Energy-Delay tradeoff in Multiple Access Channels

Achilleas Anastasopoulos
Electrical Engineering and Computer Science Department
University of Michigan
Ann Arbor, MI, 48109-2122
email:anastas@umich.edu

Abstract

In this paper we study the tradeoff between energy and delay in a multiple access channel (MAC). Messages arrive at the queues of the individual users and users decide whether to service them. Transmission of a message consumes energy and is successful only if one user transmits at a time. Users do not communicate but can observe a broadcasted feedback message from the base station indicating the success or not of the previous transmission. Delays are captured by considering the queue lengths of each user.

We formulate this problem as a decentralized stochastic control problem, the two controllers being the two users who, in the presence of limited information about each other, decide whether to transmit at each time slot. The decentralized aspect of this control problem makes it fundamentally different from the corresponding single-user counterparts and multi-user counterparts assuming a centralized controller. As a result, the tools from Markov decision processes (MDPs) and partially observed MDPs (POMDPs) cannot be directly applied.

Our contribution is twofold. First, we identify structural properties of the optimal transmission strategies for the two users so that the domain of the optimal strategies is not increasing with time. Second, based on the above structural properties, we identify the optimal strategies as the solution of a fixed point equation.

I. Introduction

Information theory has been extremely successful in characterizing the maximum information rate possible in point-to-point single-user channels. It has also been quite successful in characterizing the set of information rates that are simultaneously possible in multi-user networks, through the concept of capacity regions. This success can, to a large extend, be attributed to the very way the information theoretic problems are formulated: the quantities of interest are the exponential growth rates of number of messages (i.e., the code rates). As the codeword length grows to infinity, a concentration of measure occurs and a typical behavior emerges. In this framework, there is only a tradeoff between energy and rates: the higher the energy available for transmission, the higher the possible rates. However, the transmission delays are infinite.

A refinement of the capacity results is offered by the concept of error exponents that provide the rate of exponential decay of the error probability with respect to the codeword lengths as functions of the transmission rates. Although the best error exponent (reliability function) is not even known for a point-to-point channel for all rates, significant progress has been made in the direction of bounding the reliability function of single- and multi-user channels (see for instance [1]–[6] for the state of the art and recent advances relating to error exponents for the multiple access channel (MAC)). In this refined setup, the tradeoff between energy and delay can be thought of equivalently as the tradeoff between energy and error exponent(s) for a given set of rates. The disadvantage of this formulation is that it is still inherently asymptotic (error exponents are negative exponential growth rates of the error probabilities) and thus not well suited for real-time communication scenaria. In addition, the above approach assumes that users always have a message to transmit and cannot account for the dynamics introduced by random message arrivals and queueing.

In this paper we study the tradeoff between energy and delay in a MAC. Messages arrive at the queues of the individual users according to independent arrival processes. A simple model is used to capture the multiple access aspect of the channel: transmission is successful only if one user transmits; otherwise a catastrophic collision occurs. Every transmission incurs a cost that can be thought of as the energy consumed for the transmitted packet. Users do not communicate but can observe a broadcasted feedback message from the base station indicating the success or not of the previous transmission. Delays are captured by considering the queue lengths of each user. Clearly an energy-delay tradeoff exists since if a user chooses not to transmit, the corresponding queue length will increase; alternatively, if a user chooses to transmit it will consume energy and in addition it will run the risk of colliding with the other user's transmission.

We formulate this problem as a decentralized stochastic control problem, the two controllers being the two users that, in the presence of limited information about each other, decide whether to transmit at each time slot. Although the two users have limited information about each-other's queues, they have a common objective, i.e., to minimize the average of some linear combination of their queue lengths and consumed energy. In this respect the problem can be classified as a dynamic team problem with non-classical information structure. The decentralized aspect of this control problem makes it fundamentally different from the corresponding single-user counterparts (see for instance [7] for a study of the energy-delay tradeoff in point-to-point links) and multi-user counterparts assuming a centralized controller [8]. As a result, the problem under consideration cannot be solved applying directly the tools from Markov decision processes (MDPs) and partially observed MDPs (POMDPs). Decentralized stochastic control problems with non-classical information structure similar to the one considered here have been studied in [9], [10] and recently in [11]–[15]. Our contribution is twofold. First, we identify structural properties of the optimal transmission strategies for the two users so that the domain of the optimal strategies is not increasing with time. In general, at time t, each user can base its decision on everything that he has seen and done, i.e., on the history of his queue lengths, previous decisions and received feedback. Obviously such an increasing domain is not practical for implementation and it also generates conceptual difficulties for the infinite-horizon. Second, based on the above structural properties, we identify the optimal strategies as the solution of a fixed point equation. This solution is obtained by considering the problem from the viewpoint of a fictitious agent that observes the common information to both users.

