Communication by Sleeping: Optimizing a Relay Channel Under Wake and Transmit Power Costs

Vinod Prabhakaran and P. R. Kumar
Coordinated Science Laboratory
University of Illinois, Urbana-Champaign
Urbana, IL 61801

***Abstract*—In many low-power networks, the power cost for a node to remain ON to listen to transmissions from other nodes or to transmit to other nodes can constitute a significant part of the total power consumption by the radio. Thus, unlike the traditional relay channel model, under a low power constraint, the relay node cannot stay ON and listen to the entire duration of the transmission. We study the slope of the capacity-power function at zero power for relay channels where the power cost of remaining ON is explicitly taken into account. We show that in this low-power limit, information is conveyed primarily through the ON-OFF activity of the source. The relay node should have significantly more power compared to the source node in order to improve the slope at zero power. The rough intuition is that, in order for a relay node to be useful, it needs to stay ON and listen during a significant portion of the time during which the source node may transmit. This fact limits the utility of the relay to those cases where the it has much higher power than the source node it is trying to help.**

I. Introduction

In low-power networks (like some sensor networks), the power cost for a node to remain ON to listen to transmissions from other nodes or to transmit to other nodes can constitute a significant part of the total power consumption by the radio – typical sleep power consumption is of the order of microwatts while awake power consumption is of the order of milliwatts [1]. Thus, unlike the traditional relay channel model, under a low power constraint, the relay node cannot stay ON and listen to the entire duration of the transmission. In fact, sharply reducing power consumption is a key design objective for such sensor networks when the goal is a long lifetime for environmental monitoring. In this note we consider AWGN relay channels where the cost of remaining ON is explicitly accounted for. In particular, we will assume that whenever a relay or transmitting node is either transmitting or listening to a transmission, it incurs a fixed power cost of say Υ (watts). This is in addition to the power cost of transmission which is measured as usual. For an alternative formulation which takes processing energy into account, something we do not do here, see [2].

Let us first consider the point-to-point problem with the above cost function on the transmitter[1]. This problem was considered by Youssef-Massaad, Zheng, and Médard [3] when the transmitter is not allowed to convey information through timing, *i.e.*, the ON-OFF activity of the transmitter is not allowed to depend on the message and is the same for every codeword. Then the maximum achievable rate for power constraint P and an AWGN channel with a noise variance of unity is[2]

$$R(P) = \max \frac{\theta}{2} \log \left(1 + \left(\frac{P}{\theta} - \Upsilon\right)\right),$$

where the maximization is over $0 \leq \theta \leq \min(1, P/\Upsilon)$. Here, θ is the proportion of time the transmitter is ON. Youssef-Massaad et al [3] investigated the above optimization problem. More generally, the following facts are easy to see:

- High-SNR: The capacity $C(P)$ of the channel when the transmitter is allowed to send information through timing as well is larger than $R(P)$ by at most 1 bit/sample since the transmitter's ON-OFF state can carry at most 1 bit.
- Low-SNR: When the transmitter is prevented from sending information through timing, as in [3], the slope of capacity at $P = 0$ is 0, *i.e.*,

$$\lim_{P \downarrow 0} \frac{R(P)}{P} = 0.$$

However, when the transmitter is allowed to have its ON-OFF activity depend on the message, we can show that the slope of capacity at $P = 0$, *i.e.*, $\lim_{P \to 0} C(P)/P$ is the same as that for the regular AWGN problem where we do not impose any extra cost for remaining ON, *i.e.*,

$$\lim_{P \downarrow 0} \frac{C(P)}{P} = \frac{1}{2}.$$

One scheme which can achieve this is a pulse-position modulation (PPM) scheme where the transmitter conveys information by the time at which it sends a high amplitude

This material is based upon work partially supported by USARO under Contract Nos. W911NF-08-1-0238 and W-911-NF-0710287, and NSF under Contract Nos. ECCS-0701604, CNS-07-21992, CNS-0626584, and CNS-05-19535.

[1]It is easy to see that the problem where there is also a fixed cost on the receiver when it is turned ON to listen can be reduced to the above problem. It can be shown that there is no gain in randomizing the schedule of operation of the receiver and, without loss of optimality, the schedule can be made known to the transmitter in advance. This reduces the problem to one where the receiver is always ON and the cost function of the transmitter alone depends on its ON activity.

