

Bits through ARQs: Spectrum Sharing with a Primary Packet System

Krishnan Eswaran, Michael Gastpar, Kannan Ramchandran
 Dept. of Electrical Engineering and Computer Sciences
 University of California, Berkeley
 Berkeley, CA 94720

E-mail: {keswaran, gastpar, kannanr}@eecs.berkeley.edu

Abstract—We study a problem motivated by cognitive radio in which the primary is a packet system that employs ARQ feedback. A secondary system is allowed to transmit in the same frequency band provided it ensures that the primary attains a specified target rate. That is, the secondary has a certain “interference budget.” The crux of the problem is that the secondary does not know how much interference it creates on the primary and therefore is ignorant of its interference budget. Absent this knowledge, we propose a scheme in which the secondary eavesdrops on the primary’s ARQ and uses this knowledge to stay within its interference budget. Under certain assumptions, we show there exists an optimal rate-interference budget (RIB) tradeoff. We compare how far fixed strategies are from this RIB function as we vary the interference budget. Further, we exhibit a strategy that is optimal beyond a threshold interference budget and within 1 bit per primary packet elsewhere.

I. INTRODUCTION

A cognitive radio is a device that can sense and adjust its power, frequency band, etc. such that the primary systems for which these bands were originally allocated continue to function [1]. The FCC and international regulatory bodies are considering modifying their rules to allow for such systems to share bands with preexisting, predesigned primary communication systems that must continue to function.

There has already been work to study how easy it is to detect the presence of a primary system [2], [3]. An information-theoretic model has also been studied and is sometimes referred to as the cognitive radio channel [4], [5]. This channel is a variation on the Gaussian interference channel [6] with the modification that the cognitive radio, one of the transmitters, knows which message the primary, the other transmitter, will send. The authors of these papers also assume that the cognitive radio knows the statistics of the interference the cognitive radio generates on the primary. Under these assumptions, the authors demonstrate that it is optimal for the primary to use the same Gaussian codebook it would in the point-to-point setting. The result highlights the fact that in cognitive radio problems, one does not have the flexibility to modify the design of the primary system.

While we preserve the spirit that the primary system does not have to modify its design, we do away with modelling assumptions that require the cognitive radio to know the statistics of the interference it generates on the primary or the

message the primary intends to send. Absent such knowledge, the cognitive radio needs some way to infer this knowledge and adapt its rate so that the primary can meet its rate target. To do so, we restrict our attention to the case in which the primary system sends packets and employs ARQ feedback to determine whether or not to retransmit its packet. The cognitive radio transmitter and receiver can also listen to this ARQ¹, which allows the cognitive radio system to adjust its rate appropriately. This is summarized in Figure 1.

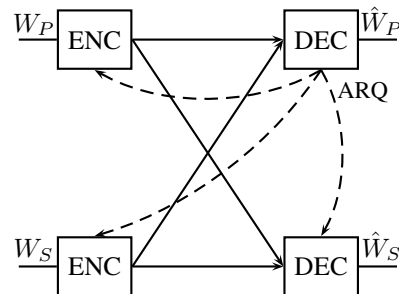


Fig. 1. An example of the type of interference channel for our cognitive radio system in which the primary has message W_P and the cognitive radio has message W_S . The cognitive radio encoder and decoder can listen to the ARQ feedback the primary receiver sends to its transmitter to adapt its transmission rate and reduce interference on the primary system.

If the interference statistics the cognitive radio generates on the primary system are stationary ergodic, then the cognitive radio can use an almost negligible fraction of packet transmissions to estimate this and select its codebook appropriately. However, the time scale of a primary packet transmission may be too large to guarantee coherence of the interference statistics. We assume that the primary’s packet erasure probabilities are larger when the cognitive radio transmits but can otherwise vary arbitrarily. Under such a setting, past performance is no guarantee of future behavior, so the cognitive radio must adopt a “do no harm” policy to ensure the primary meets its target rate. The performance of such a strategy under different conditions are summarized in plots we refer to as rate-interference budget (RIB) plots, examples of which are shown in Figures 4 and 5.

¹Future work will address limitations on who can hear the ARQ.

Before proceeding, we state assumptions necessary for our analysis. Specifically, we assume that the primary system has an inexhaustible queue from which to send packets but only wants a guaranteed packet rate R_P . For preciseness, we fix its transmission strategy at time k :

- 1) Transmit the smallest numbered packet not yet received.
- 2) ARQ feedback $A_k \in \{0, 1\}$ signals if this packet sent at time k is received.