The remaining of this paper is structured as follows. The model of the studied communication system is presented in Section II. Section III summarizes the results of an idealized system whereby an omniscient controller has perfect information about the queues of both users. The performance of this system can serve as an upper bound to the decentralized case. Structural properties of the optimal strategies and the optimal solution are developed in Section IV, and Section V concludes the paper.

II. System model and problem statement

We consider a multiple access (uplink) communication system consisting of two users and one base station. Each user is equipped with an infinite length queue. The queue length of user i at time t is denoted by x_t^i. Packets arrive at the queues of each user according to independent Bernoulli arrival processes $(a_t^i)_{t=1}^{\infty}$ ($a_t^i \in \{1,0\}$, with $a_t^i = 1$ denoting an arrival at the the i-th user's queue at time t) with identical probability mass functions (pmfs) $P_a(\cdot)$, with $P_a(1) = Pr(a_t^i = 1) = p$.

At each time $t \geq 1$ each user decides whether to transmit a packet from its queue. This decision is denoted by $u_t^i \in \{0,1\}$, with $u_t^i = 1$ denoting transmission of user i at time t. Transmission of a packet incurs a cost b which can be though of as the energy spent for the transmission. If only one of the two users transmit at time t, the transmission is successful. If both users transmit at the same time, the two transmissions are assumed to create a collision in which case, none of the transmissions is successful.

This is summarized in the following update equations for $t \geq 1$

$$x^1_{t+1} = |x^1_t - u^1_t(1-u^2_t)|^+ + a^1_t \tag{1a}$$

$$x^2_{t+1} = |x^2_t - u^2_t(1-u^1_t)|^+ + a^2_t, \tag{1b}$$

where $|x|^+ \stackrel{\text{def}}{=} \max(x,0)$. It is also assumed that $P(x^1_1, x^2_1) = Q(x^1_1)Q(x^2_1)$, for some given initial distribution $Q(\cdot)$.

The base station uses a feedback channel to inform the transmitters about the state of their transmission by broadcasting at time $t \geq 2$ the signal $y_t \in \mathcal{Y} \stackrel{\text{def}}{=} \{0, NACK, ACK1, ACK2\}$. In particular, $y_t = 0$ indicates that none of the two users attempted a transmission in the previous time slot; $y_t = NACK$ indicates that both users attempted a transmission in the previous time slot and thus they collided; $y_t = ACKi$ indicates that only user i attempted a transmission in the previous time slot and thus succeeded.

$$y_t = \begin{cases} 0 & u^1_{t-1} = u^2_{t-1} = 0 \\ NACK & u^1_{t-1} = u^2_{t-1} = 1 \\ ACK1 & u^1_{t-1} = 1 \text{ and } u^2_{t-1} = 0 \\ ACK2 & u^1_{t-1} = 0 \text{ and } u^2_{t-1} = 1. \end{cases} \tag{2}$$

Observe that y_t is in one-to-one correspondence with the pair (u^1_{t-1}, u^2_{t-1}).

We assume that each user's decision on whether to transmit at time t depends on its own queue-length history $x^{i,t}$ and the common observation y^t, i.e., $u^i_t = f^i_t(x^{i,t}, y^t)$ (at $t = 1$ we have $u^i_1 = f^i_1(x^i_1)$). At each time t, the instantaneous cost incurred at the system level is

$$c(x^1_t, x^2_t, u^1_t, u^2_t) = x^1_t + x^2_t + bu^1_t + bu^2_t. \tag{3}$$

This instantaneous cost penalizes long queues (and thus packet delay), and also accounts for the (energy) cost of a transmission. We are interested in finding strategies f^1, f^2, where $f^i \stackrel{\text{def}}{=} (f^i_t)^\infty_{t=1}$ that minimize the average discounted cost, i.e.,

$$J^* = \min_{f^1, f^2} \mathrm{E}\{\sum_{t=1}^{\infty} \lambda^{t-1} c(x^1_t, x^2_t, u^1_t, u^2_t)\}, \tag{4}$$

where $\lambda \in (0,1)$ is the discount factor.

We note that the above problem is neither an MDP nor a POMDP problem. This problem can be classified as a decentralized dynamic team problem with non-classical information structure. This information structure has some similarities to the delayed sharing pattern of [9], [10] with the difference being that the common information between the two users y_t only involves the delayed actions (u^1_{t-1}, u^2_{t-1}) and not delayed information about the queue lengths. On the other hand, the information structure of the studied problem can be thought of as a special case of model A in [15]. Although this problem does not fall under the class of MDP or POMDP problems, it is instructive to look at two simplified versions of this problem that are MDP problems: the first one is the problem of controlling the transmission of a single user in a point-to-point link; the second one is the same as the problem at hand with the only caveat that there is a centralized controller that observes both users' queue lengths x^i_t. In the following section we summarize the solutions of these two simplified problems.

