Is interference like noise when you know its codebook?

Rahul Tandra
Wireless Foundations
University of California, Berkeley
Email: tandra@eecs.berkeley.edu

Anant Sahai
Wireless Foundations
University of California, Berkeley
Email: sahai@eecs.berkeley.edu

***Abstract*— We consider a point to point communication system facing interference from other systems, with a particular focus on the case when this interference is undecodable. It is well known that when the interference is non-interactive, we can certainly treat it as additional noise at the receiver and thereby achieve certain rates. This paper asks whether any higher rates could be achieved by exploiting knowledge of the interferer's codebook.**

The main contribution of this paper is to study the converse: if the interference is undecodable, then we cannot do better than treating it as additional noise. This is proved for almost all interference codebooks when viewed under the random Gaussian codebook measure. When the interference signal is strong enough to be decodable, then codebook knowledge can be exploited at our receiver to allow higher rates to be achieved by appropriately structuring our own codebooks.

Finally, we give an example of an interference codebook that cannot be completely decoded, but whose knowledge is still useful. However, this interference codebook is bad from the perspective of the interferer's own communication system. This leads us to conjecture that when the interference signal is undecodable, the only interference codebooks worth knowing are those that are not worth using from the interfering system's point of view.

I. Introduction

We are motivated by the engineering problem (illustrated in Figure 1) of designing a point to point communication system (denoted secondary) that has to operate in the presence of interference from other (primary[1]) systems. We consider the case when the secondary system is selfish and only wants to maximize its communication rate subject to a pre-specified transmit power constraint.[2] It does not care if by doing so it causes harmful interference to the other systems.[3]

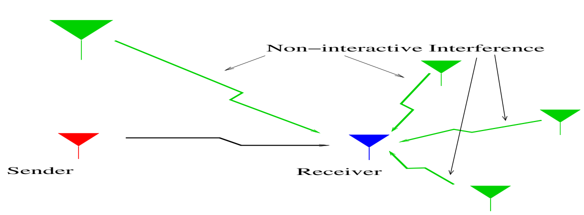

Fig. 1. Wireless interference network

[1]Our terminology reveals our motivation's origin in the study of cognitive radio systems [1], [2] and opportunistic spectrum sharing [3].

[2]This corresponds to a spectrum sharing regime in which the relevant regulation is expressed as an individual power constraint.

[3]Presumably, that concern is dealt with at the regulatory level by limiting the transmit power of the secondary system.

It is known from [4] that the secondary user can treat the interference[4] from the primary user as additional noise and achieve the corresponding AWGN capacity: $\frac{1}{2}\log\left(1+\frac{P}{\beta P_p + N}\right)$, where βP_p is the received interference power, N is the receiver noise, and P is the received power from our own transmitter. In particular, *random Gaussian coding* combined with nearest neighbor decoding can be used to achieve this, irrespective of the actual interference codebook.

We take particular interest in the cases when the interference is undecodable: our receiver could not decode the interference codeword even if our transmitter were off. Such a situation can arise whenever we are outside the primary user's service radius and we will model it as the interference codebook's rate R_I being larger than the relevant capacity $\frac{1}{2}\log\left(1+\frac{\beta P_p}{N}\right)$.

Under this setup we ask: Is there any value for our system to know the codebook of the interferer? More comically, if an industrial spy were willing to sell the interferer's codebook, how much would we be willing to pay for it?

To answer this, we assume that both the sender and receiver are aware of the interference codebook, but not the exact interference codeword that was realized. We want to find the maximum rate at which the sender can convey information to its receiver.

This paper is structured as follows: After commenting on related work in Section II, a precise information theoretic formulation is given in Section III. The main converse result for the case of a random Gaussian interferer is proved in Section IV. We extend our results to the case of multiple interferers in Section V and the discussion of general interference codebooks is done in Section VI. We give some concluding remarks in Section VII.

II. Context and Related Work

The problem considered in this paper comes from the general area of spectrum sharing. For instance, two systems tying to operate in the same frequency band will face interference from each other, even if they have some geographical separation. A topic of significant current interest is the case of cognitive radios that can sense their surroundings and adapt

[4]Here we assume that the primary user is non-interactive: it does not try to jam the secondary's transmissions as it might in an AVC formulation.

their communication strategies accordingly. The question addressed here is in what cases does it pay for the cognitive radios to adapt to the interferer's codebook?

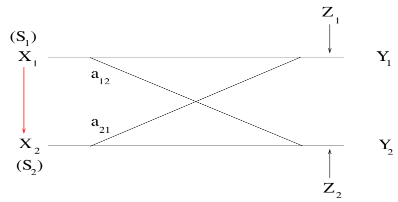

Fig. 2. The two user "cognitive radio channel" of [5].

Devroye, et. al. in [5] propose an interference channel model (illustrated in Figure 2) for cognitive radios, that they call the *cognitive radio channel*. It is defined as an interference channel in which S_2 has knowledge of the message to be transmitted by S_1. This models situations in which the cognitive radio faces very strong interference from the primary and hence can decode his message. Using ideas from dirty paper coding [6] they derive an achievable rate region in [5] for the two user cognitive radio channel. More recently, Jovicic and Viswanath in [7] have derived the maximum rate achievable for the cognitive radio channel under some special interference regimes.

