Continuous-time Optimal Rate Control for Delay Constrained Data Transmission

Murtaza Zafer and Eytan Modiano

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
email:{murtaza,modiano}@mit.edu

Abstract

We consider optimal rate control for delay constrained data transmission, over time varying channels, with the objective of minimizing the transmission energy expenditure. Energy efficiency is achieved by spreading the data transmission over time to exploit the convexity of the power-rate functions and further, by opportunistically adapting the transmission rate to the channel variations. Our system model consists of a buffer with B units of data that must be transmitted by a finite deadline T with minimum energy over a time-varying channel. Using a continuous-time stochastic control formulation, we obtain simple closed form expressions for the optimal transmission rate that minimizes the energy expenditure and meets the deadline constraint. Finally, simulation results compare the gains achieved by the optimal rate adaptation with non-adaptive policies.

Index Terms

Energy efficiency, Delay constraints, Wireless fading channel, Rate control, Quality of Service.

I. Introduction

Energy efficiency is an important concern in the design of wireless systems and is an active area of research in wireless data, ad-hoc and sensor networks. As transmission energy constitutes a significant part of the energy cost, minimizing this expenditure can lead to significant energy savings. Furthermore, with maximum delay constraints on the data, a transmission policy must also ensure that these delay requirements are satisfied. There are numerous practical scenarios that involve an optimization with the above objectives. For example, in wireless data systems, applications such as Voice over IP (VoIP), video and multimedia streaming, delay constrained file transfers etc. generate data that have strict delay constraints and must be transmitted over a time-varying channel. As energy is limited on mobile devices, this data must be transmitted in an energy efficient manner. Similarly, in sensor networks, time critical sensing applications require that the collected data must be transmitted back to a central entity within a specific time duration. Clearly, minimizing the energy cost here directly translates into an increased lifetime of the sensor.

Our focus in this work is to employ adaptive rate control, subject to deadline constraints, to minimize transmission energy costs. We consider a transmitter, such as those deployed in the 1xEV-DO system [7], whose transmission rate can be controlled over time. With rate control energy efficiency can be achieved in the following two ways. First, for many coding schemes its well known that transmitting data at a low rate and over a longer duration has less energy cost as compared to a high rate transmission; which means that spreading the data transmission over time is energy efficient. Mathematically, the above observation translates into a convex dependence of power on the transmission rate and such a model has been widely studied in the literature [1]–[3], [5]. Second, the wireless channel is time-varying and clearly

This work was supported by NSF ITR grant CCR-0325401, by DARPA/AFOSR through the University of Illinois grant no. F49620-02-1-0325 and by NASA Space Communication Project grant number NAG3-2835.

its more energy efficient to transmit under good channel conditions. Using channel estimation, one can opportunistically exploit these variations by adapting the transmission rate over the channel states.

Rate control has been studied earlier in various other contexts that involve average throughput [8], [9], average delay [1], [10] and packet loss probability [11]. These metrics are measured over an infinite time horizon and do not directly translate into any Quality of Service (QoS) guarantees over a finite time interval. Incorporating strict deadlines on the data and considering a finite-time horizon optimization makes the problem challenging. Recent work in this direction includes [2]–[6]. The work in [2] and its various extensions studied offline formulations under non-causal knowledge of future channel states; they presented an iterative algorithm that eventually converges to the offline optimal solution. The authors in [3] considered a problem that involved a fixed amount of data to be transmitted by a deadline with minimum energy and presented a numerical dynamic programming solution. The works in [4]–[6] considered a time-invariant channel formulation and in particular our earlier work in [4] used a calculus approach to obtain minimum energy transmission policies with general arrival curves and QoS constraints.

A common approach towards formulating dynamic control problems is based on applying discrete dynamic programming (DP). However, it can be easily seen that the problem we consider is mathematically intractable to solve in closed form using discrete DP. Any numerical solution would be computationally intensive due to the large state space and would not provide much insights into the problem. Hence, we adopt a different approach based on a continuous time formulation, the results of which are summarized below.

Results Summary: We consider the problem of transmitting B units of data by deadline T over a time-varying channel. The channel state evolution is modelled as a general Markov stochastic process and the power-rate functions are taken as the class of monomial functions. Using a continuous-time formulation, we obtain closed form expressions for the optimal transmission rate as a function of the present time, the data left in the queue and the present channel state. These results are presented in Theorems I and II. As an example, we then consider the Gilbert-Elliott channel model and present simulation results that provide comparisons of the energy costs and the gains achieved using optimal rate adaptation. Finally, in Section V we conclude the paper pointing out various avenues for future extensions.

II. System Model

A. Transmission Model

As discussed in the introduction, we consider transmission schemes where the rate can be controlled over time. Such control can be achieved in various ways by changing the channel encoding, the modulation scheme or the symbol duration. Maintaining a constant bit error probability, changing the transmission rate requires changing the transmission power over time. The exact power-rate relationship that we consider in this work is given below in (1). To motivate such a dependence, we briefly describe the physical layer assumptions in the context of the widely used block transmission model [1].