III. Summary of centralized control results

A. Summary of single-user results

For $b \leq \frac{\lambda}{1-\lambda}$, the optimal policy is to transmit when the queue is non-empty. The average discounted cost-to-go when the initial queue length is x is given by

$$V(x) = \frac{b(1-\lambda) - \lambda(1-p)}{(1-\lambda)^2} + \frac{1}{1-\lambda}x + \frac{\lambda(1-p) - b(1-\lambda)(1-\lambda p)}{(1-\lambda)^2}\left[\frac{\lambda - \lambda p}{1-\lambda p}\right]^x, \qquad x \geq 0. \quad (5)$$

The resulting Markov chain representing the user's queue length has a steady-state distribution with

$$p(i) = \begin{cases} 1-p & i = 0 \\ p & i = 1 \\ 0 & i \geq 2. \end{cases} \quad (6)$$

For $b \geq \frac{\lambda}{1-\lambda}$, the optimal policy is to not transmit at all. The average discounted cost-to-go when the initial queue length is x is given by

$$V(x) = \frac{\lambda p}{(1-\lambda)^2} + \frac{1}{1-\lambda}x. \quad (7)$$

The resulting Markov chain does not have a steady-state distribution (it drifts towards a longer and longer queue).

We are interested mainly in situations where λ is arbitrarily close to 1, and thus only the first scenario is of interest.

B. Summary of multi-user results with centralized controller

For $b \leq \frac{\lambda}{1-\lambda}$, the optimal policy is to service an arbitrary non-empty queue (we can always to choose to service the longest queue so that the strategy is more robust when implemented in practice). As it turns out, the average discounted cost-to-go $V(x, y)$ assuming queue lengths x, y, is only a function of $x + y$ and thus, we can study the reduced one-dimensional problem. The resulting (reduced) Markov chain has a steady-state distribution (assuming $p < 1/2$) with

$$q(i) = \begin{cases} 1 - 2p & i = 0 \\ (1 - 2p)\frac{p(2-p)}{(1-p)^2} & i = 1 \\ (1 - 2p)\frac{p^{2i-2}}{(1-p)^{2i}} & i \geq 2. \end{cases} \tag{8}$$

As in the case of one queue, for $b \geq \frac{\lambda}{1-\lambda}$, the optimal policy is to not transmit at all. The resulting Markov chain does not have a steady-state distribution (it drifts towards longer and longer total queue size).

IV. The two-user decentralized control problem

We now return to the original problem where the two users do not communicate and thus they do not have full information about each-others queue lengths. We solve this problem in several steps. First we will show that the users strategies $f_t^i(x^{i,t}, y^t)$ can be restricted without loosing optimality. Then we will look at the problem from the viewpoint of a designer and provide a sequential decomposition and thus a solution in the form of a fixed-point equation. Finally, we will study the properties of this solution and try to simplify it to arrive at either a closed form solution or one that can be obtained numerically.

In the following we make use of the notation $\sigma_t = (\sigma_t^1, \sigma_t^2)$, where σ_t^i can be any of the variables defined earlier for user i at time t. We also denote by $\Delta(S)$ the space of probability mass functions over the discrete (possibly countably infinite) set S. The set of non-negative integers is denoted by N.

One of the difficulties with the general strategies $u_t^i = f_t^i(x^{i,t}, y^t)$ is that their domain increases with time. This complicates the solution of the problem in the infinite horizon and also requires essentially infinite memory at the transmitters if the optimal solution is to be implemented. In the following we show that we can restrict ourselves to strategies with finite domains without loss of optimality.

Assume user 2 employs a fixed strategy $f^{2*} = (f_t^{2*})_{t=1}^{\infty}$. Define the process $(z_t)_{t=1}^{\infty}$ with $z_1 = x_1^1$, and $z_t \stackrel{\text{def}}{=} (x_t^1, y^t)$ for $t \geq 2$. The following is true.

Lemma 1. *$(z_t)_{t=1}^{\infty}$ is a Markov process conditioned on u_t^1, i.e.,*

$$P(z_{t+1}|z^t, u^{1,t}) = P(z_{t+1}|z_t, u_t^1). \tag{9}$$

Furthermore, the average instantaneous cost can be expressed as

$$E\{c(x_t^1, x_t^2, u_t^1, u_t^2)\} = E\{\hat{c}(z_t, u_t^1)\}, \tag{10}$$

for some function $\hat{c}(\cdot)$.

