# Energy-efficient Scheduling of Delay Constrained Traffic over Fading Channels

Juyul Lee and Nihar Jindal

Department of Electrical and Computer Engineering
University of Minnesota
E-mail: {juyul, nihar}@umn.edu

***Abstract*—A delay-constrained scheduling problem for point-to-point communication is considered: a packet of $B$ bits must be transmitted by a hard deadline of $T$ slots over a time-varying channel. The transmitter/scheduler determines how many bits to transmit, or equivalently how much energy to transmit with, during each time slot based on the current channel quality and the number of unserved bits, with the objective of minimizing expected total energy. Assuming transmission at capacity of the underlying Gaussian noise channel, a closed-form expression for the optimal scheduling policy is obtained for the case $T = 2$ via dynamic programming; for $T > 2$, the optimal policy can only be numerically determined. Thus, the focus of the work is on derivation of simple, near-optimal policies. The proposed bit-allocation policies consist of a linear combination of a delay-associated term and an opportunistic (channel-aware) term. In addition, a variation of the problem in which the entire packet must be transmitted in a single slot is studied.**

## I. Introduction

This paper considers a scheduling problem of transmitting $B$ bits over $T$ time slots, where the channel fades independently from slot to slot. During each slot, the scheduler determines how many of the $B$ bits to transmit on the basis of the current channel quality and the number of unserved bits remaining. The scheduler must balance the desire to be opportunistic, i.e., serve many bits when the channel is good, with the hard deadline. As might be expected in multimedia transmission, this setup can be used to model deterministic traffic (e.g., in VoIP systems packets arrive regularly in time and generally have hard deadlines of approximately 50 msec).

Fu et al. [1] considered the same problem and showed it can be formulated as a dynamic program. However, an explicit analytical solution is given only for the case where the energy cost is a linear function of the number of transmitted bits. We consider the case where the cost function is given by the AWGN capacity expression, in which case the cost is not linear in the number of bits. In an earlier work, Negi and Cioffi [2] studied the dual problem, but they did not described in closed-form. Berry et al. [3] and Rajan et al. [4] considered "average" delay for constantly arriving packets. This is rather different from the hard deadline case considered here.

In this paper we develop low-complexity and near-optimal policies for delay-constrained causal scheduling. Our main result is that the proposed schedulers can be cast into a single framework: a time-dependent weighted sum of a delay associated term and an opportunistic term as

$$b_t = \underbrace{\frac{1}{t}\beta_t}_{\text{delay associated}} + \underbrace{\frac{t-1}{t}\log\frac{g_t}{\eta_t}}_{\text{opportunistic}}, \tag{1}$$

where $b_t$ is the number of bits to serve (from remaining $\beta_t$ bits) at time slot $t$ ($t$ denotes the number of remaining slots), $g_t$ denotes the current channel state, and $\eta_t$ denotes a threshold level. If the current channel quality is equal to the threshold level, then a fraction $\frac{1}{t}$ of the remaining bits are transmitted. If the channel quality is better/worse than expected (intuitively represented by the threshold level $\eta_t$), then additional/fewer bits are transmitted. As one might intuitively expect, the scheduler acts very opportunistically when the deadline is far away ($t$ large) but less so as the deadline approaches.

In addition, we consider the case of one-shot allocation. When the number of bits to transmit is small, it may desirable to transmit the entire packet in one time slot rather than over multiple time slots because of potential overhead associated with multiple slot transmission.

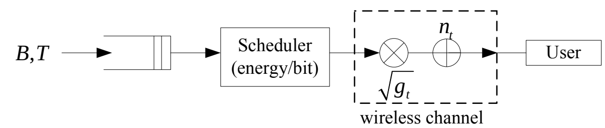



Fig. 1: Single-user delay constrained scheduling

## II. Problem Formulation

As illustrated in Fig. 1, a packet of $B$ bits must be transmitted within $T$ time slots through a fading channel, in which $T$ is referred to as the *delay-limit* or *deadline*. The purpose of the scheduler is to determine the energy, or equivalently the number of bits, to be served during each time slot, such that the expected energy is minimized and the bits are served by the deadline $T$ (no outage is allowed).