Consider a block transmission model with flat-fading and an additive white Gaussian noise channel. Data is transmitted in block codes of size N, where N is the number of channel symbols transmitted per block. Taking $N \gg 1$, one can achieve reliable communication at some rate r by varying the transmission power over each such block. For many practical systems, N is of the order of a few 1000 and the transmission duration of a block is of the order of 1 msec, [7]. In comparison, the delay requirements on the arriving packets are usually of the order of 100's of msec; for example, VoIP packets have delay requirements around $150-200$ msec. Under such scenarios, one can view the transmission block durations as negligible in comparison to the time scales of interest, thus, making a continuous time rate control model appropriate for the system.

To obtain the power-rate relationship we proceed as follows. Let h_t denote the channel gain, P_t the transmitted signal power and P_t^r the received signal power at time t. Now, with most encoding schemes the reliable rate of communication, r_t, is a concave function of the received signal power. Inverting this relationship, it implies that the required received signal power to achieve a certain rate is convex in the rate, i.e. $P_t^r = g(r_t)$, where $g(.)$ is a non-negative convex increasing function for $r_t \geq 0$. Since the received signal power is given as $P_t^r = |h_t|^2 P_t$, the required transmission power to achieve rate r_t is given by,

$$P_t = \frac{g(r_t)}{c(t)} \tag{1}$$

where $c(t) \triangleq |h_t|^2$. The quantity $c(t)$ will be referred to as the *channel state* at time t. Its value at time t is assumed known either through prediction or direct measurement but evolves stochastically in the future. Equation (1) represents a general power-rate relationship and has been widely studied earlier either in generality or in a special form in [1], [2], [5], [6], [10]. Examples of schemes where (1) applies include the optimal channel coding over the AWGN channel for which $P_t = \frac{N_0 W (2^{2r}-1)}{c(t)}$ (Shannon capacity formula) and the variable rate trellis coded M-QAM scheme [1], for which the transmitted power can be approximated as $P_t = \frac{\sigma^2}{c(t)} \left(2^{\frac{r+2\chi}{N}} \right) K_c$ (where σ, χ, K_c, N are appropriate constants).

In this paper, we consider the class of monomial functions for $g(r)$, namely, $g(r) = kr^n,\ n > 1$ ($n \in \mathbb{R}$). Without loss of generality we take the constant $k = 1$, since, any other value of k simply scales the problem without affecting the results. The choice of $g(r) = r^n$ is motivated by the fact that monomials can be used as a first approximation to general functions. In fact, in the low to moderate signal-to-noise regime, $g(r) = r^2$ is a good approximation to many transmission schemes.

B. Channel Model

We consider a general continuous time Markov model for the channel state evolution. Channel models based on Markov processes have been actively studied in the past and there is a substantial literature on modelling various fading scenarios [12]. Denote the channel state process as, $C(t)$, and let $c(t) = c$, be a particular realization at time t. Starting from state c, let $\mathcal{J}_c$ be the set of all states ($\neq c$) to which the channel can transition when the state changes. The channel transition time from state c to $\tilde{c}$ is Exponentially distributed with rate $\lambda_{c\tilde{c}}$. The sum transition rate at which the channel jumps out of state c is then, $\lambda_c = \sum_{\tilde{c} \in \mathcal{J}_c} \lambda_{c\tilde{c}}$. Clearly, the expected time $C(t)$ spends in state c is $1/\lambda_c$ and one can view $1/\lambda_c$ as the coherence time of the channel in state c.

For a general Markov model, the sum transition rate λ_c can vary over the states. To simplify the process description, one can make the sum transition rate the same for all the states. This is done by speeding up the process and introducing self-transitions to the same state such that the transition rate out of state c remains unaltered while the remaining transitions are self-transitions. Such a simplification does not lose any generality in the Markov process as it yields a stochastically identical scenario. Mathematically, let us define $\lambda \triangleq \sup_c \lambda_c$ and a random variable, $Z(c)$, as,

$$Z(c) = \begin{cases} \tilde{c}/c, & \text{with prob. } \lambda_{c\tilde{c}}/\lambda,\ \tilde{c} \in \mathcal{J}_c \\ 1, & \text{with prob. } 1 - \lambda_c/\lambda \end{cases} \tag{2}$$

The process can now be visualized to evolve in the following simple way. Given a channel state c, there is an Exponentially distributed time duration with rate λ after which the channel state changes. The new state including self-transitions is a random variable which is given as $C = Z(c)c$. Clearly, from (2) the transition rate to state $\tilde{c} \in \mathcal{J}_c$ is unchanged at $\lambda_{c\tilde{c}}$, whereas with rate $\lambda - \lambda_c$ there are indistinguishable self-transitions. As such self-transitions have no meaning in a continuous-time process since they are indistinguishable over any sample path. However, the above representation is used for notational convenience and does not change the underlying process[1]. As an example, consider the Gilbert-Elliott channel model that further