Proof: The first part of the proof is very similar to the one in [14, Th. 1]. Using the fact that y_t is in one-to-one correspondence with the pair (u_{t-1}^1, u_{t-1}^2), we can write

$$\begin{aligned} P(z_{t+1}|z^t, u^{1,t}) &= P(x_{t+1}^1, y^{t+1}|x^{1,t}, y^t, u^{1,t}) && \text{(11a)} \\ &= P(x_{t+1}^1|x^{1,t}, y^{t+1}, u^{1,t})P(y^{t+1}|x^{1,t}, y^t, u^{1,t}) && \text{(11b)} \\ &= P(x_{t+1}^1|x^{1,t}, u^{1,t}, u^{2,t})P(u^{1,t}, u^{2,t}|x^{1,t}, u^{1,t}, u^{2,t-1}). && \text{(11c)} \end{aligned}$$

Due to (1), and the fact that the quantities $x^{1,t}$, $u^{1,t}$, $u^{2,t}$ do not depend on a_t^1, the first factor becomes

$$P(x_{t+1}^1|x^{1,t}, u^{1,t}, u^{2,t}) = P(x_{t+1}^1|x_t^1, u_t^1, u_t^2) \tag{12}$$

Similarly, recognizing that $u_t^2 = f_t^{2*}(x^{2,t}, y^t) = f_t^{2*}(x^{2,t}, u^{1,t-1}, u^{2,t-1}) = f_t^{2*}(x^{2,t}, u^{t-1})$; the fact that $x^{2,t}$ is only a function of u^{t-1} and the primitive random variables x_1^2, $a^{2,t-1}$; the fact that $x^{1,t}$ is only a function of u^{t-1} and the primitive random variables x_1^1, $a^{1,t-1}$; and the fact that $u_t^1 = f_t^1(x^{1,t}, u^{1,t-1}, u^{2,t-1}) = f_t^1(x^{1,t}, u^{t-1})$, we can deduce that the second factor in (11c) can be written as

$$P(u^{1,t}, u^{2,t}|x^{1,t}, u^{1,t}, u^{2,t-1}) = P(u^{1,t}, u^{2,t}|u^{1,t}, u^{2,t-1}). \tag{13}$$

This part of the proof is completed by repeating the same derivations for $P(z_{t+1}|z_t, u_t^1)$, and noticing that the two quantities are equal.

Regarding the average instantaneous cost we have

$$\mathrm{E}\{c(x_t^1, x_t^2, u_t^1, u_t^2)\} = \mathrm{E}\{\mathrm{E}\{c(x_t^1, x_t^2, u_t^1, u_t^2)|z_t, u_t^1\}\}. \tag{14}$$

We now consider the inner expectation

$$\begin{aligned}
\mathrm{E}\{c(x_t^1, x_t^2, u_t^1, u_t^2)|z_t, u_t^1\} &= \mathrm{E}\{c(x_t^1, x_t^2, u_t^1, f_t^{2*}(x^{2,t}, u^{t-1})|x_t^1, u^{t-1}, u_t^1\} && \text{(15a)}\\
&= \sum_{x^{2,t}} c(x_t^1, x_t^2, u_t^1, f_t^{2*}(x^{2,t}, u^{t-1}))P(x^{2,t}|x_t^1, u^{t-1}, u_t^1) && \text{(15b)}\\
&= \sum_{x^{2,t}} c(x_t^1, x_t^2, u_t^1, f_t^{2*}(x^{2,t}, u^{t-1}))P(x^{2,t}|u^{t-1}) && \text{(15c)}\\
&= \hat{c}(x_t^1, u^{t-1}, u_t^1), && \text{(15d)}
\end{aligned}$$

where (15c) is derived observing that $x^{2,t}$ is only a function of the primitive random variables x_1^2 and $a^{2,t-1}$ when conditioned on u^{t-1}. ■

The implication of the above Lemma is that for a fixed strategy of user 2 (and thus for user 2's optimal strategy), the problem faced by user 1 is an appropriate MDP, thus he can restrict his strategies to $u_t^1 = f_t^1(x^{1,t}, y^t) = f_t^1(x_t^1, y^t)$ without loss of optimality. The argument can be repeated for user 2, and thus in the following we restrict ourselves to strategies of the form $u_t^i = f_t^i(x_t^i, y^t)$. Although these strategies do not require storage of the entire sequence of the queue lengths, they do depend on the entire sequence of the feedback signals y^t. In addition, identifying this problem as an appropriate MDP from the viewpoint of user 1 is not helpful in deriving globally optimal strategies (recall that the cost function $\hat{c}(\cdot)$ as well as the evolution of z_t depends implicitly on f^{2*}).

To address this problem we now consider the evolution of the system from the perspective of an agent who has access to the common information y^t of the two users [14]–[17]. In particular, we consider a fictitious agent observing y_t at time t. Each user's action $u_t^i = f_t^i(x_t^i, y^t)$ can be though of as the result of the common agent first determining pre-encoding functions $w_t^i = g_t^i(y^t)$ from y^t, where $w_t^i :\ \mathrm{N} \to \{0, 1\}$, and then each user evaluating its corresponding function at x_t^i, i.e., $u_t^i = w_t^i(x_t) = g_t^i(y^t)(x_t)$. We will now show that from the viewpoint of the fictitious agent, the original problem can be viewed as an appropriately defined POMDP problem. Define the process $(r_t)_{t=1}^{\infty}$ with $r_1 = x_1$ and $r_t = (x_t, y_t) = (x_t, u_{t-1})$ for $t \geq 2$, where $x_t \stackrel{\text{def}}{=} (x_t^1, x_t^2)$.

