International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

Impact of Processing Energy on the Capacity of Wireless Channels

Pamela YOUSSEF MASSAAD†, Muriel MEDARD†, and Lizhong ZHENG †

† Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology,
77 Massachusetts Avenue Cambridge 02139, USA
E-mail: {pmassaad, medard, lizhong}@mit.edu

Abstract

Energy efficiency is an important issue in mobile wireless networks, since the battery life of mobile terminals is limited. In this paper, we address this issue from the information theoretic point of view. Traditional information theoretic energy constraints consider only the energy used for transmission purposes. We study optimal transmission strategies by explicitly taking into account the energy expended by processes other than transmission, that run when the transmitter is in the 'on' state. We term this energy by 'processing energy'. Under these new constraints, we derive the capacity of an Additive White Gaussian Noise (AWGN) channel. We prove that, unlike the case where only transmission energy is taken into account, achieving capacity may require intermittent, or 'bursty', transmission. In particular, we show that in the low SNR regime, burstiness is optimal when the processing energy is greater than half the square of the total energy available to the transmitter.

1. INTRODUCTION

1.1. Computer and Sensor Nodes Literature for Energy Efficiency

When minimizing the total energy, it is fundamental to consider, besides the energy spent on transmission purposes, non-transmission energies. In particular, the energy cost incurred with the state of the channel being 'on', appears to constitute an important fraction of the total energy expended in wireless devices. Hence, this energy cost cannot be neglected.
In the computer and sensor nodes literature, various techniques have been proposed to reduce the mobile host's power consumption during operation. Recognizing the fact that "when inserted, many wireless communication devices consume energy continuously" and that "this energy consumption can represent over 50% of total system power for current handheld computers and up to 10% for high-end laptops", Kravertz et al. proposed in [7] software-level techniques to suspend the mobile host's device during idle periods of the communication. In [13], Chandrakasan et al. studied power-aware techniques to minimize power consumption of wireless microsensor systems.
At the intersection between the communication theory and the networking fields, El Gamal et al. proposed an optimal scheduling algorithm to minimize transmission energy by maximizing the transmission time for buffered packets, [2]. In [10], Rulnick and Bambos studied mobile power management for maximum battery life in wireless communication networks. In [1], Cui et al. considered wireless applications, where nodes operate on batteries, and analyzed the best modulation strategy to minimize the total energy consumption, when error-control codes are used. In [3], the authors analyzed the best modulation strategy to minimize the total energy consumption, while satisfying throughput and delay requirements. Various other interesting power-aware components and algorithms for wireless networks can be found in ([9, 8, 11, 12]).
The information theoretic literature has also extensively considered the energy cost of being 'on'. In particular, Verdú investigated in [5], the minimum cost incurred by the transmission of one bit of information through a noisy channel, characterizing the most economical way to communicate reliably. This approach is well suited for continuous changes and presents very general results. Applying those to include a cost associated with the 'on' state, would not lead, however, to continuous results since, in this case, a step function would be involved.
In this paper, we take explicitly into account the energy of being 'on', which we term 'processing energy', in ad-

This research is supported by HP, Advanced Concepts in Wireless Networking, Research Alliance Agreement Award 008542-008 and NSF NRT Award ANI-0335217.

dition to the energy used for transmission purposes. In the presence of an additional cost associated with the 'on' state, achieving capacity may require intermittent transmission, which we term 'bursty'. We derive the optimal burstiness of signaling as well as the optimal amount of energy that should be transmitted while in the 'on' state.
We neglect the energy cost associated with the transitioning to the 'on' state; the effect of state transition energy cost would only be manifested if a delay metric, such as an error exponent, was used.

1.2. Background

The classical information theory problem of characterizing capacity for an Additive White Gaussian Noise (AWGN) channel under an energy constraint strives to answer the question: How much information can be transmitted through a channel, and how?
This traditional approach considers that all the available energy without inefficiency or overhead is consumed as radiated energy for transmission. Owing to the concavity of $\log(1+x)$, transmission should have constant energy. Unlike the case when transmission accounts for all the expended energy, we take explicitly into account the processing energy. We show that the processing energy may lead, instead, to making bursty transmission capacity achieving. We consider a bandlimited AWGN channel, as shown in figure 1, with one sender and one receiver.