[2]We will use nats and the natural logarithm throughout unless otherwise stated.

pulse and remains OFF for the rest of the transmission; the receiver can be a simple energy detector. This scheme is along the lines of Kennedy [4]. The effect of imposing this cost function at low SNRs is to force the transmission strategy towards schemes where the transmitter remains on for very short periods of time and the information is conveyed exclusively through the timing of transmission rather than the amplitude.

This fact motivates us to ask how the slope of the capacity of an AWGN relay channel behaves at $P = 0$ when the cost of remaining ON is also taken into account[3]. In this short note, we are interested in finding out under what conditions (and how) the relay node can make a difference to the slope of capacity at $P = 0$ when the cost of remaining ON to listen and transmit is also taken into consideration.

For relay channels with wake power costs, the fact that, in general, allowing the activity of the nodes to depend on the message can lead to improved rates of operation was demonstrated in a related work by Kramer [5]. For relay channels with a cost function which depends on the sleep-listen-or-talk activity of the nodes, achievable rates using decode-and-forward and partial-decode-and-forward schemes where the activity of the nodes may depend on the message were considered in the paper. In contrast, optimal strategies in the low power regime of interest in this paper will turn out to be amplify-and-forward.

We show that when the relay node has comparable or lower power than the source node, the relay node cannot improve the slope of the capacity at $P = 0$ (compared to the point-to-point (no-relay) scenario). The relay node should have a significantly higher power compared to the source node, in a sense which we make clear, to improve the slope of the capacity at $P = 0$. The rough intuition is that, at low power, the information is always conveyed over timing rather than amplitude. A relay node obviously cannot anticipate exactly when the source node it is helping will be ON, since the ON-OFF activity of the source also carries the message from the source. Hence, for the relay node to be useful, it needs to stay ON and listen during a significant portion, if not the entire duration, during which the source node may transmit. This limits the utility of the relay to only those cases where the relay node has significantly more power than the node it is trying to help.

II. Problem Statement and Main Result

A. Channel Model

Consider the three-node network in Fig. 1. Nodes 1 and 2 have information they want to communicate to node 3 which has no transmit capabilities. Nodes 1 and 2 may help each other communicate. We assume that the channels are all

[3]Note that the slope of the capacity of an AWGN relay channel under certain settings has been investigated (without the ON-cost, *i.e.*, when $\Upsilon = 0$) by El Gamal et al [6].

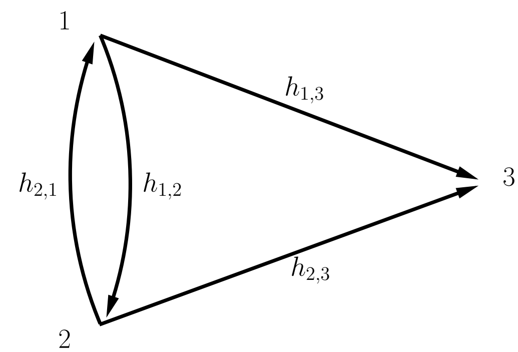

Fig. 1: Problem Setup. Nodes 1 and 2 have information they want to communicate to node 3 which has no transmit capabilities.

additive white Gaussian noise channels:

$$Y_k(i) = A_k(i)\left(\sum_{j\neq k} h_{jk}X_j(i) + N_k(i)\right),$$

where $X_j(i)$ and $Y_k(i)$ are, respectively, what node j transmits and node k receives at discrete time $i \in \{1, 2, \dots\}$. Also, h_{jk} is the channel co-efficient (a real number, assumed non-zero) of the channel from node j to node k, and the noise $N_k(j)$ is i.i.d. Gaussian of unit variance. $A_k(i) \in \{0, 1\}$ is a random variable which indicates the *activity* of node k at time i which we will define presently. At each time i, node $k \in \{1, 2\}$ chooses to be either ON or OFF based on the message W_k it wants to send and everything it has heard so far (Y_k^{i-1}). When node k is ON at time i, we indicate that by $A_k(i) = 1$; otherwise $A_k(i) = 0$. Also, we set $A_3(i) = 1$ for all i; *i.e.*, we assume that the receiver is always in the ON state. We consider both half- and full-duplex modes of operation for the nodes 1 and 2. In the *half-duplex mode*, at each time i, these nodes may either transmit or receive, but not do both simultaneously; whereas in *full-duplex mode* they may simultaneously listen and transmit. In full-duplex mode, $A_k(i) = 0$ and $X_k(i)$ must be 0 if the node chooses to be in the OFF state at time i; $A_k(i) = 1$ in the ON state. The power constraint is