Lemma 2. *$(r_t)_{t=1}^{\infty}$ is a Markov process conditioned on $w_t \stackrel{\text{def}}{=} (w_t^1, w_t^2)$, i.e.,*

$$P(r_{t+1}|r^t, w^t) = P(r_{t+1}|r_t, w_t). \tag{16}$$

Furthermore, the instantaneous cost $c(\cdot)$ is only a function of x_t and w_t.

Proof: We have

$$P(r_{t+1}|r^t, w^t) = P(x_{t+1}, y_{t+1}|x^t, y^t, w^t) \tag{17a}$$

$$= P(x_{t+1}, u_t|x^t, u^{t-1}, w^t) \tag{17b}$$

$$= P(x_{t+1}|x^t, u^t, w^t)P(u_t|x^t, u^{t-1}, w^t) \tag{17c}$$

$$= P(x_{t+1}|x_t, u_t)P(u_t|x_t, w_t) \tag{17d}$$

$$= P(x_{t+1}|x_t, u_t)\delta(u_t^1 - w_t^1(x_t^1))\delta(u_t^2 - w_t^2(x_t^2)) \tag{17e}$$

$$= P(r_{t+1}|r_t, w_t), \tag{17f}$$

where the first factor in (17d) is due to (1), and the fact that the quantities x^t, u^t, w^t do not depend on a_t^i, and the second factor is due to that fact that given w_t and x_t, both outputs are determined.

Furthermore, the instantaneous cost can be written as

$$c(x_t^1, x_t^2, u_t^1, u_t^2) = x_t^1 + x_t^2 + bu_t^1 + bu_t^2 \tag{18a}$$

$$= x_t^1 + x_t^2 + bw_t^1(x_t^1) + bw_t^2(x_t^2) \tag{18b}$$

$$= \tilde{c}(x_t, w_t) \tag{18c}$$

■

Based on the above Lemma, from the point of view of the fictitious agent the original problem is a POMDP with state $r_t = (x_t, u_{t-1})$, observation $y_t = u_{t-1}$, actions w_t and instantaneous cost $\tilde{c}(\cdot)$. This problem can be solved using standard techniques in POMDPs. In particular, one can define the information state $\pi_t \in \Delta(\mathrm{N} \times \mathrm{N})$, where

$$\pi_t(x_t) \stackrel{\text{def}}{=} P(x_t|u^{t-1}, w^{t-1}), \tag{19}$$

and solve the following fixed-point equation for $w^*(\pi)$

$$V(\pi) = \inf_w \{\sum_x \tilde{c}(x, w)\pi(x) + \lambda \sum_u V(T(w, u)\pi)\}, \tag{20}$$

where $T(w, u)$ is a linear operator on $\Delta(\mathrm{N} \times \mathrm{N})$ determined by the problem primitives.

In the following we show that due to the special structure of the problem, a solution can be obtained based on the fixed point of an equation over $\Delta(\mathrm{N}) \times \Delta(\mathrm{N})$, resulting in significant computational savings compared to (20). Towards this goal, we define the sequences $(\xi_t^i)_{t=1}^\infty$ for $i = 1, 2$, with $\xi_t^i \in \Delta(\mathrm{N})$, where

$\xi_1^i(x_1^i) \stackrel{\text{def}}{=} P(x_1^i) = Q(x_1^i)$ and $\xi_t^i(x_t^i) \stackrel{\text{def}}{=} P(x_t^i|u^{t-1}, w^{t-1})$. The following lemma shows that $\xi_t \stackrel{\text{def}}{=} (\xi_t^1, \xi_t^2)$ can be considered an information state for the problem under consideration.

Lemma 3. *There exist functions $\Phi^i(\cdot)$, $C(\cdot)$ such that $\xi_{t+1}^i = \Phi^i(\xi_t^i, w_t^i, u_t)$ and $E\{c(x_t, u_t)|u^{t-1}, w^t\} = C(\xi_t^1, w_t^1) + C(\xi_t^2, w_t^2)$ for all $t \geq 1$. Furthermore, $\Phi^1(\xi_t^1, w_t^1, u_t^1, u_t^2) = \Phi^2(\xi_t^1, w_t^1, u_t^2, u_t^1)$.*

Proof: Consider ξ_{t+1}^1.

