, $P(\pi_{2k+1} \neq e_T|\pi_{2k-1}) \leq (1 - \alpha)$. Thus

$$P(\pi_{2k+1} \neq e_T|\pi_1) \leq (1 - \alpha)P(\pi_{2k-1} \neq e_T|\pi_1) \leq (1 - \alpha)^k$$

which implies that

$$\begin{aligned} M(\pi_1) &= \sum_{k=1}^{\infty} P(\pi_k \neq e_T|\pi_1) \\ &= P(\pi_1 \neq e_T) + \sum_{k=1}^{\infty}[P(\pi_{2k} \neq e_T|\pi_1) + P(\pi_{2k+1} \neq e_T|\pi_1)] \\ &\leq 2\sum_{k=1}^{\infty}(1 - \alpha)^k = 2/\alpha \end{aligned}$$

■

DERIVATION OF THE NORM BOUND (27)

The expected number of retransmissions, $M(\pi_n)$, is finite, and therefore from (25):

$$M(\pi_n) = \sum_{k=1}^{\infty} k\begin{bmatrix}1 & 1 & 0\end{bmatrix} Q^{\text{Tr}}(\text{ACK})\Big(R^{\text{NAK}\prime} Q^{\text{Tr}}(\text{NAK})\Big)^{k-1}\pi_n \tag{37}$$

$$= \left|\left|\sum_{k=1}^{\infty} k\begin{bmatrix}(1 - P_e) & 0 & 0\end{bmatrix} A_{\text{NAK}}^{k-1}\,\pi\right|\right|_2 \tag{38}$$

$$\leq \sum_{k=1}^{\infty} k\,(1 - P_e)\left|\left|\,A_{\text{NAK}}^{k-1}\,\pi\right|\right|_2 \tag{39}$$

$$\leq \sum_{k=1}^{\infty} k\,(1 - P_e)\left|\left|A_{\text{NAK}}\right|\right|_2^{k-1}\,||\pi||_2 \tag{40}$$

where $A_{\text{NAK}} \triangleq R^{\text{NAK}\prime} Q^{\text{Tr}}(\text{NAK})$, and $\|A_{\text{NAK}}\|_2$ denotes the l_2 induced norm of the matrix A_{NAK}, i.e., the maximum singular value $\sigma_{\max}$ of A_{NAK}. Note that the first inequality above follows from the triangle inequality and the second inequality from the sub-multiplicative property of matrix induced norms. Substitution of the l_2 norm, $\sigma_{\max}$, gives

$$M(\pi_n) \leq \sum_{k=1}^{\infty} k\,(1 - P_e)\,\sigma_{\max}^{k-1}\sqrt{\pi(1)^2 + \pi(2)^2} \tag{41}$$

which straightfowardly results in the bound (27) upon recognizing the geometric sequence derivative in (41).

Leigh Johnston received the BSc and BE degrees in computer science and electrical engineering in 1997 and the PhD degree in electrical engineering in 2000 from the University of Melbourne, Australia. She was a postdoctoral researcher at the Centre for Systems Engineering and Applied Mechanics at the Universite Catholique de Louvain, Beligum from 200-2001, and in the Centre for Vision Research, York University, Canada from 2002-2004. Leigh is currently with the Howard Florey Institute in Melbourne, Australia, pursuing research interests in neurogenomics and brain imaging.

Vikram Krishnamurthy (Fellow'05) was born in 1966. He received his bachelor's degree from the University of Auckland, New Zealand in 1988, and Ph.D. from the Australian National University, Canberra, in 1992. Since 2002, he has been a professor and Canada Research Chair at the Department of Electrical Engineering, University of British Columbia, Vancouver, Canada. Prior to this he was a chaired professor at the Department of Electrical and Electronic Engineering, University of Melbourne, Australia, where he also served as Deputy head of department. His research interests span several areas including ion channels and nanobiology, stochastic scheduling and control, statistical signal processing and wireless telecommunications.

Dr. Krishnamurthy is currently an associate editor for *IEEE Transactions on Signal Processing*, *IEEE Transactions Aerospace and Electronic Systems*, *IEEE Transactions Nanobioscience* and *Systems and Control Letters*. He was a guest editor of a special issue of *IEEE Transactions on NanoBioScience*, March 2005 on bio-nanotubes.