Both the cognitive radio encoder and decoder also have access to ARQ feedback from the primary system, which can be thought of as the output to a time-varying DMC.

$$P(A_k = 1|T_k) = \begin{cases} 1 - \epsilon_1(k), & T_k = 1 \\ 1 - \epsilon_0(k), & T_k = 0 \end{cases}, \quad (1)$$

where the $\{\epsilon_i(k)\}$ are unknown sequences. We assume the channel factorizes as

$$P(\mathbf{Y}_k, A_k|\mathbf{X}_k, T_k) = P(\mathbf{Y}_k|\mathbf{X}_k, T_k)P(A_k|T_k). \quad (2)$$

The cognitive radio transmits across a discrete memoryless channel with the following two properties.

- 1) The cognitive radio has L channel uses over a primary packet, and the channel factorizes as follows.

$$P(\mathbf{Y}^n|\mathbf{X}^n, T^n, A^{n-1}) = \prod_{k=1}^n P(\mathbf{Y}_k|\mathbf{X}_k, T_k) \quad (3)$$

where $P(\mathbf{Y}_k|\mathbf{X}_k, T_k) = \prod_{j=1}^L P(Y_{k,j}|X_{k,j}, T_k)$.

- 2) Given that no transmission occurs during time k ($T_k = 0$), \mathbf{X}_k is independent of \mathbf{Y}_k .

$$I(\mathbf{X}_k; \mathbf{Y}_k|T_k = 0) = 0. \quad (4)$$

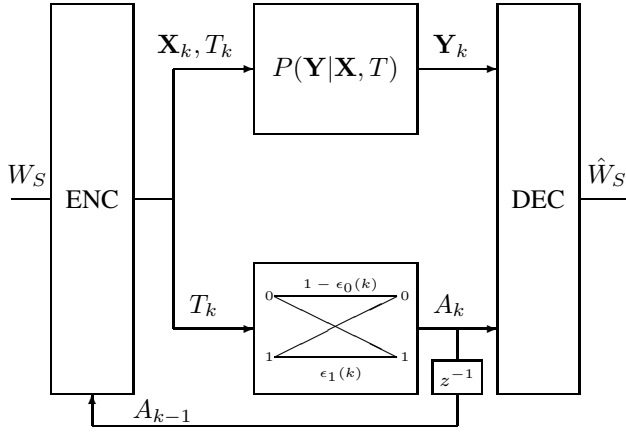


Fig. 2. Our assumptions lead to the following model for the cognitive radio's channel. The cognitive radio must adapt its rate in such a way to ensure that the primary meets its target rate R_P , so $n^{-1} \sum_{k=1}^n A_k \geq R_P$ must hold with high probability as n gets large.

Given that the strategy of the primary system is fixed a priori, the main constraint placed on the cognitive radio is that the primary must meet its rate target R_P . We ignore delay and simply require that for all $\delta > 0$ and large enough n ,

$$\frac{1}{n} \sum_{k=1}^n A_k > R_P - \delta \quad (5)$$

with high probability.

When the $\epsilon_i(k)$ are not fixed for all k , we must handle two problems. The first is to make sure the message is decoded correctly. The second is to make sure that the primary system receives enough packets to meet its target rate R_P . At the cost of 1 bit per primary packet, we consider encoders that separate these two functions in a way that sacrifices the potential to communicate over the ARQ channel. This is made precise in the next section.

II. PROBLEM SETUP

Definition 1: A schedule $\Xi(h)$ is a sequence of functions

$$h_i : \{0, 1\}^{i-1} \rightarrow \{0, 1\}$$

such that

$$\tau_k = h_k(A^{k-1}). \quad (6)$$

Note that given a schedule, the decoder can determine τ_k . Rather than refer to the schedule directly, we will refer to it obliquely in the definition of τ_k .

Definition 2: A block coding strategy $\Delta(n_0, M, f, g_i)$ with M codewords is composed of an encoding function

$$f : \{1, \dots, M\} \rightarrow \mathcal{X}^{m \times n_0} \times \{0, 1\}^{n_0}$$

and decoding functions

$$g_i : \mathcal{Y}^{m \times i} \times \{0, 1\}^i \times \{0, 1\}^i \rightarrow \{1, \dots, M\}$$

over the positive integers such that, for message $W \in \{1, \dots, M\}$, $\tilde{\mathbf{X}}^{n_0}, \tilde{T}^{n_0} = f(W)$ and $\hat{W}_i = g_i(\mathbf{Y}^i, A^i, \tau^i)$.