$$\xi_{t+1}^1(x_{t+1}^1) = P(x_{t+1}^1|u^t, w^t) \tag{21a}$$

$$= \sum_{x_t^1, x_t^2} P(x_{t+1}^1|x_t^1, x_t^2, u^t, w^t) P(x_t^1, x_t^2|u^t, w^t) \tag{21b}$$

$$= \sum_{x_t^1, x_t^2} P(x_{t+1}^1|x_t^1, x_t^2, w_t) P(x_t^1, x_t^2|u^t, w^t), \tag{21c}$$

since conditioned on $w_t = (w_t^1, w_t^2)$ and x_t^1, x_t^2, the value of $u_t = (u_t^1, u_t^2)$ is exactly determined, and making use of (1). The second factor can be written as

$$P(x_t^1, x_t^2|u^t, w^t) = \frac{P(x_t^1, x_t^2, u_t|u^{t-1}, w^t)}{P(u_t|u^{t-1}, w^t)} \tag{22a}$$

$$= \frac{P(u_t|x_t^1, x_t^2, u^{t-1}, w^t) P(x_t^1, x_t^2|u^{t-1}, w^t)}{P(u_t|u^{t-1}, w^t)} \tag{22b}$$

$$= \frac{P(u_t|x_t^1, x_t^2, w_t) P(x_t^1, x_t^2|u^{t-1}, w^t)}{P(u_t|u^{t-1}, w^t)} \tag{22c}$$

$$= \frac{\delta(u_t^1 - w_t^1(x_t^1))\delta(u_t^2 - w_t^2(x_t^2)) P(x_t^1|u^{t-1}, w^{t-1}) P(x_t^2|u^{t-1}, w^{t-1})}{P(u_t|u^{t-1}, w^t)} \tag{22d}$$

$$= \frac{\delta(u_t^1 - w_t^1(x_t^1))\delta(u_t^2 - w_t^2(x_t^2))\xi_t^1(x_t^1)\xi_t^2(x_t^2)}{\sum_{x_t^1, x_t^2} \delta(u_t^1 - w_t^1(x_t^1))\delta(u_t^2 - w_t^2(x_t^2))\xi_t^1(x_t^1)\xi_t^2(x_t^2)}, \tag{22e}$$

where the $\delta(\cdot)$ function appears for the reason mentioned above; x_t^1 and x_t^2 are independent conditioned on u^{t-1} by making recursive use of (1); and w_t is eliminated from the conditioning since it is a deterministic function of u^{t-1}.

Combining the above equations we get

$$\xi^1_{t+1}(x^1_{t+1}) = \sum_{x^1_t, x^2_t} P(x^1_{t+1}|x^1_t, x^2_t, w_t) \frac{\delta(u^1_t - w^1_t(x^1_t))\delta(u^2_t - w^2_t(x^2_t))\xi^1_t(x^1_t)\xi^2_t(x^2_t)}{\sum_{x^1_t, x^2_t} \delta(u^1_t - w^1_t(x^1_t))\delta(u^2_t - w^2_t(x^2_t))\xi^1_t(x^1_t)\xi^2_t(x^2_t)} \tag{23a}$$

$$= \sum_{x^1_t, x^2_t} P(x^1_{t+1}|x^1_t, u_t) \frac{\delta(u^1_t - w^1_t(x^1_t))\delta(u^2_t - w^2_t(x^2_t))\xi^1_t(x^1_t)\xi^2_t(x^2_t)}{\sum_{x^1_t, x^2_t} \delta(u^1_t - w^1_t(x^1_t))\delta(u^2_t - w^2_t(x^2_t))\xi^1_t(x^1_t)\xi^2_t(x^2_t)} \tag{23b}$$

$$= \sum_{x^1_t} P(x^1_{t+1}|x^1_t, u_t) \frac{\delta(u^1_t - w^1_t(x^1_t))\xi^1_t(x^1_t)\sum_{x^2_t}\delta(u^2_t - w^2_t(x^2_t))\xi^2_t(x^2_t)}{\sum_{x^1_t, x^2_t} \delta(u^1_t - w^1_t(x^1_t))\delta(u^2_t - w^2_t(x^2_t))\xi^1_t(x^1_t)\xi^2_t(x^2_t)} \tag{23c}$$

$$= \sum_{x^1_t} P(x^1_{t+1}|x^1_t, u_t) \frac{\delta(u^1_t - w^1_t(x^1_t))\xi^1_t(x^1_t)}{\sum_{x^1_t} \delta(u^1_t - w^1_t(x^1_t))\xi^1_t(x^1_t)}, \tag{23d}$$

where (23b) is due to (1) and the presence of the $\delta(\cdot)$ functions. Observe that $\Phi^1(\xi^1_t, w^1_t, u^1_t, u^2_t) = \Phi^2(\xi^1_t, w^1_t, u^2_t, u^1_t)$, since $P(x^1_{t+1}|x^1_t, u^1_t, u^2_t) = P(x^2_{t+1}|x^2_t, u^2_t, u^1_t)$ while the rest of the terms in (23d) are completely symmetric with respect to the users. Thus $\xi^1_{t+1} = \Phi^1(\xi^1_t, w^1_t, u_t)$.