Time is indexed in descending order, i.e., $t = T$ is the initial slot, $t = T - 1$ is the 2nd slot, ..., and $t = 1$ is the final slot before the deadline. The channel state, in power units, is denoted as $g_t$. We assume that the channel states $\{g_t\}_{t=1}^{T}$ are i.i.d. and the scheduler has causal knowledge of these channel states (i.e., at time $t$, $g_T, g_{T-1}, \cdots, g_t$ are known

Juyul Lee is the recipient of a Motorola Partnerships in Research Grant.

but $g_{t-1}, \cdots, g_1$ are unknown). In this context, we refer to this type of scheduler as a *causal scheduler* that exploits the statistical quantities of the future channel state instead of actual realizations. Specifically, it utilizes the fractional moments that are defined as

$$\nu_m = \left(\mathbb{E}\left[\left(\frac{1}{g}\right)^{\frac{1}{m}}\right]\right)^m, \quad m = 1, 2, \ldots. \tag{2}$$

(See [5] for properties of the sequence $\{\nu_m\}$.) We also assume that the additive noise $n_t$ is Gaussian with zero mean and unit variance. Let $b_t (\in \mathbb{R})$ be the number of bits allocated at time slot $t$ and $\beta_t$ be the remaining bits at time $t$; thus recursively as $\beta_t = \beta_{t+1} - b_{t+1}$. Assuming transmission at capacity, the number of transmitted bits at slot $t$ is given by $b_t = \log_2(1 + g_t E_t)$, where $E_t$ is the energy of transmission during the time slot. By solving for $E_t$, we have

$$E_t(b_t, g_t) = \frac{2^{b_t} - 1}{g_t}. \tag{3}$$

Given this setup, a scheduler is a sequence of functions $\{b_t\}_{t=1}^T$ that maps from the remaining bits and the current channel state to the number of bits served, i.e., $b_t(\beta_t, g_t) \in [0, \beta_t]$, $t = T, \cdots, 1$. Then, the optimal energy-efficient scheduler is the set of scheduling functions $\{b_t^{\text{opt}}(\cdot,\cdot)\}_{t=1}^T$ that minimizes the total expected energy cost: i.e.,

$$\min_{b_T, \cdots, b_1} \mathbb{E}\left[\sum_{t=1}^T E_t(b_t, g_t)\right] \tag{4}$$

subject to $\sum_{t=1}^T b_t = B$ and $b_t \geq 0$ for all $t$.

The optimization (4) can be formulated sequentially (via dynamic programming) with the remaining bits $\beta_t$ as a state variable that summarizes the bit allocation up until the previous time step:

$$b_t^{\text{opt}}(\beta_t, g_t) = \arg\min_{0 \leq b_t \leq \beta_t} \left\{E_t(b_t, g_t) + \mathbb{E}\left[\sum_{s=1}^{t-1} E_s(b_s, g_s) \middle| b_t\right]\right\} \tag{5}$$

for $t = T, \ldots, 2$. At the last time slot $t = 1$, $b_1^{\text{opt}}(\beta_1, g_1) = \beta_1$ for all $g_1$ trivially since no optimization is required. Note that the optimization (5) should be performed for all possible pairs $(\beta_t, g_t)$. In fact, finding a general solution for (5) is not easy in the context of optimization theory.

## III. Optimal Scheduling

In this section we derive the optimal causal scheduler for $T = 2$ and discuss the difficulty in obtaining an analytical form of the optimal causal scheduler for $T > 2$. We also derive a scheduler assuming the channel states are known non-causally. From the derived schedulers, we interpret how the delay constraint and the channel state affect the scheduling process, which will also turn out be the case for suboptimal schedulers later.

### A. Causal Scheduling

Under perfect causal CSI, we derive the optimal scheduler using the conventional dynamic programming technique [6].

*1) $T = 2$:* In the last time slot ($t = 1$), the scheduler is required to transmit $\beta_1$ unserved bits regardless of the channel state $g_1$. Thus, $E_1(\beta_1, g_1) = (2^{\beta_1} - 1)/g_1$ for all $g_1$. At $t = 2$, $g_2$ is known but $g_1$ is unknown. The scheduler needs to determine $b_2$, which should be a function of $g_2$ and $B$, while balancing between the current energy cost and the *expected* future cost, i.e., minimizing the sum:

$$b_2^{\text{opt}}(B, g_2) = \arg\min_{0 \leq b_2 \leq B} \left(\frac{2^{b_2} - 1}{g_2} + \mathbb{E}\left[E_1(B - b_2, g_1)\right]\right) \tag{6}$$

The solution to (6) can be easily found since the expected future cost has a simple form in this case:

$$b_2^{\text{opt}}(B, g_2) = \left\langle \frac{1}{2}B + \frac{1}{2}\log_2(g_2 \nu_1)\right\rangle_0^B, \tag{7}$$

where $\langle\cdot\rangle_0^B$ denotes truncation from below at 0 and truncation from above at $B$. Note that this policy is only meaningful when $\nu_1 (= \mathbb{E}[1/g])$ is finite; this rules out Rayleigh fading, in which case $g$ is exponentially distributed and thus $\mathbb{E}[1/g]$ is not finite.