[1]The technical assumptions in the model are as follows. The channel state space, $\mathcal{C}$, is a countable space (it could be infinite), and $\mathcal{C} \subseteq \mathbb{R}^+$. The states $c = 0, \infty$ are excluded from $\mathcal{C}$ since each of this state leads to a singularity in (1). The set $\mathcal{J}_c, \forall c$, is a finite subset of $\mathcal{C}$. Transition rate $\lambda_c, \forall c$ is bounded which ensures that λ defined as the supremum is finite. For all c, the support of $Z(c)$ lies in $[z_l, z_h]$, where $0 < z_l \leq z_h < \infty$. This ensures that $C(t)$ does not hit 0 or ∞, a.s. (almost surely), over a finite time interval.

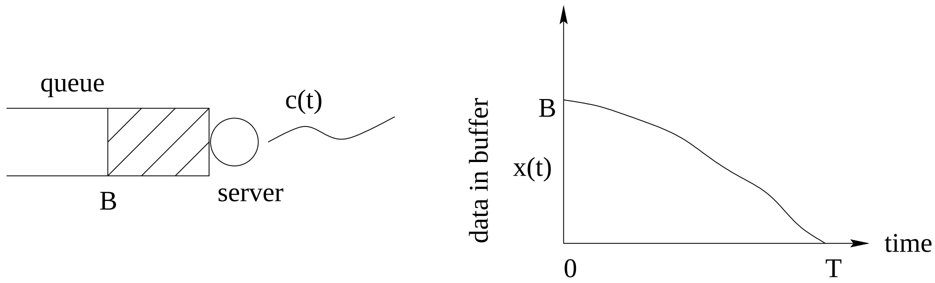

Fig. 1. BT-problem; B units of data to be transmitted by deadline T.

clarifies the notations used above.

Example: Consider the standard Gilbert-Elliott channel model [12] that has two states b and g denoting the bad and the good channel conditions respectively. Let the transition rate from the good to the bad state be λ_{gb} and from the bad to the good state be λ_{bg}. Let the ratio of the bad to the good channel quality be denoted as γ, i.e. $\gamma = c_b/c_g$. Using the earlier notations, $\lambda = \max(\lambda_{bg}, \lambda_{gb})$. For state c_g we have,

$$Z(c_g) = \begin{cases} \gamma, & \text{with prob. } \lambda_{gb}/\lambda \\ 1, & \text{with prob. } 1 - \lambda_{gb}/\lambda \end{cases} \tag{3}$$

To obtain $Z(c_b)$ simply replace γ with $1/\gamma$ and λ_{gb} with λ_{bg} in (3) above.

III. PROBLEM FORMULATION

We formulate the problem in the framework of continuous time optimal control theory. As mentioned earlier, the queue has B units of data that must be transmitted by deadline T, with minimum energy over a time varying channel. We refer to this as the "BT-problem" where the notation implies that the amount of data under consideration is B, and the deadline is T. The system is depicted in Figure 1.

A. Optimal Control Formulation

Consider the BT-problem and let $x(t)$ denote the amount of data left in the queue at time t. The system state can be described as (x, c, t), where the notation means that at time t, we have $x(t) = x$ and $c(t) = c$. The state space for the process (x, c, t) is $\Omega = [0, B] \times \mathcal{C} \times [0, T]$. Let $r(x, c, t)$ denote the chosen transmission rate for the corresponding system state (x, c, t). Since the underlying process is Markovian, it is sufficient to restrict attention to transmission policies that depend only on the present system state [15]. Clearly then, (x, c, t) is a Markov process.

Given a policy $r(x, c, t)$, let us now see how the system evolves over a particular sample path of the channel process. We are given $x(0) = B$ and $c(0) = c_0$. Until τ_1, where τ_1 is the first time instant after $t = 0$ at which the channel changes, the buffer is reduced at the rate $r(x(t), c_0, t)$. Hence, over the interval $[0, \tau_1]$, $x(t)$ satisfies the ordinary differential equation (ODE), $\frac{dx(t)}{dt} = -r(x(t), c_0, t)$. Equivalently,

$$x(t) = x(0) - \int_0^t r(x(s), c_0, s)ds \quad , \; t \in [0, \tau_1] \tag{4}$$

Now, starting from the new state $(x(\tau_1), c_1, \tau_1)$, the above procedure repeats until $t = T$ is reached.