$$\frac{1}{n}\sum_{i=1}^{n} A_k(i)\Upsilon + X_k^2(i) \le P_k,$$

where n is the blocklength and P_k is the power available to node k. In *half-duplex mode*, when a node chooses to be in the ON state, it also decides to be in one of either LISTEN or SEND states (again based on its message and everything it has heard so far). $A_k(i) = 1$ if the node is in ON and LISTEN state, and 0 otherwise. $X_k(i)$ is required to be 0 unless the node is in ON and SEND state. The power constraint is

$$\frac{1}{n}\sum_{i=1}^{n} B_k(i)\Upsilon + X_k^2(i) \le P_k,$$

where $B_k(i) \in \{1, 0\}$ indicates whether the node is in ON ($B_k(i) = 1$) or OFF ($B_k(i) = 0$) state.

We consider blocklength n codes defined by functions $f_{k,i}, g_{k,i}^A, g_{k,i}^B$, such that $X_k(i) = f_{k,i}(W_k, Y_k^{i-1})$, and also $A_k(i) = g_{k,i}^A(W_k, Y_k^{i-1})$, and $B_k(i) = g_{k,i}^B(W_k, Y_k^{i-1})$ and the power constraints above are satisfied. Definition of achievable pairs of rates (R_1, R_2) and the rate region is as usual.

We consider a scaling of the powers motivated by the fact that the nodes could have widely different available powers (say, a remote sensor node versus a wired relay node). In this paper we consider $P_1 = P^{\alpha_1}$ and $P_2 = P^{\alpha_2}$ with $0 < \alpha_1 \leq \alpha_2$. When $\alpha_1 = \alpha_2$, we investigate the scaling $P_1 = P, P_2 = \gamma P$. We will be interested in $\frac{R_k(P)}{P^{\alpha_k}}$ in the limit of $P \to 0$. Investigating this scaling is justified by the easy-to-see fact that $\frac{1}{R_k(P)} = \Theta\left(\frac{1}{P^{\alpha_k}}\right)$.

B. Main Result

(a) $\alpha_1 < \alpha_2$*:*
There is a sequence of half-duplex encoders and decoders (one set for each P) which can (simultaneously) achieve $(R_1(P), R_2(P))$ such that

$$\lim_{P \downarrow 0} \frac{R_1(P)}{P^{\alpha_1}} = \frac{h_{13}^2}{2},$$
$$\lim_{P \downarrow 0} \frac{R_2(P)}{P^{\alpha_2}} = \frac{h_{23}^2 + h_{21}^2}{2}.$$

Moreover, for any sequence of (half- or full-duplex) encoders and decoders, the following two bounds apply if $\Upsilon > 0$:

$$\limsup_{P \downarrow 0} \frac{R_1(P)}{P^{\alpha_1}} \leq \frac{h_{13}^2}{2},$$
$$\limsup_{P \downarrow 0} \frac{R_2(P)}{P^{\alpha_2}} \leq \frac{h_{23}^2 + h_{21}^2}{2}.$$

Thus, node 1 (which has significantly more power than node 2, since $\alpha_1 < \alpha_2$) can in effect act as a second receiver antenna for the transmission from node 2 without affecting its own transmission. However, node 2 cannot improve by any amount the slope of $R_1(P)$ for node 1's transmission. As we explain in the next section, our scheme is a PPM scheme with only node 1 acting as a relay for node 2's transmission. When acting as the relay, node 1 performs amplify-and-forward. The transmissions from the two nodes are orthogonalized through time-division multiplexing.