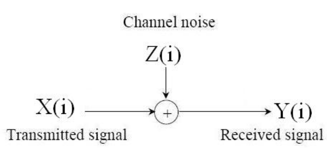

Figure 1: AWGN channel

For this channel, we use a sampled time that we denote by the variable i.
The input signal is constrained in energy and in bandwidth and is distorted by additive and bandlimited white Gaussian noise. The noise samples are independent and identically distributed Gaussian random variables with zero-mean and variance σ^2: $Z(i) \sim \mathcal{CN}(0, \sigma^2)$.
The output of the channel at time i is given by:

$$Y(i) = X(i) + Z(i). \tag{1}$$

The remainder of the paper is organized as follows. In Section 2, we modify the energy constraint of an AWGN channel to include the processing energy. Then, we derive the capacity of the channel in terms of the optimal transmitted energy and the optimal burstiness of signaling. In Section 3, we determine the conditions that make burstiness capacity achieving, in terms of the total available energy and the processing energy. Section 4 summarizes our conclusions and presents our future work. Finally, a review of the Lambert$W$ function can be found in the Appendix.

2. INCLUSION OF PROCESSING ENERGY IN ENERGY CONSTRAINT

2.1. Capacity in Terms of Transmission Energy

Let us first review the capacity of an AWGN channel [6], with energy constraint $\mathcal{E}$, noise variance σ^2 (product of bandwidth and noise spectral density), and under no processing energy:

$$C = \log(1 + \frac{\mathcal{E}}{\sigma^2}). \tag{2}$$

The energy constraint is given by:

$$\frac{\sum_{i=1}^{n} P_i}{n} \leq \mathcal{E}, \tag{3}$$

where $P_i = \overline{X^2(i)}$ and $X(i)$ is the input to the channel at time i.
In what follows, we let 1_i be the indicator function taking the value 0 or 1 according to whether at time i the transmitter is sending or not, P_i be the actual energy used for transmission at time i and $\mathcal{E}$ be the total energy available per channel use.
When the processing energy is taken into account and is given by ϵ, the system that we seek to solve becomes:

$$\max \frac{1}{n} \sum_{i=1}^{n} 1_i \log(1 + \frac{P_i}{\sigma^2}) \tag{4}$$

$$\text{s.t. } \frac{\sum_{i=1}^{n} 1_i (P_i + \epsilon)}{n} \leq \mathcal{E}. \tag{5}$$

Remarks:

- We are considering arbitrarily large time n while maximizing the sum in (4).
- For a fixed n, this is equivalent to dividing the channel into n subchannels with power constraints $P_1 ... P_n$ respectively and assuming independent coding over the subchannels.
- In writing (4), we assumed that the transmitter and the receiver have agreed on the time at which transmission should occur.

- Setting ϵ to 0, (5) leads back to the constraint in (3) and the corresponding maximized value of (4) would be given by (2).

When we transmit, i.e. when $1_i = 1$, the concavity of the function $\log(1+x)$ implies that P_i should, at any time i for which $1_i = 1$, be equal to a constant, say ν. We denote the *burstiness of signaling* by

$$\Theta \hat{=} \frac{1}{n}\sum_{i=0}^{n} 1_i. \tag{6}$$

When $\Theta = 1$, a constant transmission strategy is being used; when $\Theta < 1$, bursty signals are being transmitted. The smaller the Θ, the burstier the transmission. Given a certain total amount of energy $\mathcal{E}$, how would an increase of the processing energy, ϵ, affect the optimal transmission mode? Clearly, since the optimal strategy should avoid paying too much overhead owing to the processing energy cost, sending more bursty signals would result in higher rates. On the other hand, transmission should not be too bursty, since putting all the energy in one slot may result in a loss of rate. Therefore, there is a certain tradeoff between sending bursty signals and adopting a continuous transmission strategy. This tradeoff is controlled by the values of $\mathcal{E}$ vs. ϵ.
Back to the constraint in (5), since additional total energy can only be beneficent, we take the constraint to equality:

$$\Theta(\nu + \epsilon) = \mathcal{E}. \tag{7}$$

Hence, we denote (4) in terms of ν by,

$$C_1(\nu) \hat{=} \frac{\mathcal{E}}{\nu+\epsilon}\log(1+\frac{\nu}{\sigma^2}). \tag{8}$$