A transmission policy, $r(x, c, t)$, is *admissible*, if it satisfies the following,

(a) $0 \leq r(x, c, t) < \infty$, (non-negativity)

(b) $r(x, c, t) = 0$, if $x = 0$ (no data left to transmit) and,

(c) $x(T) = 0$, a.s. (deadline constraint)

An additional technical requirement is that $r(x, c, t)$ be continuous and locally lipschitz in x (for $x > 0$) which ensures that $x(t)$ is the unique solution of the ODE presented earlier [16].

We now obtain the expected energy expenditure for an admissible transmission policy. Define a cost-to-go function, $J_r(x,c,t)$, as the expected energy cost under an admissible policy $r(\cdot)$ starting at time $t<T$ with the system state being (x,c,t). Then,

$$J_r(x,c,t) = E\left[\int_t^T \frac{1}{c(s)} g(r(x(s),c(s),s))ds\right] \tag{5}$$

where the above expectation is conditioned on $x(t)=x$, $c(t)=c$. Define a *value function*, $J(x,c,t)$, as the infimum of $J_r(x,c,t)$ over the set of admissible transmission policies.

$$J(x,c,t) = \inf_r J_r(x,c,t), \quad r(x,c,t) \text{ admissible} \tag{6}$$

The optimization problem, then, is to compute $J(x,c,t)$ and obtain a policy $r^*(x,c,t)$, that achieves it. In particular, the least expected energy expenditure over the entire interval, $[0,T]$, is given as, $J(B,c_0,0)$.

B. Optimality Conditions

A standard approach towards studying continuous time problems is to investigate their behavior over a small time interval. In the context of the BT-problem, this methodology applies as follows. Suppose that the system is in state (x,c,t). We first apply a transmission policy, $r(\cdot)$, in the small interval $[t,t+h]$ and thereafter, starting from $(x(t+h),c(t+h),t+h)$ we assume that the optimal policy is followed. By assumption, the energy cost is optimal over $[t+h,T]$, hence, investigating the system over $[t,t+h]$ would give conditions for the optimality of the chosen rate at time t. Since t is arbitrary, we obtain formal conditions for an optimal policy.

Following the above approach, consider $t\in[0,T)$ and a small interval $[t,t+h]$, where $t+h<T$. Then, clearly, the value function $J(x,c,t)$ satisfies,

$$J(x,c,t) = \min_{r(\cdot)}\left\{E\int_t^{t+h} \frac{1}{c(s)} g(r(x(s),c(s),s))ds + EJ(x_{t+h},c_{t+h},t+h)\right\} \tag{7}$$

where x_{t+h} is a short-hand notation for $x(t+h)$. The expression within the minimization bracket in (7) denotes the total cost with policy $r(\cdot)$ being followed over $[t,t+h]$ and the optimal policy thereafter. This cost must be clearly no more than the cost of applying the optimal policy from the starting state (x,c,t). Thus for an admissible policy $r(\cdot)$ we get,

$$J(x,c,t) \le E\int_t^{t+h} \frac{1}{c(s)} g(r(x(s),c(s),s))ds + \ E\left[J(x_{t+h},c_{t+h},t+h)\right] \tag{8}$$

$$E[J(x_{t+h},c_{t+h},t+h)] - J(x,c,t) + E\int_t^{t+h} \frac{1}{c(s)} g(r(x(s),c(s),s))ds \ge 0 \tag{9}$$

Now, divide (9) by h and take the limit $h\downarrow 0$. In the limit we have,

$$\frac{E\int_t^{t+h}\left(\frac{g(r(x_s,c_s,s))}{c_s}\right)ds}{h} \to g(r)/c \tag{10}$$

where r is the transmission rate at time t, i.e. $r=r(x,c,t)$. Define $\lim_{h\downarrow 0}\frac{EJ(x_{t+h},c_{t+h},t+h)-J(x,c,t)}{h} = A^r J(x,c,t)$, then in the limit (9) simplifies to,

$$A^r J(x,c,t) + \frac{1}{c} g(r) \ge 0 \tag{11}$$

The quantity $A^r J(x,c,t)$ is called the differential generator of the Markov process $(x(t),c(t))$ for policy $r(\cdot)$. Intuitively, the differential generator is a natural generalization of the ordinary time derivative for a function that depends on a stochastic process. An elaborate discussion on this topic can be found in [13]–[15]. For the process $(x(t), c(t))$, using the time evolution as in (4), the quantity $A^r J(x, c, t)$ can be evaluated as,

$$A^r J(x,c,t) = \frac{\partial J(x,c,t)}{\partial t} - r\frac{\partial J(x,c,t)}{\partial x} + \sum_{\tilde{c} \in \mathcal{J}_c} \lambda_{c\tilde{c}}[J(x,\tilde{c},t) - J(x,c,t)] \tag{12}$$

Furthermore, using the definition of $Z(c)$ as in (2) it can be simplified to,

$$A^r J(x,c,t) = \frac{\partial J(x,c,t)}{\partial t} - r\frac{\partial J(x,c,t)}{\partial x} + \lambda(E_z[J(x,Z(c)c,t)] - J(x,c,t)) \tag{13}$$

where E_z is the expectation with respect to the Z variable.