Given a schedule and corresponding block coding strategy, channel inputs are determined as follows.

$$\mathbf{X}_k, T_k = \begin{cases} x_0^m, 0 & \tau_k = 0 \\ \tilde{\mathbf{X}}_{\bar{\tau}_k}, \tilde{T}_{\bar{\tau}_k} & \tau_k = 1 \end{cases}, \quad (7)$$

where $\bar{\tau}_k = \sum_{i=1}^k \tau_i$ and $x_0 \in \mathcal{X}$. Note that the choice of x_0 can be arbitrary given (4).

We take a compound channel view of successful decoding of the message.

Definition 3: A rate R is achievable if for every $\delta > 0$, there exists a schedule $\Xi(h)$, block coding strategy $\Delta(n_0, M, f, g_i)$, and $n, n_0 \leq n < \frac{\log M}{R}$ such that for all $\{\epsilon_i(k)\}$,

$$P(W \neq \hat{W}_n) \leq \delta. \quad (8)$$

Definition 4: A rate R is R_P -achievable for the sequence $\{\epsilon_i(k)\}$ if it is achievable with $\Delta(n_0, M, f, g_i)$, and for every $\delta > 0$ and for large enough M , there exists some $n, n_0 \leq n < \frac{\log M}{R}$, schedule $\Xi(h)$, such that for $\Delta(n_0, M, f, g_i)$, $P(W \neq \hat{W}_n) < \delta$ and for all $n_* \geq n_0$

$$P\left(n_*^{-1} \sum_{i=1}^{n_*} A_i < R_P - \delta\right) < \delta. \quad (9)$$

Definition 5: Let $\mathcal{R}(R_P, \{\epsilon_i(k)\})$ be the set of all R_P -achievable rates. Then

$$R_{IB}(R_P, \{\epsilon_i(k)\}) := \sup \mathcal{R}(R_P, \{\epsilon_i(k)\}) \quad (10)$$

is the rate-interference budget function. For sequences with limits $\lim_{k \rightarrow \infty} \epsilon_i(k) = \epsilon_i$, we use the shorthand $R_{IB}(R_P, \epsilon_0, \epsilon_1)$.

Definition 6:

$$R_{IB}(\gamma) = \max_{p(\mathbf{x}, t); p(t=1) \leq \gamma} I(\mathbf{X}, T; \mathbf{Y}). \quad (11)$$

Our main result is the following.

Theorem 1: If $1 - \epsilon_1 < R_P < 1 - \epsilon_0$, then for interference budget $\gamma = \frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0}$,

$$R_{IB}(R_P, \epsilon_0, \epsilon_1) = R_{IB}(\gamma). \quad (12)$$

III. ACHIEVABLE STRATEGY

We now consider our achievable strategy. We start by specifying the scheduling algorithm. Let

$$S_k = \sum_{i=1}^k (A_i - R_P), \quad (13)$$

which denotes the difference between the number of packets received by the primary and the target number of packets. We use this to construct the following schedule.

$$\tau_k = \begin{cases} 0 & S_{k-1} \leq 0 \\ 1 & S_{k-1} > 0 \end{cases}. \quad (14)$$

Given this schedule, we state results about achievable strategies. To conserve space, we argue the case where $\epsilon_i(k) = \epsilon_i$, which is given in the Appendix. It is straightforward to extend the arguments for sequences $\limsup_k \epsilon_i(k) = \epsilon_i$.

Theorem 2: The scheduling algorithm applied above leads to the following R_P achievable rates. For $1 - \epsilon_1 < R_P < 1 - \epsilon_0$,

$$\begin{aligned} R_{IB}(R_P, \epsilon_0, \epsilon_1) \\ \geq \frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} \max_{p(\mathbf{x}|T=1)} I(\mathbf{X}; \mathbf{Y}|T=1) \end{aligned} \quad (15)$$

Note that we are not using T to convey information in Theorem 2. Indeed, we can do at least as well as the above strategy if we do.