Taking the derivative with respect to ν and setting it to 0, we obtain the optimal value of the energy for which the capacity is maximized in the presence of processing energy

$$\nu_{opt} = \frac{\epsilon-\sigma^2}{W(\frac{\epsilon-\sigma^2}{e\sigma^2})} - \sigma^2. \tag{9}$$

As seen in the expression of ν_{opt}, we have used the Lambert$W$ function which we denoted by W. Refer to the Appendix for a brief review of the function.
Note that the optimal value of the transmission energy is independent of the total available energy $\mathcal{E}$ and is merely dependent on the processing energy ϵ.
The maximum value of (4) that can be achieved in the presence of processing energy is

$$C_1(\nu_{opt}) = \frac{\mathcal{E}}{\nu_{opt}+\epsilon}\log(1+\frac{\nu_{opt}}{\sigma^2}), \tag{10}$$

where ν_{opt} is given by (9). Refer to 2.2 for discussion of second derivative constraint. There are two constraints on the value of the energy,

$$\nu \geq 0 \tag{11}$$

$$\nu \geq \mathcal{E} - \epsilon \text{ from } \Theta \leq 1. \tag{12}$$

For unit variance, we can show that $\nu_{opt} \geq 0$ for all $\epsilon \geq 0$. The proof is omitted for the sake of brevity.
For the second constraint, however, we should find the values of ϵ and $\mathcal{E}$ for which $\nu_{opt} \geq \mathcal{E} - \epsilon$.
Substituting for ν_{opt} from (9) and simplifying, we obtain the equivalent constraint in terms of $\mathcal{E}$ and ϵ, i.e.

$$\mathcal{E} \leq \frac{(\epsilon-1)(1+W(\frac{\epsilon-1}{e}))}{W(\frac{\epsilon-1}{e})}. \tag{13}$$

Failing to meet this condition would lead to having optimality when $\nu_{opt} = \mathcal{E} - \epsilon$.
For instance, in the case of $\mathcal{E} = 1$, it can be shown that $\nu_{opt} \geq \mathcal{E} - \epsilon$ for all $\epsilon >\approx 0.237$. In that case, the value in (9) achieves capacity. Again, we omit the proof for the sake of brevity.

2.2. Capacity in Terms of the Burstiness of Signaling

We can also maximize the capacity with respect to Θ. Hence, using (7), we denote (4) by:

$$C_2(\Theta) \hat{=} \Theta\log(1+\frac{\frac{\mathcal{E}}{\Theta}-\epsilon}{\sigma^2}). \tag{14}$$

Note that, although different functions, C_2 above and C_1 in (10) are describing the same entity, namely capacity in the presence of processing energy, but in terms of different variables. Calculating the derivative with respect to Θ, then, setting it to 0, we obtain the corresponding Θ_{opt} that maximizes the capacity in the presence of processing energy.

$$\Theta_{opt} = \frac{\mathcal{E}W(\frac{\epsilon-\sigma^2}{\sigma^2\times e})}{(\epsilon-\sigma^2)(W(\frac{\epsilon-\sigma^2}{\sigma^2\times e})+1)}. \tag{15}$$

For the remainder of the paper, we assume $\sigma^2 = 1$; then we refer to $\mathcal{E}$ by the Signal to Noise Ratio (SNR) and to ϵ by the Normalized Cost (NC), or by the 'processing energy' as above. Then, Θ_{opt} can be reexpressed merely in terms of the SNR and the NC by:

$$\Theta_{opt} = \mathcal{E}\frac{W(e^{-1}(\epsilon-1))}{(\epsilon-1)(W(e^{-1}(\epsilon-1))+1)}. \tag{16}$$

Note that the optimal burstiness of signaling is directly proportional to the SNR.

Next, we would like to know how Θ_{opt}/SNR varies with the NC. Figure 2 shows the plot of $\log_{10}(\Theta_{opt}/\mathcal{E})$ vs. ϵ. As illustrated in figure 2, for a given SNR, the burstiness of signaling decreases with increasing NC. This agrees with our intuition, since the higher the cost of being 'on', the lesser the fraction of time we want to spend transmitting.