Notice that the optimal scheduling function (7) has two additive terms: (a) $\frac{1}{2}B$ corresponds to an equal distribution to time slots $t = 1$ and $t = 2$, and (b) $\frac{1}{2}\log_2(g_2\nu_1)$ associated with a measure of the channel quality at $t = 2$. That is, if the channel quality $g_2$ is bigger than a threshold $1/\nu_1$, then more bits are allocated than $\frac{1}{2}B$; if $g_2$ is smaller than the threshold then fewer bits are allocated and more bits are deferred to the final slot.

*2) $T > 2$:* The optimization that the scheduler solves at each time step $t$ is:

$$J_t(\beta_t, g_t) = \begin{cases} \min_{0 \leq b_t \leq \beta_t}\left(\frac{2^{b_t}-1}{g_t} + \bar{J}_{t-1}(\beta_t - b_t)\right), & t \geq 2 \\ E_1(\beta_1, g_1), & t = 1 \end{cases} \tag{8}$$

where $\bar{J}_{t-1}(\beta_{t-1}) = \mathbb{E}_{g_{t-1}}[J_{t-1}(\beta_{t-1}, g_{t-1})]$. This is simply a one-dimensional optimization over $b_t$, but it does not seem feasible to derive an analytic solution (see [5] for discussion). When the energy cost is a linear function of the control $b_t$, obtaining an analytic form is possible and is given in [1].

Alternately, we can numerically find the optimal scheduler by the discretization method [7]. Having sufficiently fine discretization is computationally burden and requires large memory size, however. Furthermore, this discretization gives little insight on how the delay constraint and channel state affect the scheduling function.

### B. Non-causal Scheduling

If the channel states are known non-causally, i.e., $g_T, \ldots, g_1$ are known at $t = T$, the optimal scheduling/allocation is determined by waterfilling because each channel at each time slot serves as a parallel channel. While conventional waterfilling maximizes rate subject to a power constraint, this is rather the dual of minimizing power/energy subject to a rate/bit constraint. This will be referred to as *inverse-waterfilling* and the solution is obtained through the standard

Lagrangian method as:

$$b_t^{\text{IWF}} = \frac{1}{t'}\beta_t + \frac{t'-1}{t'}\log_2\frac{g_t}{\eta_t^{\text{IWF}}}, \quad \text{if } g_t \geq g_{\text{th}} \tag{9}$$

where $t' = \sum_{i=1}^{t} 1_{\{g_i \geq g_{\text{th}}\}}$ denotes the number of *active* time slots that participate in the waterfilling up to $t$,

$$\eta_t^{\text{IWF}} = \left(\prod_{i=1}^{t-1} g_i^{1\{g_i \geq g_{\text{th}}\}}\right)^{1/(t'-1)}, \tag{10}$$

and $g_{\text{th}} = \left(\prod_{i=1}^{T} g_t^{1\{g_i \geq g_{\text{th}}\}}\right)^{1/T'} / 2^{B/T'}$ [5]. Note that $g_{t-1}, \cdots, g_1$ are relatively future quantities at slot $t$.

In (9), the bit allocation process is described in two stages: first the $\beta_t$ bits are divided equally amongst the active slots at $t$ (i.e., $\beta_t/t'$ term) and then bits are added/subtracted depending on the channel state (i.e., $g_t/\eta_t^{\text{IWF}}$ term). We will soon see that a very similar interpretation can be given to the proposed suboptimal causal scheduling policies.