In the above steps from (8)-(11) if policy $r(\cdot)$ is replaced with the optimal policy $r^*(\cdot)$, there is equality throughout and we get,

$$A^{r^*} J(x,c,t) + \frac{1}{c} g(r^*) = 0 \tag{14}$$

Hence, for a given system state (x, c, t), the optimal transmission rate, r^*, is the value that minimizes (11) and the minimum value of the expression equals zero. Thus, we get,

$$\min_{r \in [0,\infty)} \left[\frac{g(r)}{c} + A^r J(x,c,t) \right] = 0 \tag{15}$$

Substituting $A^r J(\cdot)$ from (13), we get a partial differential equation (PDE) in $J(x, c, t)$[2] which is also referred as the Hamilton-Jacobi-Bellman (HJB) equation.

$$\min_{r \in [0,\infty)} \left\{ \frac{g(r)}{c} + \frac{\partial J(x,c,t)}{\partial t} - r\frac{\partial J(x,c,t)}{\partial x} + \lambda(E_z[J(x,Z(c)c,t)] - J(x,c,t)) \right\} = 0 \tag{16}$$

The boundary conditions for the above PDE are, $J(0, c, t) = 0$, and $J(x, c, T) = \infty,$ if $x > 0$. The last condition follows due to the deadline constraint of T on the data. Now, given any smooth solution of the HJB equation in (16), we need to verify that such a solution actually represents the value function. This is indeed the case, however, the verification theorems are omitted for brevity.

IV. OPTIMAL TRANSMISSION POLICY

The optimality conditions, derived in the last section, apply to general channel models and power-rate functions. We now specialize to the class of monomial functions, $g(r) = r^n$.

A. Constant Drift Channel

We, first, consider a simplified channel model referred to as the *constant drift channel* and then extend the results to the general Markov channel case. As mentioned earlier, there are two mechanisms that affect the minimization of the energy cost. First, the channel variations, $c(t)$, which the optimal policy can exploit by increasing the transmission rate under good states. Second, the convexity of $g(r)$, which dictates that high rate transmissions are energy costly. Hence, the data transmission must also be spread over time.

Constant Drift Channel: The channel process evolves as described in Section II-B with the additional assumption that the expected value of the random variable, $1/Z(c)$, is independent of the channel state, i.e. $E[1/Z(c)] = \beta$ (a constant). Since $\tilde{c} = Z(c)c$, at the next transition we have $E\left[\frac{1}{\tilde{c}}\right] = E\left[\frac{1}{Z(c)}\right]\frac{1}{c} = \beta/c$. If we look at the process $1/c(t)$, the above assumption means that over the interval of interest, the expected value of the next state (given the present state $1/c$) is a constant multiple of the present state. We refer to β

[2]Note, the minimization on the left eliminates the variable r and gives an equation in a single function $J(x, c, t)$.

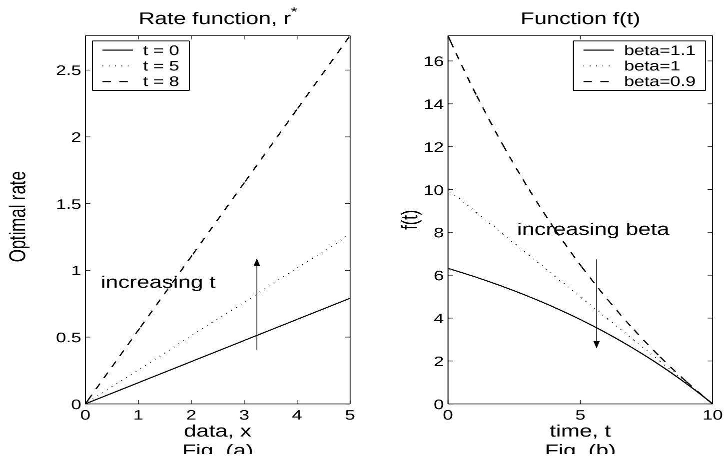

Fig. 2. (a) Plot of $r^*(x, c, t)$ versus x with $\beta = 1.1$, (b) plot of $f(t)$ versus t. Other parameters include $g(r) = r^2$, $T = 10$ and $\lambda = 1$.

as the "drift" parameter of the channel process. If $\beta > 1$, the process $1/c(t)$ drifts upwards in an expected sense, if $\beta = 1$, there is no expected drift and if $\beta < 1$, the drift is downwards. As an example, suppose that the channel transition either improves the channel by a factor u with probability p_u or worsens it by a factor d. This gives, $Z(c) = u \ or \ d$ with the respective probabilities and $\beta = \frac{p_u}{u} + \frac{1-p_u}{d}$, independent of c.