Theorem 3: When

$$\gamma := p(t=1) < \frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} \quad (16)$$

is satisfied, then rates

$$R_{IB}(R_P, \epsilon_0, \epsilon_1) \geq \max_{\substack{p(\mathbf{x}, t): \\ p(t=1) = \gamma}} I(\mathbf{X}, T; \mathbf{Y}) \quad (17)$$

are R_P -achievable. Otherwise, under the distribution $p(\mathbf{x}, t)$,

$$\begin{aligned} R_{IB}(R_P, \epsilon_0, \epsilon_1) \geq & I(\mathbf{X}; Y|\mathbf{T}=1) \frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} \\ & + \frac{1 - \epsilon_0 - R_P}{\gamma(\epsilon_1 - \epsilon_0)} I(T; \mathbf{Y}). \end{aligned} \quad (18)$$

Proof: We construct our random codebook by drawing our codewords independently from the distribution

$$p(\mathbf{x}^{n_0}, t^{n_0}) = \prod_{i=1}^{n_0} p(\mathbf{x}_i, t_i). \quad (19)$$

The number of codewords M will be

$$M = 2^{n_0(\max_{p(\mathbf{x}, t), p(t=1)=\gamma} I(\mathbf{X}, T; \mathbf{Y}) - \delta)}. \quad (20)$$

Note that the error probability bound for this code given in (34) continues to hold.

When (16) is not satisfied, the argument follows almost identically to the proof of Theorem 2, except that when applying Theorem 6, we replace ϵ_1 with $\gamma\epsilon_1 + (1 - \gamma)\epsilon_0$ to give an effective transmit fraction of

$$\frac{1 - \epsilon_0 - R_P}{\gamma(\epsilon_1 - \epsilon_0)}, \quad (21)$$

showing that (18) is R_P -achievable.

Under (16), one can lower bound A_k by i.i.d. Bernoulli($1 - \bar{\epsilon}$) random variables B_k , where $\bar{\epsilon} = \gamma\epsilon_1(1 - \gamma)\epsilon_0$. Then the probability that the decoder does not always have the option to transmit after time ℓ is given by

$$\sum_{k=\ell}^{\infty} \mathbb{P}(S_k < 0) \leq \sum_{k=\ell}^{\infty} e^{-k(R_P - 1 + \gamma\epsilon_1 + (1 - \gamma)\epsilon_0)^2 / 2} \quad (22)$$

$$= \frac{e^{-\ell(R_P - 1 + \bar{\epsilon})^2 / 2}}{1 - e^{-(R_P - 1 + \bar{\epsilon})^2 / 2}} \quad (23)$$

For $\ell = \frac{\sqrt{n_0}}{(R_P - 1 + \bar{\epsilon})^2}$, one can achieve the rate

$$\max_{p(\mathbf{x}, t), p(t=1)=\gamma} I(\mathbf{X}, T; \mathbf{Y}) - \eta(n_0), \quad (24)$$

with error probability no more the sum of (34) and

$$\sum_{k=\ell}^{\infty} \mathbb{P}(S_k < 0) \leq \frac{e^{-\sqrt{n_0}/2}}{1 - e^{-n_0^{-1/4}/2}}, \quad (25)$$

where $\eta(n_0) \rightarrow 0$ as $n_0 \rightarrow \infty$. ■

IV. CONVERSE

We now turn to a proof of the converse. The following result will be useful in the sequel.

Lemma 1: $R_{IB}(\gamma)$ is a nondecreasing, concave function.

Proof: Showing it is increasing is trivial. Let $\lambda\gamma_1 + (1 - \lambda)\gamma_2 = \gamma$. Then

$$\lambda R_{IB}(\gamma_1) + (1 - \lambda)R_{IB}(\gamma_2) \leq I_{p_{\lambda}(\mathbf{x}, t)}(\mathbf{X}, T; \mathbf{Y}) \quad (26)$$

$$\leq R_{IB}(\gamma). \quad (27)$$

where (26) follows since mutual information is concave in p [8, Theorem 2.7.4] and (27) by definition. ■

Theorem 4: For $\epsilon_i(k) = \epsilon_i$ such that $1 - \epsilon_1 < R_P < 1 - \epsilon_0$, then

$$R_{IB}(R_P, \epsilon_0, \epsilon_1) \leq R_{IB}\left(\frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0}\right). \quad (28)$$