## IV. Suboptimal Scheduling Algorithms

### A. Suboptimal I Algorithm

If we compare the optimal causal scheduler for $T = 2$ (Section III-A) to the non-causal scheduler (Section III-B), we can immediately notice that the optimal scheduler determines $b_2^{\text{opt}}$ by inverse-waterfilling over channels $g_2$ and $1/\nu_1$. This is because of the particularly simple form of the expected future cost. Although the expected future cost does not take on such a simple form for $T > 2$, we can get a suboptimal scheduler by simply applying this inverse-waterfilling at every time slot $t$. In other words, at time step $t$, perform inverse-waterfilling over the following $t$ channels:

$$g_t, \underbrace{\frac{1}{\nu_1}, \ldots, \frac{1}{\nu_1}}_{t-1}.$$

Since $t-1$ of the $t$ channels are equal, performing this inverse-waterfilling is extremely simple and the solution is given by

$$b_t^{(\text{I})}(\beta_t, g_t) = \left\langle \frac{1}{t}\beta_t + \frac{t-1}{t}\log_2\frac{g_t}{\eta_t^{(\text{I})}} \right\rangle_0^{\beta_t}, \tag{11}$$

where $\eta_t^{(\text{I})} = 1/\nu_1$ serves as the channel threshold.

When the deadline is far away (large $t$), the first term in (11) is negligible and the bit allocation is nearly completely dependent on the channel measure. As the deadline approaches ($t$ decreases toward 1), the weight of the channel-dependent second term decreases and the weight of the delay-associated first term increases.

### B. Suboptimal II Algorithm

The inability to find a general closed-form solution to the original optimization (8) is due to complications induced by the boundary constraints $0 \leq b_t \leq \beta_t$ for every $t$ in dynamic programming. However, if we relax this constraint in (8), we can in fact find a simple analytical solution.

From the relaxation, we can show the expected future cost is given by [5]

$$\bar{J}_{t-1}^{(\text{II})}(\beta_{t-1}) = (t-1)2^{\frac{\beta_{t-1}}{t-1}}\,\mathbb{G}(\nu_{t-1}, \ldots, \nu_1) - (t-1)\nu_1 \tag{12}$$

where $\mathbb{G}(\nu_{t-1}, \ldots, \nu_1)$ denotes the geometric mean of $\nu_{t-1}, \ldots, \nu_1$. Since we relax the constraint "$0 \leq b_t \leq \beta_t$" in (8), the solution can be found by solving an unconstrained optimization. Then by truncating both from below and above, we obtain a scheduling policy and we refer this to as the suboptimal II scheduler:

$$b_t^{(\text{II})} = \left\langle \frac{1}{t}\beta_t + \frac{t-1}{t}\log_2\frac{g_t}{\eta_t^{(\text{II})}} \right\rangle_0^{\beta_t}, \tag{13}$$

where $\eta_t^{(\text{II})} = 1/\,\mathbb{G}\left(\nu_{t-1}, \nu_{t-2}, \cdots, \nu_1\right)$ denotes a threshold that depends only on the statistics not the realizations.

Notice that the only difference between suboptimal I and II is in the threshold values. For suboptimal I, the threshold $\eta_t^{(\text{I})}$ is constant for all $t$. On the other hand, the threshold for suboptimal II $\eta_t^{(\text{II})}$ actually decreases as $t \to 1$ (see [5]). That is, suboptimal II has a higher threshold in the beginning time steps and the level of threshold get smaller as time goes.

### C. General Framework

The algorithms thus far considered can be cast into a single framework:

$$b_t(\beta_t, g_t) = \left\langle \frac{1}{t}\beta_t + \frac{t-1}{t}\log_2\frac{g_t}{\eta_t} \right\rangle_0^{\beta_t}, \tag{14}$$

where $\eta_t$ is a threshold determined by the individual algorithms. This simple allocation strategy reveals how the delay constraint works on the scheduling algorithms: at time step $t$ serve a fraction $1/t$ of the remaining bits plus/minus a quantity that depends on the strength of the current channel compared to a channel threshold. If the current channel is good (i.e., $g_t$ is bigger than the threshold $\eta_t$), additional bits are served (up to $\beta_t$), while fewer bits are served when the current channel is poor. Furthermore, note that when $t$ is large (i.e., far from the deadline), the first term $\beta_t/t$ is very small and the number of bits served is almost completely determined by the current channel conditions. This agrees with intuition that we should make aggressive, almost completely channel dependent (and deadline independent) decisions when the deadline is far away, while we should make more conservative (more deadline dependent, less channel dependent) decisions when the deadline is approaching (small $t$).