For the second part we observe that

$$\mathrm{E}\{c(x_t, u_t)|u^{t-1}, w^t\} = \sum_{x_t}[x^1_t + x^2_t + bw^1_t(x^1_t) + bw^2_t(x^2_t)]P(x_t|u^{t-1}, w^t) \tag{24a}$$

$$= \sum_{x_t}[x^1_t + x^2_t + bw^1_t(x^1_t) + bw^2_t(x^2_t)]P(x^1_t|u^{t-1}, w^{t-1})P(x^2_t|u^{t-1}, w^{t-1}) \tag{24b}$$

$$= \sum_{x^1_t}[x^1_t + bw^1_t(x^1_t)]\xi^1_t(x^1_t) + \sum_{x^2_t}[x^2_t + bw^2_t(x^2_t)]\xi^2_t(x^2_t) \tag{24c}$$

$$= C(\xi^1_t, w^1_t) + C(\xi^2_t, w^2_t), \tag{24d}$$

where we have used similar arguments as above. ■

We are now ready to state the basic result of the paper. Consider the process $(\xi_t)^\infty_{t=1} = (\xi^1_t, \xi^2_t)^\infty_{t=1}$. The following theorem shows that from the viewpoint of a designer that designs the optimal policy for the fictitious agent, ξ_t is a sufficient statistic for control of the original problem.

Proposition 1. *The process $(\xi_t)^\infty_{t=1}$ is a MDP with actions w_t. Furthermore, the average instantaneous cost of the original problem can be written as $E\{c(x_t, u_t)\} = E\{C(\xi^1_t, w^1_t) + C(\xi^2_t, w^2_t)\}$. The optimal policies for the original problem can be found by solving the fixed-point equation*

$$V(\xi^1, \xi^2) = \inf_{w^1, w^2}\{C(\xi^1, w^1) + C(\xi^2, w^2) + \lambda \sum_{u^1, u^2} V(T(w^1, u^1, u^2)\xi^1, T(w^2, u^2, u^1)\xi^2)\}, \tag{25}$$

where $\xi^i \in \Delta(N)$, $w^i : N \to \{0, 1\}$, and $T(\cdot)$ is a linear operator on $\Delta(N)$ defined as a scaled version *of the function* $\Phi^1(\cdot)$*, i.e.,*

$$T(w^1, u^1, u^2)(x', x) = P(x^1_{t+1} = x' | x^1_t = x, u_t = (u^1, u^2))\delta(u^1 - w^1(x)). \tag{26a}$$

Proof: The first part of the proposition follows from Lemma 3 and the fact that conditioned on u^{t-1}, $x^{1,t}$ and $x^{2,t}$ are independent. Once this is established, the rest follows from well known facts about MDPs (see for instance [18]). The details are omitted due to space limitations. ■

Several comments are in order. First, suppose the above fixed point equation has been solved and the optimal solutions are denoted by $w^{i*} = w^{i*}(\xi^1, \xi^2)$, $i = 1, 2$. The exact on-line implementation of the optimal policy by each user is as follows. Each user follows the evolution of the states (ξ^1_t, ξ^2_t) using (Q, Q) as the initial state; $w^{i*} = w^{i*}(\xi^1, \xi^2)$ to evaluate the optimal pre-encoding functions; $u^i_t = w^{i*}(\xi^1_t, \xi^2_t)(x^i_t)$ to decide whether to transmit or not; and after receiving the feedback $y_{t+1} = u_t$, Lemma 3 to evaluate the next state $(\xi^1_{t+1}, \xi^2_{t+1})$ from $\xi^i_{t+1} = \Phi^i(\xi^i_t, w^{i*}_t, u_t)$. Observe that the above updates will evolve differently for the two users, due to the presence of x^i_t in the evaluation of u^i_t.

Second, due to the time homogeneity of the problem in Proposition 1, the optimal solution will be of the form $w^{i*} = w^{i*}(\xi^1, \xi^2)$, $i = 1, 2$. This however does not mean that the optimal policies is time invariant. To see that consider the evolution of the information state for user 1.

$$(\xi^1_1, \xi^2_1) = (Q, Q) \overset{w^1_1 = w^{1*}(\xi^1_1, \xi^2_1),\ w^2_1 = w^{2*}(\xi^1_1, \xi^2_1)}{\longrightarrow} (\xi^1_2, \xi^2_2) \overset{w^1_2 = w^{1*}(\xi^1_2, \xi^2_2),\ w^2_2 = w^{2*}(\xi^1_2, \xi^2_2)}{\longrightarrow} (\xi^1_3, \xi^2_3) \longrightarrow \cdots \tag{27}$$

Third, even with the above simplifications, the solution is given by a functional optimization problem (recall that w^i are functions $\mathrm{N} \to \{0, 1\}$). Motivated by this observation and by the solution of the centralized version of the problem, one would like to show that the optimal policy is a threshold policy, i.e., $w^{i*}(\xi^1, \xi^2)$ is such that $w^{i*}(\xi^1, \xi^2)(x) = 0$ for all $x < x_{th}(\xi^1, \xi^2)$ and $w^*(\xi^1, \xi^2)(x) = 1$ for all $x \geq x_{th}(\xi^1, \xi^2)$, for some threshold $x_{th}(\xi^1, \xi^2)$. If a property like this holds, then the above fixed point equation reduces from a functional optimization problem to a parameter optimization over the parameter x_{th}.