Proof: The argument is divided into two parts. First, we bound on the rate so that decoding is successful for all $\{\epsilon_i(k)\}$. Second, we constrain the channel input distribution for specific ϵ_i and apply this to our bound. Let us consider the rate of the cognitive radio user. Fano's inequality and standard arguments show that

$$nR \leq \sum_{i=1}^n (I(\mathbf{X}_i, T_i; \mathbf{Y}_i) + I(A_i; T_i)) + n\delta_n. \quad (29)$$

By Definition 3, the message has to decode with arbitrarily small error probability for all $\{\epsilon_i(k)\}$. We take the infimum over all possible sequences $\{\epsilon_i(k)\}$ to give the tighter bound

$$nR \leq \sum_{i=1}^n I(\mathbf{X}_i, T_i; \mathbf{Y}_i) + n\delta_n, \quad (30)$$

where $\delta_n \rightarrow 0$. Now

$$R \leq \sum_{i=1}^n \frac{1}{n} R_{IB}(\gamma_i) + \delta_n \quad (31)$$

$$\leq R_{IB} \left(\frac{1 - \epsilon_0 - R_P + \delta}{\epsilon_1 - \epsilon_0} \right) + \delta_n \quad (32)$$

where (31) follows from Lemma 2 for some γ_i such that $n^{-1} \sum \gamma_i \leq \frac{1 - \epsilon_0 - R_P + \delta}{\epsilon_1 - \epsilon_0}$, and (32) follows from Lemma 1. The result follows by taking the limit as $n \rightarrow \infty$ and then as $\delta \rightarrow 0$. ■

While we do not claim there exists one strategy that universally attains this upper bound, Theorem 3 implies that for every point in the upper bound, there exists a strategy that achieves it. This is sufficient to prove Theorem 1.

V. EXAMPLES

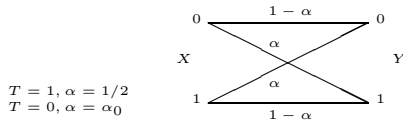


Fig. 3. In the above binary symmetric channel, the position of transmissions over primary packets conveys no information to the cognitive radio receiver. For this setting, the achievable strategy universally attains all points on the rate-interference budget (RIB) function.

Consider the binary symmetric channel given in Figure 3. Here, $L = 1$, and given a realization of T , we have a binary symmetric channel. For this channel, the strategy in Theorem 2 attains all points on the RIB function $R_{IB}(R_P, \epsilon_0, \epsilon_1) = \frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} (1 - h_b(\alpha_0))$.

As a second example, consider a channel in which T is perfectly received at Y , so $P(Y = t|X, T = t) = 1$. Here, $m = 1$, $|\mathcal{X}| = 1$. For this case, we plot different block coding strategies in Figure 4 based on Theorem 3, which hit different points on the RIB function. The gap reflects the regret of particular strategy.

One might expect that the number of channel uses available to the cognitive radio over a primary packet would be large. In our third example, we consider a cognitive radio with L uses of a binary symmetric over the run of a single packet for the primary that we describe with the channel

$$P(Y|X, T) = \begin{cases} 1 & Y = \text{OFF}, T = 0 \\ \gamma^k (1 - \gamma)^{L-k} & T = 1, d_H(Y, X) = k \end{cases} \quad (33)$$

Here, $d_H(\cdot, \cdot)$ denotes Hamming distance. For the values of L plotted in Figure 5, the gap between the achievable strategies and converses are negligible for transmissions on the order of a packet for the cognitive radio.

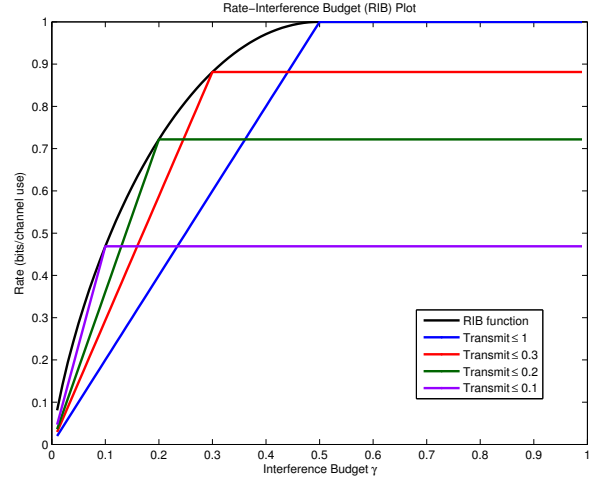


Fig. 4. This RIB plot for a channel in which it is only possible to communicate via which primary packet positions the cognitive radio chooses to transmit. While no strategy shown can universally attain the RIB function (black), each can meet at least one point on it.