The behavior among the algorithms is characterized by the channel thresholds. Comparing the threshold-associated terms of the suboptimal I and II, we have $\mathbb{E}\left[\log_2(g_t/\eta_t^{(\text{I})})\right] = \mathbb{E}\left[\log_2\left(g_t\nu_1\right)\right] > 0$ for all $t$ and $\lim_{t\to\infty}\mathbb{E}\left[\log_2(g_t/\eta_t^{(\text{II})})\right] = 0$, respectively. This implies that the bit allocation for the suboptimal I is overly aggressive and thus the bit allocation may finish early.

## V. Asymptotic Analysis & Numerical Results

We compare the performance between the causal and the non-causal schedulers as well as between the causal and the equal-bit schedulers. The comparison analysis is performed asymptotically for two extreme cases: $B \to 0$ and $B \to \infty$.

### A. $T = 2$

From (7), we can see that the packet is split over both time slots (i.e., $b_2 > 0$ and $b_1 > 0$) if $2^{-B}/\nu_1 < g_2 < 2^B/\nu_1$. As $B \to 0$, the probability of this event clearly goes to zero: if $g_2 < 1/\nu_1$ then all bits are deferred to the final slot, while if $g_2 > 1/\nu_1$ all bits are served at $t = 2$. As a result, the expected energy cost takes on a rather simple form as $B \to 0$. A very similar statement can be made about non-causal inverse-waterfilling (both slots are used if $2^{-B}g_1 \leq g_2 \leq 2^B g_1$).

$$\bar{J}_2^{\text{eq}}(B) = 2(2^{\frac{B}{2}} - 1)\,\mathbb{E}\left[\frac{1}{g}\right] \tag{15}$$

$$\bar{J}_2^{\text{opt}}(B) \cong (2^B - 1)\,\mathbb{E}\left[\min\left(\frac{1}{g_2}, \nu_1\right)\right], \tag{16}$$

$$\bar{J}_2^{\text{IWF}}(B) \cong (2^B - 1)\,\mathbb{E}\left[\min\left(\frac{1}{g_2}, \frac{1}{g_1}\right)\right], \tag{17}$$

where $\cong$ represents equivalence in the limit (i.e., the difference between both sides converges to 0 as $B \to 0$). The corresponding effective channel of the causal and the non-causal is given as $\max(g_2, 1/\nu_1)$ and $\max(g_2, g_1)$, respectively. Thus, the performance offsets as $B \to 0$ are quantified by the following theorem:

*Theorem 1:* If $g$ is a continuous random variable, the ratio of the expected energy costs for $T = 2$ schedulers as $B \to 0$ are given by:

$$\lim_{B\to 0} \frac{\bar{J}_2^{\text{eq}}(B)}{\bar{J}_2^{\text{opt}}(B)} = \frac{\mathbb{E}\left[\frac{1}{g}\right]}{\mathbb{E}\left[\min\left(\frac{1}{g_2}, \nu_1\right)\right]} \tag{18}$$

$$\lim_{B\to 0} \frac{\bar{J}_2^{\text{opt}}(B)}{\bar{J}_2^{\text{IWF}}(B)} = \frac{\mathbb{E}\left[\min\left(\frac{1}{g_2}, \nu_1\right)\right]}{\mathbb{E}\left[\min\left(\frac{1}{g_2}, \frac{1}{g_1}\right)\right]}. \tag{19}$$

Similarly at the other extreme, when $B \to \infty$:

*Theorem 2:* If $g$ is a continuous random variable. the ratio of the expected energy costs for $T = 2$ schedulers as $B \to \infty$ are given by:

$$\lim_{B\to \infty} \frac{\bar{J}_2^{\text{eq}}(B)}{\bar{J}_2^{\text{opt}}(B)} = \sqrt{\frac{\nu_1}{\nu_2}} \tag{20}$$

$$\lim_{B\to \infty} \frac{\bar{J}_2^{\text{opt}}(B)}{\bar{J}_2^{\text{IWF}}(B)} = \sqrt{\frac{\nu_1}{\nu_2}} \tag{21}$$

*Proof:* See [5]. ■

### B. $T > 2$

We extend the asymptotic results for any finite $T$ with the suboptimal II scheduler instead of the optimal scheduler because no analytical form for the optimal scheduler is available when $T > 2$.