In practice, one observes that the wireless channel fluctuates between intervals of upward and downward drifts. For example, when a mobile device is moving in the direction of the base station, the channel has an expected drift towards improving conditions (equivalently $1/c(t)$ is drifting downwards or $\beta < 1$) and vice-versa for movement away from the base station. Similarly, in case of satellite channels, changing weather conditions (such as cloud cover) makes the channel drift on average towards worsening conditions and vice-versa. For slow fading channels, the time scale of these drift changes is usually longer than the packet deadlines in which case over the time interval of interest a constant drift channel is an appropriate model.

Theorem *I:* Consider the BT-problem with $g(r) = r^n$, $n > 1, n \in \mathbb{R}$ and a constant drift channel with parameter β. The optimal policy, $r^*(x, c, t)$, and the value function, $J(x, c, t)$, $\forall (x, c, t) \in \Omega, t < T$ are,

$$r^*(x, c, t) = \frac{x}{f(t)} \tag{17}$$

$$J(x, c, t) = \frac{x^n}{c(f(t))^{n-1}} \tag{18}$$

where $f(t) = \frac{(n-1)}{\lambda(\beta-1)}(1 - \exp(-\frac{\lambda(\beta-1)}{n-1}(T - t)))$.

Figure 2 plots the optimal rate function versus x for different t values and the function $f(t)$ for different β values. The various features of the optimal policy can be observed as follows. First, the rate function in (17) increases linearly in x; thus, with more data left in the queue the transmission rate is proportionally higher. Second, for any fixed t, we can view the slope $1/f(t)$ as the "urgency" of transmission since a higher $1/f(t)$ implies a steeper rate function. It can be easily verified both analytically and from the plot that $1/f(t)$ increases (or equivalently $f(t)$ decreases) as t gets closer to the deadline. Third, to quantify the dependence on β, observe that the expected future cost $\left(E\left[\frac{g(r)}{Z(c)c}\right] = \frac{g(r)}{c}\beta\right)$ is β times the present cost, $\frac{g(r)}{c}$. For $\beta > 1$, the channel is drifting towards higher expected cost or worsening conditions and vice-versa. Thus, as β increases it becomes more energy efficient to utilize the present conditions and hence the urgency of transmission $1/f(t)$ increases with β. Again, this can be verified both analytically and from the plot. Interestingly, when $\beta = 1$, the expected future energy cost does not change and in this case the optimal policy reduces to $r^*(x, c, t) = \frac{x}{T-t}$. This is simply emptying the x units in time $T - t$ by spreading the data transmission over the time left without adjusting for the channel variations. Such a policy will be referred to as the *Simple Drain* policy. Similarly, when $\lambda = 0$[3], there are no channel fluctuations and, as expected, the optimal policy is again simple drain, i.e. $r^*(x, c, t) = \frac{x}{T-t}$.

Another interesting feature of the optimal rate is that it depends on the channel process through the parameter β but does not depend on the present channel state $c(t)$. To understand such a behavior, first note that the optimal rate should depend only on the relative future channel evolution starting from the present state and not on the present channel value. For the constant drift channel, the relative future statistics in terms of the channel drift does not depend on the present channel state; hence the observed result. A practical advantage of this feature is that the average parameter β is easy to measure in real time by time averaging.

B. General Markov Channel Model

The results for the constant drift channel motivate much of the discussion in this section where we consider general Markov channel models. Before proceeding further, we present additional notations regarding the channel process. The channel state space is assumed to be finite and the various states $c \in \mathcal{C}$ are denoted as $c^1, c^2, \ldots, c^m$. Given a channel state c^i, the values taken by the random variable $Z(c^i)$ are denoted as $\{z_{ij}\}$, where $z_{ij} = c^j/c^i$. The probability that $Z(c^i) = z_{ij}$ is denoted as p_{ij}. Clearly, if there is no transition from state c^i to c^j, $p_{ij} = 0$.

We know from the last section that the optimal rate for the constant drift channel is given as, $r^*(x, c, t) = \frac{x}{f(t)}$. As a first guess for the general case, one might consider a similar functional form of the optimal rate and check if the HJB equation is satisfied. Towards this end, let $r^*(x, c^i, t) = \frac{x}{f_i(t)}$, where the function $f_i(t)$ is associated with the channel state c^i. Unlike the constant drift channel, the rate function $r^*(\cdot)$ now depends on the channel state through the different functions $\{f_i(t)\}_{i=1}^m$ and over an interval when the channel state is c^i, policy $\frac{x}{f_i(t)}$ is followed. The following theorem proves the optimality of such a policy and obtains the functions $\{f_i(t)\}_{i=1}^m$ as the solution of a set of ODE's, that satisfy the HJB equation. This system of ODE's can be easily solved offline numerically using standard techniques and no further computation is required during the system operation.