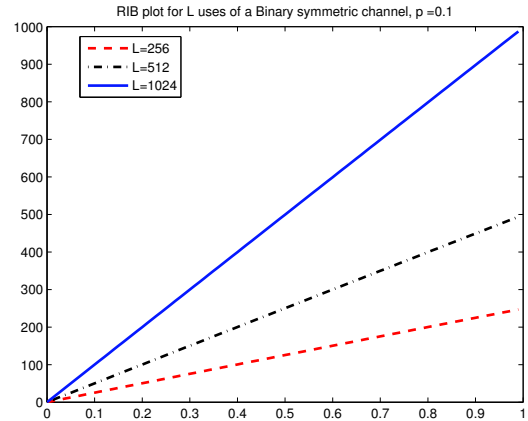


Fig. 5. In this example, we plot the rate-interference budget (RIB) function for cases in which the cognitive radio has L uses of a BSC over a primary packet transmission. When L is large enough, the gap from an achievable strategy to the RIB function is negligible.

VI. DISCUSSION

We introduced a problem motivated by cognitive radio in which the primary is not a second user in the communication system but rather a constraint on the first user. Based on our encoder restrictions, we were able to show a class of achievable strategies and a converse. One may also have introduced common randomness between the encoder and decoder and allowed for general encoder structures. Such topics will be areas of future exploration.

ACKNOWLEDGMENTS

We would like to acknowledge comments and suggestions from qualifying exam committee members Anant Sahai and Sourav Chatterjee (Stat, UC Berkeley). We are also thankful for stimulating discussions with Mubaraq Mishra. This work was supported by the NSF under award CNS-0326503.

APPENDIX

We begin with a proof of Theorem 2.

Proof: We construct a random codebook according to the distribution $p(\mathbf{x}^{n_0} | t^{n_0} = 1) = \prod_{i=1}^{n_0} p(\mathbf{x}_i | t_i = 1)$ and $M = 2^{n_0(I(\mathbf{X}; \mathbf{Y} | T=1) - \delta)}$ codewords.

Let E_n denote the event of a decoder failure at time n . Given the event that $\sum_{i=1}^n \tau_i \geq n_0$, for any $\delta > 0$, a random coding argument can be used to show [9, p. 539, Problem 5.23]

$$\mathbb{P}(E_n | \sum_{i=1}^n \tau_i \geq n_0) \leq e^{-n_0 \frac{\delta^2}{8/\epsilon^2 + 4(\log |\mathcal{Y}|)^2}}. \quad (34)$$

When the error does not occur, this results in a rate of

$$\frac{n_0}{n} (I(\mathbf{X}; \mathbf{Y} | T = 1) - \delta). \quad (35)$$

It remains to understand for which ratios $\frac{n_0}{n}$ we can guarantee with high probability that $\sum_{i=1}^n \tau_i \geq n_0$ and the primary is close to its rate target R_P . We now consider R_P -achievability for the schedule above. We set $n \geq n_0$ to satisfy

$$\frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} - \delta - n_0^{-1} \leq \frac{n_0}{n} \leq \frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} - \delta. \quad (36)$$

For $\frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} - \delta > 0$, the existence of such an n is guaranteed, and under the convention that decoding need not occur to meet rate 0, a series of manipulations (omitted for space) can be used to show that the rate

$$\frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0} I(\mathbf{X}; \mathbf{Y} | T = 1) - \eta(n_0, \delta) \quad (37)$$

is achievable with rate loss $\eta(n_0, \delta)$ with error probability no more than the sum of (34),

$$\mathbb{P}\left(\sum_{i=1}^n \tau_i < n_0\right) \leq 4e^{-\delta^2 n_0^{1/3}/2} + 2e^\lambda e^{-\lambda(\delta n_0^{2/3} - n_0^{1/3})} \quad (38)$$

by Theorem 6, and

$$\mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n A_k < R_P - \delta\right) \leq e^\lambda e^{-\lambda(\delta n_0^{2/3} - n_0^{1/3})} \quad (39)$$

by Theorem 5. For $\delta = n_0^{-1/12}$ and as $n_0 \rightarrow \infty$, all of these errors vanish and $\eta(n_0, n_0^{-1/12}) \rightarrow 0$. ■