*Theorem 3:* The ratio of the two expected energy costs for finite $T$ converges for small $B$ and large $B$ respectively as

$$\lim_{B\to 0} \frac{\bar{J}_T^{(\text{II})}}{\bar{J}_T^{\text{IWF}}} = \tag{22}$$

$$\frac{\mathbb{E}\left[\min\left(\frac{1}{g_T}, \mathbb{E}\min\left(\frac{1}{g_{T-1}}, \cdots \mathbb{E}\min\left(\frac{1}{g_2}, \nu_1\right)\right)\right)\right]}{\mathbb{E}\left[\min\left(\frac{1}{g_T}, \frac{1}{g_{T-1}}, \cdots, \frac{1}{g_1}\right)\right]},$$

$$\lim_{B\to \infty} \frac{\bar{J}_T^{(\text{II})}}{\bar{J}_T^{\text{IWF}}} = \frac{G_T}{\nu_T}, \tag{23}$$

where $G_t = \mathbb{G}(\nu_t, \nu_{t-1} \cdots, \nu_1)$.

*Proof:* The result is a straightforward extension of the proof for the $T = 2$ case. ■

In (23) we can see that the average energy of the suboptimal II scheduler approaches the energy used by (non-causal) inverse-waterfilling due to the property $\lim_{T\to\infty} \nu_T = \lim_{T\to\infty} G_T$ under $B/T = R$ fixed with large $R$ [5].

### C. Additional Numerical Results

So far we have investigated the asymptotic performance in two extremes. In this subsection, we examine the performances in mild conditions by simulations. Throughout the simulations, we assume that the distribution of the channel state $g_t$ is a truncated exponential with parameter $\lambda = 1$ and threshold $\gamma_0 = 0.001$.

From Fig. 2a we see that both suboptimal I and II perform nearly as well as the optimal scheduler, although suboptimal II performs slightly better than I for small values of $B$. There are significant differences between the equal-bit, optimal causal, and non-causal schedulers, which is to be expected given the time diversity available over the five time slots. In Fig. 2b we see even larger differences between equal-bit, optimal causal, and optimal non-causal, which can be explained by the even larger degree of time diversity ($T = 50$). Furthermore, suboptimal II significantly outperforms suboptimal I for $T = 50$ due to the over-aggressive nature of suboptimal I as discussed in Section IV-C. Suboptimal II performs nearly as well as the optimal scheduler when $B$ is approximately 50 or larger (i.e., $B/T \geq 1$), but is sub-optimal for smaller values of $B$. It is also interesting to note that the there is only a very slight difference between optimal causal and non-causal scheduling when $T = 50$; intuitively, the large amount of time diversity reduces the advantage of having non-causal CSI.

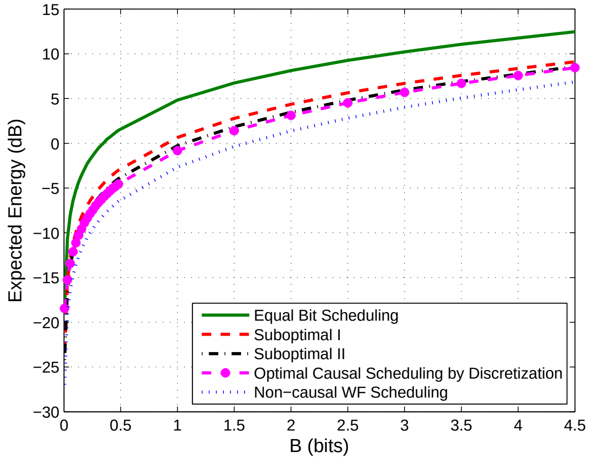
## VI. One-shot Allocation

When the number of bits $B$ to allocate during $T$ time slots is small, splitting bit allocation across multiple time slots may not be wanted. In this setting, it is desirable to find only one time slot among the $T$ slots for the transmission of $B$ bits.