V. Conclusions

The problem of optimal tradeoff between energy and delay in MAC channels have been studied as a decentralized stochastic control problem. We have shown that the optimal policies do not have an increasing domain and can be obtained as the solution of a fixed-point equation. The complexity of solving such an equation is still high (functional optimization); however our experience and relevant literature indicates that this difficulty is inherent in all decentralized stochastic control problems. One fruitful direction towards reducing the computational burden both for the off-line design and the on-line operation, is to show that the optimal policies for each user are threshold policies with respect to their queue lengths.

REFERENCES

[1] E. A. Haroutunian, “Lower bound for the error probability of multiple-access channels,” *Problemy Peredachi Informatsii*, vol. 11, no. 3, pp. 23–36, June 1975.

[2] R. G. Gallager, “A perspective on multiaccess channels,” *IEEE Trans. Information Theory*, vol. 31, no. 2, pp. 124–142, Mar. 1985.

[3] Y. Liu and B. L. Hughes, “A new universal random coding bound for the multiple access channel,” *IEEE Trans. Information Theory*, vol. 42, pp. 376–386, 1996.

[4] L. Weng, S. Pradhan, and A. Anastasopoulos, “Error exponent regions for Gaussian broadcast and multiple access channels,” *IEEE Trans. Information Theory*, vol. 8, pp. 2919 – 2942, July 2008.

[5] A. Nazari, S. Pradhan, and A. Anastasopoulos, “A new sphere-packing bound for maximal error exponent for multiple-access channels,” in *Proc. International Symposium on Information Theory*, Toronto, Canada, July 2008, pp. 1513–1517.

[6] A. Nazari, A. Anastasopoulos, and S. Pradhan, “A new universal random-coding bound for average probability error exponent for multiple-access channels,” in *Proc. Conference on Information Sciences and Systems (CISS)*, 2009, (Submitted).

[7] J. Cao and E. Yeh, “Power-delay tradeoff analysis for communication over fading channels with feedback,” in *Proc. International Symposium on Information Theory*, Toronto, ON, July 2008, pp. 614 – 618.

[8] N. Ehsan and T. Javidi, “Delay optimal transmission policy in a wireless multi-access channel,” *IEEE Trans. Information Theory*, vol. 54, no. 8, pp. 3745–3751, Aug. 2008.

[9] H. S. Witsenhausen, “Separation of estimation and control for discrete time systems,” *Proc. IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.

[10] P. Varaiya and J. Walrand, “On delayed sharing patterns,” *IEEE Trans. Automatic Control*, vol. 23, no. 3, pp. 443–445, June 1978.

[11] A. Mahajan and D. Teneketzis, “On jointly optimal encoding, decoding and memory update for noisy real-time communication systems,” *submitted to IEEE Trans. Inf. Theory*, Jan. 2006, available as Control Group Report CGR-06-03, Department of EECS, University of Michigan, Ann Arbor, MI 48109-2122.

[12] ——, “Optimal performance of feedback control systems with limited communication over noisy channels,” *submitted to SIAM Journal of Control and Optimization*, Dec. 2006, available as Control Group Report CGR-06-07, Department of EECS, University of Michigan, Ann Arbor, MI 48109-2122.

[13] ——, “On the design of globally optimal communication strategies for real-time communcation systems with noisy feedback,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 580–595, May 2008.

[14] A. Nayyar and D. Teneketzis, “On jointly optimal real-time encoding and decoding strategies in multi-terminal communication systems,” in *Proc. IEEE Conf. on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 1620–1627.

[15] A. Mahajan, A. Nayyar, and D. Teneketzis, “Identifying tractable decentralized problems on the basis of information structures,” in *Proc. Allerton Conf. Commun., Control, Comp.*, Monticello, Illinois, Sept. 2008.

[16] J. C. Walrand and P. Varaiya, “Optimal causal coding—decoding problems,” *IEEE Trans. Inf. Theory*, vol. 29, no. 6, pp. 814–820, Nov. 1983.

[17] ——, “Causal coding and control of markov chains,” *System and Control Letters*, vol. 3, pp. 189–192, 1983.

[18] P. R. Kumar and P. Varaiya, *Stochastic systems: estimation, identification, and adaptive control*. Englewood Cliffs, NJ: Prentice-Hall, 1986.