Theorem *II:* Consider the BT-problem with $g(r) = r^n$, $n > 1, n \in \mathbb{R}$ and the Markov channel model. The optimal policy, $r^*(x, c, t)$, and the value function, $J(x, c, t)$, for all $(x, c, t) \in \Omega, t < T$ are,

$$r^*(x, c^i, t) = \frac{x}{f_i(t)}, \; i = 1, \ldots, m \tag{19}$$

$$J(x, c^i, t) = \frac{x^n}{c^i(f_i(t))^{n-1}}, \; i = 1, \ldots, m \tag{20}$$

where $\{f_i(t)\}_{i=1}^m$ is the solution of the following ODE with the boundary conditions $f_i(T) = 0, f_i'(T) = -1, \forall i$ [4],

$$f_1'(t) = -1 - \frac{\lambda f_1(t)}{n-1} + \frac{\lambda}{n-1}\sum_{k=1}^{m}\frac{p_{1k}}{z_{1k}}\frac{(f_1(t))^n}{(f_k(t))^{n-1}} \tag{21}$$

$$\vdots$$

$$f_m'(t) = -1 - \frac{\lambda f_m(t)}{n-1} + \frac{\lambda}{n-1}\sum_{k=1}^{m}\frac{p_{mk}}{z_{mk}}\frac{(f_m(t))^n}{(f_k(t))^{n-1}} \tag{22}$$

[3]For both $\lambda = 0$ and $\beta = 1$, $f(t)$ is evaluated by taking appropriate limit.

[4]For numerical purposes, these two conditions can be incorporated by taking a small $\epsilon > 0$, letting $f_i(T - \epsilon) = \epsilon, \forall i$ and solving the ODE as an Initial Value problem over $[0, T - \epsilon]$.

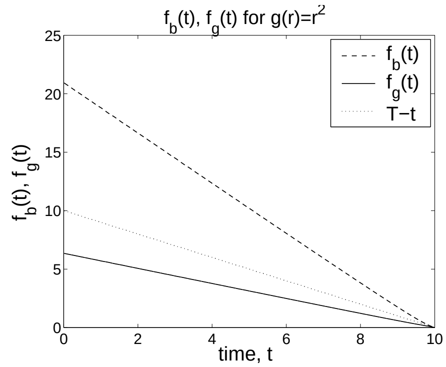

Fig. 3. $f_b(t)$ and $f_g(t)$ plot for the bad and the good channel respectively. Other parameters include, $T = 10, \lambda = 5, \gamma = 0.3$.

As an example, consider the Gilbert-Elliott (GE) channel model referred to in Section II-B. Let $g(r) = r^2$ and for simplicity take $\lambda_{bg} = \lambda_{gb} = \lambda$. Denoting $\gamma = c_b/c_g$, we have, $Z(c_g) = \gamma$, and $Z(c_b) = 1/\gamma$. As the Z representation does not have any self-transitions, we have $p_{gg} = p_{bb} = 0$. Denoting $f_b(t), f_g(t)$ as the respective functions in the bad and the good states, the ODE becomes,

$$f_b'(t) = -1 - \lambda f_b(t) + \frac{\gamma\lambda(f_b(t))^2}{f_g(t)} \tag{23}$$

$$f_g'(t) = -1 - \lambda f_g(t) + \frac{\lambda(f_g(t))^2}{\gamma f_b(t)} \tag{24}$$

Figure 3 plots these functions for $T = 10, \lambda = 5, \gamma = 0.3$. First, as expected $f_g(t) \leq f_b(t), \forall t$, which implies that given x units of data in the buffer and time t, the rate $\frac{x}{f_g(t)}$ is higher under the good channel state. Second, it can be shown that, $f_g(t) \leq T-t \leq f_b(t)$, where the function, $f(t) = T-t$, gives the rate, $\frac{x}{T-t}$, corresponding to the simple drain (SD) policy. Thus, compared to SD policy, the optimal policy, as expected, increases the transmission rate under the good state and reduces the rate when the channel is bad. The optimal adaptation, however, is governed by the respective $f_i(t)$.

As a comparison, we now present simulation results for the optimal and the SD policy. The transmission rate for the SD policy is given as, $r^*(x, c, t) = \frac{x}{T-t}$ and such a policy is optimal when the channel is time-invariant. In the simulations, we consider $g(r) = r^2$ and the GE channel model as described above. We take $T = 10$ and partition the interval $[0, 10]$ into slots of length $dt = 10^{-3}$, thus, having $10,000$ time slots. A channel sample path is simulated using a Bernoulli process as follows. In a time slot, the channel transitions with probability λdt and with probability $1 - \lambda dt$ there is no transition. At each transition, the new state is $\tilde{c} = Z(c)c$ which for the GE model amounts to jumps between the two states. Expected energy cost is computed by taking an average over 10^4 sample paths.