From the definition of Θ in (6), we should satisfy

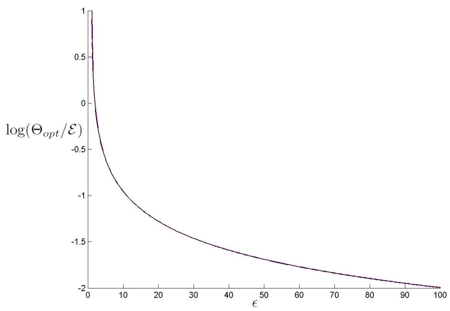

Figure 2: Plot of $\log_{10}(\Theta_{opt}/\mathcal{E})$ vs. ϵ

$0 \leq \Theta_{opt} \leq 1$.
We can show that $\Theta_{opt} \geq 0$ for all $\epsilon \geq 0$. We omit the proof for the sake of brevity. Next, we seek the values of $\mathcal{E}$ and ϵ for which $\Theta_{opt} \leq 1$. Let us illustrate the effect of these constraints when $\mathcal{E} = 1$ and check the conditions on ϵ. As seen in 2.1, for $\epsilon \geq\approx 0.237$, $\Theta_{opt} \leq 1$ and the optimal value is Θ_{opt} as given in (15), otherwise $\Theta_{opt} = 1$.
Given certain values of ϵ and $\mathcal{E}$, we should calculate the second derivatives of (10) and (14) and verify that they are non-positive when evaluated at $\nu = \nu_{opt}$ and $\Theta = \Theta_{opt}$ respectively, otherwise corresponding values of the boundaries would be biting, hence maximizing the sum in (4).
Finally, this example shows that, unlike the case where no processing energy is considered, Θ_{opt} is not always equal to 1. Therefore, burstiness is, in some cases, capacity achieving. We investigate the conditions of such a transmission mode in the following section.

3. BURSTY TRANSMISSION

In this section, we determine when achieving the capacity requires bursty transmission. For this purpose, we would like to find a relationship between the SNR ($\mathcal{E}$) and the processing energy (ϵ) that allows the transition from the bursty transmission regime to the constant transmission one. This transition occurs at $\Theta_{opt} = 1$.

Recall the expression of Θ_{opt} in (15). Then,

$$\Theta_{opt} = 1 \Leftrightarrow \mathcal{E} = \frac{(\epsilon - 1)(W(\frac{\epsilon-1}{e}) + 1)}{W(\frac{\epsilon-1}{e})}. \tag{17}$$

We consider the curve $(\mathcal{E}, \epsilon)$ which separates the regions in which burstiness and constant transmission become capacity achieving. As seen in figure 5, the region below the function $\Theta_{opt} = 1$ denotes the region where we would like to be bursty. The region where $\Theta_{opt} = 1$ is where continuous transmission is appropriate and the processing energy can be neglected. We are then back to the original idealized model of wireless channels where there is no need to consider the cost of being 'on'. Next, we would like to find numerical bounds between these two models.
Recognizing that,

$$\lim_{x \to -e^{-1}} W(x) = -1,$$

$$\lim_{x \to 0} W(x - e^{-1}) = -1$$

and

$$\lim_{x \to 0} \frac{W(x - e^{-1}) + 1}{\sqrt{x}} = \sqrt{2e}, \tag{18}$$

we can approximate the expression in (17) as $\epsilon \to 0$, or $\Theta \to 0$ as follows.
Let $y = ex$. Use change of variables on (18) to obtain:

$$\lim_{y \to 0} W(e^{-1}(y - 1)) - \sqrt{2y} = -1, \tag{19}$$

Therefore, (19) shows that for arbitrarily small ϵ, we can substitute $W(e^{-1}(\epsilon - 1))$ by $(\sqrt{2\epsilon} - 1)$; (17) can then be approximated by:

$$\mathcal{E} \approx \frac{\sqrt{2\epsilon}(\epsilon - 1)}{\sqrt{2\epsilon} - 1}. \tag{20}$$

(20) can be furthermore approximated by using:

$$\lim_{\epsilon \to 0} \frac{\sqrt{2\epsilon}(\epsilon - 1)}{\sqrt{2\epsilon} - 1} - \sqrt{2\epsilon} = 0. \tag{21}$$