Lemma 2: If the constraint (9) is satisfied, then

$$\frac{1}{n} \sum_{i=1}^n P(T_i = 1) \leq \frac{1 - \epsilon_0 - R_P + \delta}{\epsilon_1 - \epsilon_0}. \quad (40)$$

Proof: If (9) holds, then $\frac{1}{n} \sum_{i=1}^n P(A_i = 1) \geq (1 - \delta)(R_P - \delta) \geq R_P - \delta$ must hold since $0 \leq R_P \leq 1$. Recognizing that $P(A_i = 1) = \epsilon_1 P(T_i = 1) + \epsilon_0 P(T_i = 0)$ and rearranging terms completes the proof. ■

Theorem 5: Let $\sup_k 1 - \epsilon_1(k) < R_P < \inf_k 1 - \epsilon_0(k)$. If $T_k = \tau_k$ for all k , then for all $\delta > 0$, $\lambda > 0$,

$$\mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n A_k < R_P - \delta\right) \leq e^{\lambda v(R_P, \epsilon) - n\lambda\delta}, \quad (41)$$

where $v(R_P, \epsilon) = 1 + \max\left\{\frac{1}{\inf_k 1 - \epsilon_0(k) - R_P}, \frac{1}{\sup_k R_P - (1 - \epsilon_1(k))}\right\}$.

Proof: Details are omitted due to lack of space, but the result follows from an application of Markov's inequality on the stopping times

$$N_{2k-1} = \inf\{n > N_{2k-2} : S_n > 0\}, \quad (42)$$

$$N_{2k} = \inf\{n > N_{2k-1} : S_n \leq 0\}, \quad (43)$$

which involves bounding the expectation of the submartingale $e^{\alpha \tilde{S}_n}$, where \tilde{S}_n is the martingale

$$\tilde{S}_n = S_n - n(1 - R_P) + \sum_{k=1}^n \epsilon_0(k) + \sum_{k=1}^n \tau_k(\epsilon_1(k) - \epsilon_0(k)). \quad (44)$$

Theorem 6: Let $\epsilon_i(k) = \epsilon_i$ for all k and $1 - \epsilon_1 < R_P < 1 - \epsilon_0$. If $T_k = \tau_k$ for all k , then for all $\delta > 0$, $\lambda > 0$,

$$\mathbb{P}\left(\left|\sum_{k=1}^n \tau_k - \frac{1 - \epsilon_0 - R_P}{\epsilon_1 - \epsilon_0}\right| > \frac{\delta}{\epsilon_1 - \epsilon_0}\right) \leq 4e^{-nD\left(\frac{1+\delta/2}{2} \parallel \frac{1}{2}\right)} + 2e^{\lambda v(R_P, \epsilon) - \lambda n\delta/2}, \quad (45)$$

where

$$v(R_P, \epsilon) = 1 + \max\left\{\frac{1}{1 - \epsilon_0 - R_P}, \frac{1}{R_P - (1 - \epsilon_1)}\right\}.$$

Proof: By the triangle inequality, $|\sum_{k=1}^n \tau_k - n(1 - R_P) + n\epsilon_0| \leq |\tilde{S}_n| + |S_n|$, where \tilde{S}_n is defined in (44). Theorem 5 gives a bound on S_n violated with a small probability. Furthermore, results from Dembo and Zetouni [10, Cor. 2.4.7, p. 57] on bounded martingales and a union bound complete the proof. ■

REFERENCES

- [1] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology, 2000.
- [2] R. Tandra and A. Sahai, "Fundamental limits on detection in low snr under noise uncertainty," in *WirelessCom 05 Symposium on Signal Processing*, June 2005.
- [3] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Allerton Conference on Communication, Control, and Computing*, Oct. 2004.
- [4] N. Dervoye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.
- [5] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," 2006. [Online]. Available: <http://www.citebase.org/abstract?id=oai:arXiv.org:cs/0604107>
- [6] M. H. M. Costa, "On the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 31, pp. 607–615, 1985.
- [7] I. Csiszár and J. G. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, FL, USA: Academic Press, Inc., 1982.
- [8] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley and Sons, 1991.
- [9] R. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons, 1968.
- [10] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications, 2nd ed*, ser. Stochastic Modelling and Applied Probability. Springer, 1998.