(b) $P_1 = P, P_2 = \gamma P$ *(so* $\alpha_1 = \alpha_2$*):*
For any sequence of (half- or full-duplex) encoders and decoders, if $\Upsilon > 0$,

$$\limsup_{P \downarrow 0} \frac{R_k(P)}{P} \leq \frac{h_{k3}^2}{2}, \ k = 1, 2.$$

A simple point-to-point (no relaying) scheme with the transmissions from the two sources orthogonalized through time-division multiplexing achieves this.
Thus, when the powers of the nodes are comparable, they are unable to help each other improve the slopes of the achievable rates at $P = 0$. The intuition is that, the relays do not have enough energy to stay awake to listen to significant enough parts of the other node's transmission to help.

III. Proof Sketch

For ease of notation, we prove our result below for the case of $h_{jk} = 1$. The extension to the general case is straightforward.

A. Upper bounds

a) $\alpha_1 < \alpha_2$*:* For the upper bounds, we allow the nodes to act in full-duplex mode. We apply cut-set upper bounds.
Node 1's rate: For the transmission of node 1's message, we will let nodes 2 and 3 cooperate. We will also let node 2 listen for the maximum duration of time possible which is P^{α_2}/Υ. It is easy to see that this gives the following upper bound on R_1:

$$R_1(P) \leq \max \frac{1 - (P^{\alpha_2}/\Upsilon)}{2} \log(1 + p_1) + \frac{(P^{\alpha_2}/\Upsilon)}{2} \log(1 + 2p_2),$$

where the maximization is over $p_1, p_2 \geq 0$ subject to

$$p_1(1 - (P^{\alpha_2}/\Upsilon)) + p_2(P^{\alpha_2}/\Upsilon) \leq P^{\alpha_1}.$$

Hence,

$$\begin{aligned} R_1(P)/(P^{\alpha_1}) \leq & \frac{1 - (P^{\alpha_2}/\Upsilon)}{2} \frac{\log(1 + \frac{P^{\alpha_1}}{1-(P^{\alpha_2}/\Upsilon)})}{P^{\alpha_1}} \\ & + \frac{P^{\alpha_2}/\Upsilon}{2} \frac{\log(1 + 2\frac{P^{\alpha_1}}{P^{\alpha_2}/\Upsilon})}{P^{\alpha_1}} \\ & \to 1/2 \end{aligned}$$

in the limit as $P \downarrow 0$.
Node 2's rate: Here we will use a trivial upper bound where node 1 cooperates with the receiver node 3, and no operating cost is incurred by node 1. *i.e.*, we let node 1 remain active for the whole duration. Now

$$R_2(P)/(P^{\alpha_2}) \leq \frac{1}{2} \frac{\log(1 + 2P^{\alpha_2})}{P^{\alpha_2}} \to 1,$$

as $P \downarrow 0$.

b) $P_1 = P, \ P_2 = \gamma P$*:* For the transmission of node 1, we will let nodes 2 and 3 cooperate. Also, node 2 will stay ON for the longest proportion of time, which is $P_2/\Upsilon = \gamma P/\Upsilon$. Moreover, no cost is imposed on the transmitter. It is easy to see that the following is an upper bound on the rate of transmission from node 1.

$$R_1(P) \leq \max \frac{(\gamma P/\Upsilon)}{2} \log(1 + 2p_1) + \frac{1 - (\gamma P/\Upsilon)}{2} \log(1 + p_2),$$

where the maximization is over $p_1, p_2 \geq 0$ subject to

$$p_2(1 - (\gamma P/\Upsilon)) + p_1(\gamma P/\Upsilon) \leq P.$$

Since $0 \leq p_1 \leq \Upsilon/\gamma$, let us define: $p = \Upsilon/\gamma - p_1$, and

$$\begin{aligned} R_1(P,p)/P \stackrel{\text{def}}{=} & \frac{\gamma/\Upsilon}{2} \log\left(1 + 2\left(\frac{\Upsilon}{\gamma} - p\right)\right) \\ & + (1/P - \gamma/\Upsilon) \log\left(1 + \frac{\gamma P p}{\Upsilon(1 - \gamma P/\Upsilon)}\right). \end{aligned}$$

We can show that for $p \in [0, \Upsilon/\gamma]$,

$$\frac{R_1(P,p)}{P} \leq \frac{\gamma}{2\Upsilon}\left(\log(1+2(\Upsilon/\gamma)-2p)+p\right)+2(\gamma/\Upsilon)P$$
$$\leq 1/2+2(\gamma/\Upsilon)P.$$