Therefore, for low SNR we obtain:

$$\mathcal{E} \approx \sqrt{2\epsilon}. \tag{22}$$

We illustrate in figure 3, the equation in (17) and its approximation in (22).
In general, (15) and (21) give:

$$\mathcal{E} \approx \Theta_{opt}\sqrt{2\epsilon}. \tag{23}$$

Replacing $\Theta_{opt} \leq 1$ in (23), we obtain

$$\mathcal{E} \leq \sqrt{2\epsilon}. \tag{24}$$

The above equation shows that burstiness is capacity achieving when the total energy available to the channel (/SNR) is less than the square root of double the processing energy (/NC). This fact is illustrated in figure 5 for small ϵ.
Moreover, the slope of (22) at $\epsilon = 0$ goes to ∞. Therefore, in the low SNR regime, when the processing energy is *any percentage* of the total energy, burstiness is capacity achieving.

On the other hand, to remain non-bursty, ϵ and $\mathcal{E}$ should satisfy the relationship: $\epsilon \leq \frac{\mathcal{E}^2}{2}$. Under this constraint, constant transmission over the channel is optimal.

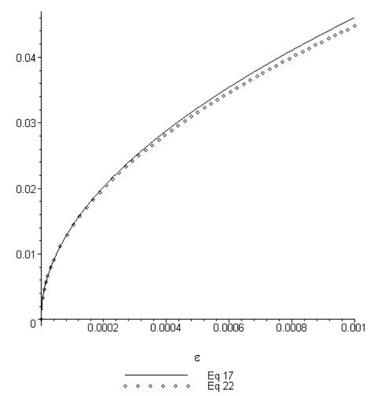

Figure 3: Bound for low SNR regime

For large SNR, when $\epsilon \to \infty$,

$$\frac{d\mathcal{E}}{d\epsilon} = 1 + \frac{1}{W(\frac{\epsilon-1}{e})} - \frac{1}{W(\frac{\epsilon-1}{e})(1 + W(\frac{\epsilon-1}{e}))} \to 1. \quad (25)$$

Thus, the slope of the curve that separates the bursty transmission regime from the non-bursty region tends to 1. In figure 4, we plot $\frac{d\mathcal{E}}{d\epsilon}$ as a function of ϵ.
Therefore, for large SNR, given a certain ratio $\frac{\mathcal{E}}{\epsilon}$, we can directly deduce the region in which we are situated in figure 5 and decide whether burstiness is optimal or not. For instance, for $\epsilon = 0.01\mathcal{E}$, we are clearly above the curve $(\mathcal{E}, \epsilon)$, hence we want to have continuous transmission. This, of course, agrees with our intuition that for a small cost of being 'on', processing energy can be neglected and transmission should remain constant.

4. CONCLUSIONS AND FUTURE WORK

We study the optimal strategy in using the total energy when the processing energy of the transmitter being 'on' is also taken into account. We derive the capacity of an AWGN channel in terms of the total energy constraint $\mathcal{E}$ and the processing energy ϵ. We find the optimal value of transmission energy (ν_{opt}) for which capacity is maximized. We also calculate the optimal burstiness of signaling(Θ_{opt}) at which channels should be transmitting. We distinguish two modes for transmission: constant transmission and bursty transmission. We describe the conditions that render bursty transmission capacity achieving in terms of the total energy available to the input and the processing energy.
In low SNR regime, we prove that capacity achieving requires burstiness when the processing energy is greater than half the square of the total energy. In particular, if the processing energy is any percentage of the total available energy, burstiness achieves capacity.
In high SNR, we show that comparing the ratio $\frac{\mathcal{E}}{\epsilon}$ to the slope of the curve $(\mathcal{E}, \epsilon)$, which tends to 1 at infinity, would lead to making a decision about the transmission mode.
Future work includes extending the results obtained to an AWGN channel shared by multiple users in the presence of processing cost. Moreover, delay or transitioning costs can also be taken into account. In