Figure 4(a) plots the energy costs of the two policies as λ is varied with $\gamma = 0.3$, $B = 10$. When λ is small the channel is essentially time-invariant over the deadline interval and the two policies are comparable. As λ increases, the optimal cost substantially decreases due to the channel adaptation. In Figure 4(b), γ is varied with $\lambda = 5$, $B = 10$. As γ decreases the good and the bad channel quality differ significantly and the optimal rate adaptation leads to a substantially lower energy cost as compared to SD policy.

V. Conclusion

We considered optimal rate control for delay constrained data transmission over time-varying channels. Specifically, we looked at the problem of transmitting B units of data by deadline T over a general time-varying Markov channel. Using a novel approach based on a continuous-time formulation and stochastic

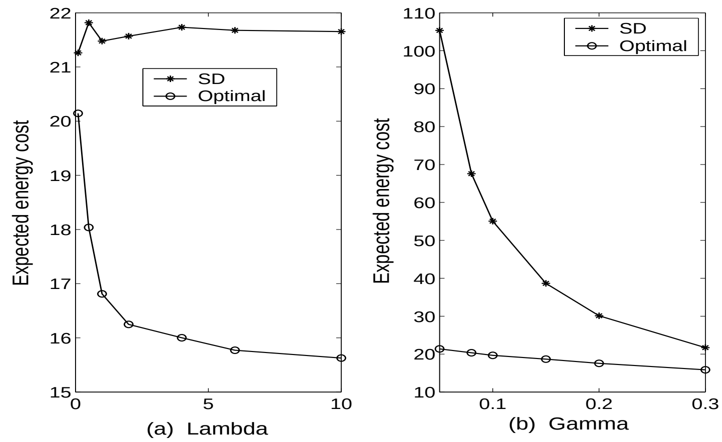

Fig. 4. Expected energy cost for the optimal and SD policies.

control theory we obtained closed form expressions for the optimal rate that minimizes the transmission energy expenditure. We believe that the continuous-time framework holds promise for various extensions addressing quality of service constrained data transmission in wireless systems. Some of the natural extensions include incorporating a stochastic model for the arrival process, down-link scenarios in cellular networks with multiple mobile devices and multi-hop transmissions with end to end constraints.

REFERENCES

[1] R. Berry, R. Gallager, "Communication over fading channels with delay constraints", *IEEE Tran. on Information Theory*, vol. 48, no. 5, May 2002.

[2] B. Prabhakar, E. Biyikoglu, A. El Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling", *IEEE INFOCOM 2001*, vol. 1, pp 386 - 394, April 2001.

[3] A. Fu, E. Modiano, J. Tsitsiklis, "Optimal energy allocation for delay constrained data transmission over a time-varying channel", *IEEE INFOCOM 2003*, vol. 2, pp 1095 - 1105, April 2003.

[4] M. Zafer, E. Modiano, "A Calculus Approach to Minimum Energy Transmission Policies with Quality of Service Guarantees", *IEEE INFOCOM 2005*, March 2005.

[5] M. Khojastepour, A. Sabharwal, "Delay-constrained scheduling: power efficiency, filter design and bounds", *IEEE INFOCOM 2004*, March 2004.

[6] P. Nuggehalli, V. Srinivasan, R. Rao, "Delay constrained energy efficient transmission strategies for wireless devices", *IEEE INFOCOM 2002*, vol. 3 , pp. 1765 - 1772,June 2002.

[7] A. Jalali, R. Padovani, R. Pankaj, "Data throughput of CDMA-HDR a high efficiency high data rate personal communication wireless system", *IEEE Vehicular Technology Conf.*, vol. 3, 2000.

[8] M. Neely, E. Modiano, C. Rohrs, "Power and server allocation in a multi-beam satellite with time varying channels", *IEEE INFOCOM 2002*, vol. 3, pp 1451-1460, June 2002.

[9] A. Goldsmith, "The capacity of downlink fading channels with variable rate and power", *IEEE Tran. Vehicular Technology*, vol. 45, pp. 1218-1230, Oct. 1997.

[10] B. Collins, R. Cruz, "Transmission policies for time varying channels with average delay constraints", Proceedings, 1999 Allerton Conf. on Comm., Control and Comp., Monticello, IL, 1999.

[11] B. Ata, "Dynamic power control in a wireless static channel subject to a quality of service constraint", *To appear in Operations Research*.

[12] Q. Zhang, S. Kassam, "Finite-State Markov Model for Rayleigh Fading Channels", *IEEE Tran. on Communications*, vol. 47, no. 11, Nov. 1999.

[13] W. Fleming and H. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, 1993.

[14] M. Davis, *Markov Models and Optimization*, Chapman and Hall, 1993.

[15] B. Oksendal, *Stochastic Differential Equations*, Springer, 5^{th} edn., 2000.

[16] Hurewicz W., *Lectures on Ordinary Differential Equations*, Dover Publications.