Thus, we can conclude that

$$\lim_{P\downarrow 0}\frac{R_k(P)}{P} \leq 1/2.$$

B. Achievability

a) $\alpha_1 > \alpha_2$: The upperbound of $\lim_{P\downarrow 0} R_1(P)/P^{\alpha_1} \leq 1/2$ is trivially achieved without relaying. Node 2 may shut off during the time node 1 transmits to node 3. Now we will show how $\lim_{P\downarrow 0} R_2(P)/P^{\alpha_2} = 1$ can be achieved simultaneously by demonstrating that node 1 needs to invest only a negligible part of its available energy to help relay node 2's message, and the transmissions of the two messages can be done through orthogonalization (time-division multiplexing). This second fact is hardly surprising since the regime of operation is power-limited and not degrees-of-freedom limited. We use a pulse-position modulation (PPM) scheme where the nodes act in half-duplex.

Let us consider a block of T symbols. Over this block, node 2 will remain powered off during symbols $S+1, S+2, \ldots, T$, *i.e.*, node 2 transmits only during the first S symbols. Node 2 sends its message using PPM: in particular, it sends $(a, 0, 0, \ldots, 0)$ for message 1, $(0, a, 0, \ldots, 0$ for message 2 and so on, where $a = \sqrt{P^{\alpha_2}T - \Upsilon}$ is the amplitude of the pulse. Since there are S such codewords, the rate of transmission is $R_2 = (\log S)/T$. We let node 1 listen to the transmission from node 2 by staying ON during the entire duration in which node 2 may transmit. Let us define β to be such that

$$\Upsilon S = P^{\beta}T,$$

where we note that the left-hand-side is the amount of energy expended by node 1 by staying ON to listen to node 2's transmission. We will constrain S such that $\alpha_1 < \beta < \alpha_2$. Node 1's strategy is amplify-and-forward, *i.e.*, to record everything it heard during the S symbols and replay it over the next S symbols. The average energy received by node 1 is

$$(P^{\alpha_2}T-\Upsilon)+S\mathbb{E}[N_1^2] = (P^{\alpha_2}T-\Upsilon)+S$$
$$= \left(P^{\alpha_2}+\frac{P^{\beta}}{\Upsilon}\right)T-\Upsilon$$
$$= P^{\beta}(1/\Upsilon+P^{\alpha_2-\beta})T-\Upsilon.$$

Note that this is approximately $P^{\beta}T/\Upsilon$ when T is suitably large as in the sequel. Node 1 will spend a total energy of $P^{\gamma}T$ to replay what it heard, where

$$\alpha_1 < \gamma < \beta < \alpha_2.$$

i.e., during symbols $S+1, S+2, \ldots, 2S$, node 1 sends

$$\sqrt{\frac{P^{\gamma}}{P^{\beta}/\Upsilon}}Y_1^S,$$

where Y_1^S is the vector of S observations made by node 1. The receiver declares the message to be $k \in \{1, 2, \ldots, S\}$ if it finds a unique k satisfying

$$\left(y(k)+\frac{y(S+k)}{\sqrt{\frac{P^{\gamma}}{P^{\beta}/\Upsilon}}}\right)^2 \geq 4\lambda(P^{\alpha_2}T-\Upsilon).$$

Above, $0 < \lambda < 1$ is an arbitrary constant. Note that the signal-to-noise ratio of this detector is $2(P^{\alpha_2}T-\Upsilon)/(1+\frac{1}{2P^{\gamma-\beta}\Upsilon})$. Using standard techniques (the union bound and Chernoff bound) it is easy to show that the probability of error in decoding the message from node 2 is upper bounded by

$$\Pr(\text{error}) \leq \exp\left(\log S - \frac{\Psi(P,T)-1-\log\Psi(P,T)}{2}\right),$$

where

$$\Psi(P,T) = 2\lambda\frac{P^{\alpha_2}T-\Upsilon}{1+\frac{1}{2P^{\gamma-\beta}\Upsilon}}.$$