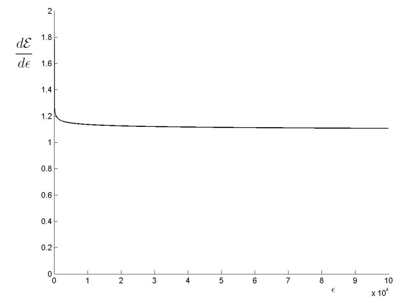

Figure 4: Slope Bound

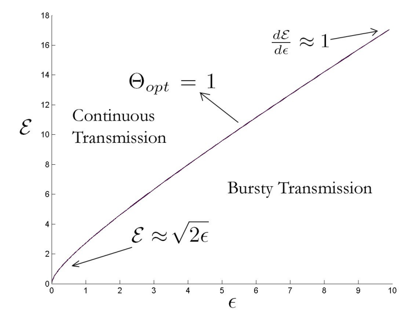

Figure 5: Summary of main approximations

this case, a delay metric such as error exponents would need to be considered.

APPENDIX

The Lambert$W(x)$ function [4], also called the Omega function, is the inverse function of $f(w) = w \times e^w$ for complex numbers w; This means that for every complex number x, we have

$$\text{Lambert}W(x) \times e^{\text{Lambert}W(x)} = x. \qquad (26)$$

Since the function f is not injective in $(-\infty, 0)$, the Lambert$W$ is multivalued in $[-\frac{1}{e}, 0)$.
For $x \geq -e^{-1}$, the funtion $W(x)$ is real and single valued.
For example, $W(0) = 0$ and $W(-e^{-1}) = -1$.

Furthermore, the plot of $W(x)$ as a function of x, for $x \geq -e^{-1}$, is shown in figure 6.

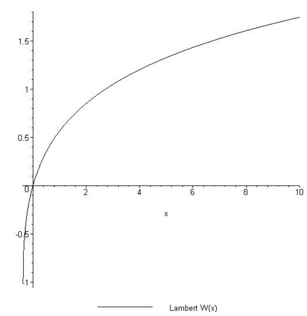

Figure 6: W(x) as a function of x

References

[1] S. Cui, A.J. Goldsmith, and A. Bahai: "Energy-constrained Modulation Optimization for Coded Systems," In *Globecom*, 2003.

[2] A.El Gamal, C. Nair, B. Prabhakar, E. Uysal-Biyikoglu, and S. Zahedi: "Energy-efficient scheduling of packet transmissions over wireless networks," In *Proc. IEEE Infocom*, 2002.

[3] S. Cui, A.J. Goldsmith, and A. Bahai: "Energy-efficiency of MIMO and Cooperative MIMO Techniques in Sensor Networks," *IEEE journal on selected areas on communications*, 2004.

[4] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth: "On the Lambert W Function," In *Advances in Computational Mathematics*, vol.5, pp. 329–359, 1996.

[5] S. Verdú: "On Channel Capacity per Unit Cost," In *IEEE Trans. on Info. Theory*, vol.36, no.5, 1990.

[6] T.M. Cover and J.A. Thomas: *Elements of Information Theory*, New York: Wiley, 1991.

[7] R. Kravertz and P. Krishman: "Power Management Techniques for Mobile Communication," In *Proc. of The Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking*, 1998.

[8] P. Agrawal, J-C. Chen, S. Kishore, P. Ramanathan, and K. Sivalingam: "Battery Power Sensitive Video Processing in Wireless Networks," In *Proc. IEEE PIMRC*, 1998.

[9] P.J.M. Havinga, G.J.M. Smit: "Minimizing Energy Consumption for Wireless Computers in Moby Dick," In *Proc. IEEE International Conference on Personal Wireless Communication*, 1997.

[10] J.M. Rulnick and N. Bambos: "Mobile power management for maximum battery life in wireless communication networks," In *Proc. of IEEE Infocom 96*, 1996.

[11] A. Sampath, P.S. Kumar, and J. Holtzman: "Power control and resource management for a multimedia cdma wireless system," In *The Sixth International Symposium on Personal, Indoor and Mobile Radio Communications*, 1995.

[12] M. Stemm and R. Katz: "Reducing power consumption of network interfaces in hand-held devices," In *Third International Workshop on Mobile Multimedia Communications*, 1996.

[13] A. Chandrakasan, R. Min, M. Bhardwaj, S-H Cho, and A. Wang: "Power aware wireless microsensor systems," *Keynote Paper ESSCIRC*, 2002.