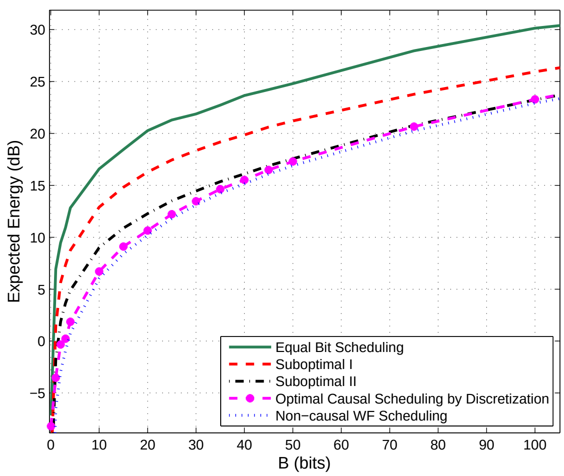
The dynamic program in this setting can be written as an optimal stopping problem [6]:

$$J_1(\beta_1) = \frac{2^{\beta_1} - 1}{g_1}, \tag{24}$$

$$J_t(\beta_t) = \min\left\{\frac{2^{\beta_t} - 1}{g_t}, \mathbb{E}[J_{t-1}(\beta_t)]\right\},\ t = 2, \cdots, T, \tag{25}$$




(a) $T = 5$




(b) $T = 50$

Fig. 2: Average total energy consumption

Thus, the optimal policy can be stated:

$$b_t = \begin{cases} B, & t = \max\{s : g_s > 1/\omega_s\}, \\ 0, & \text{elsewhere.} \end{cases} \tag{26}$$

The threshold can be calculated recursively as follows (See [5] for derivation):

$$\omega_t = \mathbb{E}\left[\frac{1}{g}\Big|\frac{1}{g_t} < \omega_{t-1}\right] \Pr\left\{\frac{1}{g_t} < \omega_{t-1}\right\} + \omega_{t-1} \Pr\left\{\frac{1}{g_t} \geq \omega_{t-1}\right\} \tag{27}$$

It is obvious that no threshold is required at the last time $t = 1$ ($\omega_1 = \infty$). Notice that the threshold $1/\omega_t$ depends only on the channel statistics and does not depend on $B$. From the update formula of $\omega_t$, we can see that $\omega_t \leq \omega_{t-1}$ (i.e., $1/\omega_t \geq 1/\omega_{t-1}$). Intuitively, this makes sense; the threshold get smaller as getting closer to the deadline.

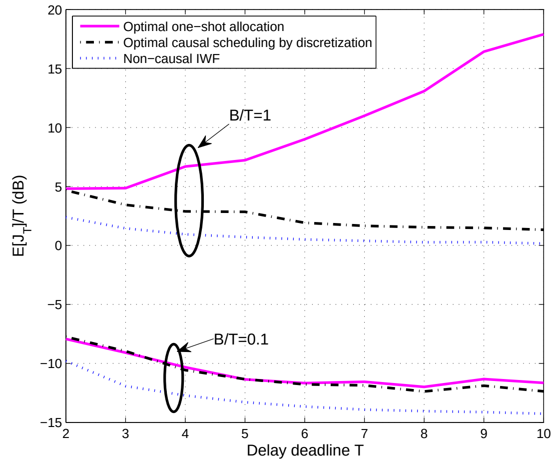



Fig. 3: Performance of the optimal one-shot allocation

Figure 3 illustrates the energy performance.When $B$ is relatively small, the power (energy/time) of the one-shot allocation is nearly the same as the optimal causal policy that allows for multiple slots to be used. (The performance of the inverse-waterfilling yields the theoretical bound). However, this one-shot allocation is not appropriate when $B$ is relatively large because the required energy grows exponentially with $B$.

## VII. Conclusion

In this paper we considered a bit allocation over finite time horizon, assuming perfect instantaneous channel state information is available. The proposed schedulers have a simple and intuitive form that gives insight into the optimal balance between channel-awareness (i.e., opportunism) and deadline-awareness in a delay-limited setting. We also considered the same problem under the additional constraint that only a single of the available time slots can be used, and in this case found the optimal threshold-based policy.

## References


[1] A. Fu, E. Modiano, and J. N. Tsitsiklis, "Optimal transmission scheduling over a fading channel with energy and deadline constraints," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 630–641, Mar. 2006.

[2] R. Negi and J. M. Cioffi, "Delay-constrained capacity with causal feedback," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2478–2494, Sep. 2002.

[3] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1135–1149, May. 2002.

[4] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 125–144, Jan. 2004.

[5] J. Lee and N. Jindal, "Energy-efficient scheduling of delay constrained traffic over fading channels," in preparation.

[6] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Mass.: Athena Scientific, 2005, vol. 1.

[7] ——, "Convergence of discretization procedures in dynamic programming," *IEEE Trans. Automat. Contr.*, vol. AC-20, no. 3, pp. 415–419, Jun. 1975.