We now choose the block length T such that $\log S = \log(P^{\beta}T/\Upsilon)$ satisfies

$$\log S = \log\frac{P^{\beta}T}{\Upsilon} = \left(\frac{\Psi(P,T)-1-\log\Psi(P,T)}{2}\right)(1-\delta),$$

where $\delta > 0$ is an arbitrary constant. As P gets closer to 0, the T which satisfies the above condition grows such that $P^{\alpha_2}T \to \infty$. Note that with the above choice, the probability of error goes to 0 as $P \to 0$. We have

$$\frac{R_2(P)}{P^{\alpha_2}} = \frac{(\log S)/T}{P^{\alpha_2}}$$
$$= \frac{\Psi(P,T)}{2P^{\alpha_2}T}(1-\delta)-\left(\frac{1}{2P^{\alpha_2}T}+\frac{\log\Psi(P,T)}{2P^{\alpha_2}T}\right)(1-\delta)$$
$$\to \lambda(1-\delta),$$

in the limit as $P \downarrow 0$. Since $0 < \lambda < 1$ and $\delta > 0$ are arbitrary constants, $\lim_{P\downarrow 0} R_2(P)/P^{\alpha_2}$ can be made arbitrarily close to its upper bound of 1. Moreover, the fraction of time used in sending the message from node 2 is

$$2S/T = 2P^{\beta}/\Upsilon \to 0,$$

and the fraction of the energy used by node 1 is

$$\frac{P^{\beta}+P^{\gamma}}{P^{\alpha_1}} \to 0.$$

IV. Discussion

A few points are worth noting about the scheme.

- Each node sends its own message though its ON-OFF activity. This is consistent with the strategy which achieves the slope of capacity at $P = 0$ for the point-to-point problem.
- The above fact means that in order to be useful, a relay node (which obviously cannot anticipate exactly when the node it is helping will be ON) needs to be ON

and listening during a significant portion of the time, essentially the entire duration, during which the node it is helping may transmit. This limits the utility of the relay to only those cases where the relay node has significantly more power than the node it is trying to help.

- It has already been noted by El Gamal, Mohseni, and Zahedi [6] that for AWGN relay channels (without any power cost on the nodes for being ON), a strategy which time-shares between direct transmission and amplify-and-forward outperforms one which only performs amplify-and-forward, especially at low powers. Moreover, the power allocation is such that the amplify-and-forward phase lasts for a short proportion of the time, and operates at an effective SNR which does not go to zero. The intuition is that, since the relay forwards its observation to the destination, if what the relay receives has a low SNR, it will be forwarding mostly the noise it has observed. Hence, the portion of the transmission where the relay is active should not be at low SNR.

  In our case, the above intuition is doubly true. Besides the reason pointed out above, the transmitter will itself want to transmit only occasionally in order to reduce the amount of time it needs to stay ON, and thus incur the cost of staying ON.
- While the above discussion assumes a non-fading scenario, the pulse-position transmission scheme and energy detection receiver lend themselves to a direct application in a fading scenario with no channel knowledge at the nodes. For the application of PPM to a non-coherent scenario, but without the cost for staying ON, see [7].

REFERENCES

[1] J. Polastre, R. Szewczyk, C. Sharp, and D. Culler. “The Mote Revolution: Low Power Wireless Sensor Network Devices,” Hot Chips, August, 2004.

[2] P. Grover and A. Sahai. “Green codes: Energy-efficient short-range communication,” IEEE Symp. on Inform. Theory, 2008.

[3] P. Youssef-Massaad, L. Zheng, and M. Médard, “Bursty transmission and glue pouring: on wireless channels with overhead costs,” *IEEE Trans. Wireless Commun.* vol. 7, no. 12, pt. 2, pp. 5188–5194, 2008.

[4] R. S. Kennedy. *Fading Dispersive Communication Channels*, Wiley-Interscience, New York, 1969

[5] G. Kramer. “Models and theory for relay channels with receive constraints,” 42nd Annual Allerton Conf. on Commun., Control, and Computing, 2004.

[6] A. El Gamal, M. Mohseni, and S. Zahedi. “Bounds on Capacity and Minimum Energy-Per-Bit for AWGN Relay Channels,” *IEEE Trans. Inform. Theory.* IT-52(4), pp. 1545–1561, 2006.

[7] A. S. Avestimehr and D. N. C. Tse. “Outage Capacity of the Fading Relay Channel in the Low-SNR Regime,” *IEEE Trans. Inform. Theory.* IT-53(4), pp. 1401–1415, 2007.