In this line of prior work, the main assumption is that the primary's message can be decoded by the secondary user. Furthermore, the underlying values are 'polite' in that the secondary user tries to listen to the primary for a while, decode it and try to take advantage of the decoded message in a way that does not cause harmful interference to the primary. In contrast, we focus on the 'rude' approach in which the secondary user does not care about the success or failure of the primary's message, but is willing to exploit whatever it knows to its own advantage.

Another difference is that we implicitly consider the case of secondary users that are outside the service area of the primary and so may not be able to decode the primary at all. An analogous question with respect to detection rather than decoding has been treated in [8]–[10]. At low SNR, the sample complexity of detecting the presence or absence of a primary user's signal does not change significantly with additional knowledge of the signal constellation (assumed to be zero-mean). From a sample complexity point of view, treating the primary signal as noise and doing a simple energy detection is nearly optimal. The result here suggests that the same philosophy applies at the level of code-design — it is not worth adapting to the details of the primary codebook when the primary signal's SNR is too low to decode. In a sense, this validates the current assumption that given enough geographical separation, systems can be designed independently of each other and should only end up seeing each other as noise.

III. Problem formulation

The boxed area in Figure 3 shows the channel of interest. The secondary transmitter and receiver are referred to as sender and receiver respectively. For obvious practical reasons, we make the half-duplex assumption that our sender cannot receive and transmit at the same time in the same frequency band.

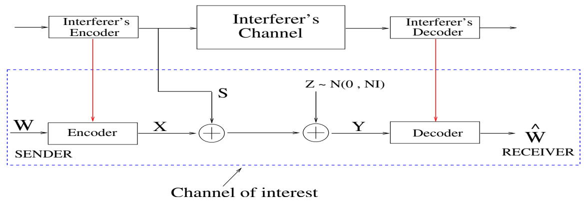

Fig. 3. Information theoretic problem formulation

The channel of interest is expressed as:

$$Y_i = X_i + S_i + Z_i$$

where X_i and Y_i denote the input and output of the channel, S_i is the interference and Z_i is the receiver noise at time i. We assume that the noise process, interference and the input are independent of each other.

- *Noise process*: White Gaussian $Z_i \sim \mathcal{N}(0, N)$.
- *Interference*: Rate R_I, with codewords uniformly chosen from a $(2^{nR_I}, n)$ code book, $\mathcal{S}_n$. Each codeword $\mathbf{s}_i = (s_{i1}, s_{i2}, \cdots, s_{in})$ must satisfy:

$$\frac{1}{n}\sum_{j=1}^{n} s_{ij}2 \leq I, \qquad i = 1, 2, \cdots, 2^{nR_I} \tag{1}$$

We assume that $R_I > \frac{1}{2}\log\left(1 + \frac{I}{N}\right)$, i.e., the interference is undecodable even when the sender is off.

For each block, we assume that the interference codebook $\mathcal{S}_n$ is revealed to both the sender and receiver and thus the channel input is allowed to depend on the interference codebook, but not on the interference message. Our goal is to send a message $W \in \{1, 2, \cdots, 2^{nR}\}$ to the receiver in n uses of the channel.[5] Thus we must specify a $(2^{nR}, n)$ code $\mathcal{C}_n$, with codewords $\mathbf{x}(W) = (x_1(W), x_2(W), \cdots, x_n(W)), W \in \{1, 2, \cdots 2^{nR}\}$ satisfying the power constraint $\frac{1}{n}\sum_{i=1}^{n} x2_i(W) \leq P$ together with a decoding function $g_n : \mathbf{R}^n \rightarrow \{1, 2, \cdots 2^{nR}\}$ such that the average error probability $P_e^{(n)}$ is low.

$$\begin{aligned} P_e^{(n)}(\mathcal{C}_n, \mathcal{S}_n) &= \frac{1}{2^{nR}}\sum_{i=1}^{2^{nR}} Pr\left\{g_n(\mathbf{Y}) \neq i | \mathbf{X} = \mathbf{x}(i)\right\} &(2)\\ &= Pr\{g_n(\mathbf{Y}) \neq W)\} &(3)\end{aligned}$$

where it is assumed that the message W is uniformly distributed over $\{1, 2, \cdots 2^{nR}\}$ and is independent of the interference codebook $\mathcal{S}_n$ and the noise sequence $\mathbf{Z}$.

The aim is to find the set of rates R, at which $P_e^{(n)}(\mathcal{C}_n, \mathcal{S}_n) \rightarrow 0$ as $n \rightarrow \infty$ for some sequence of

[5]Throughout this paper we assume that the interference blocklength is same as the blocklength used by our encoder. All our results should easily extend to the case when the blocklengths are different.

encoder/decoder designs. From results in [4], it follows that the sender can communicate at all rates $R < \frac{1}{2}\log\left(1 + \frac{P}{I+N}\right)$

IV. Main result

Trying to prove a converse for all "reasonable" interference codebooks is very tricky since there are silly interference codebooks for which we can in fact do better. Instead, we consider the product Gaussian measure on the set of all interference codebooks satisfying the power constraint in (1), i.e., we assume that each coordinate of an interference codeword is generated using a Gaussian distribution. Under this measure we are interested in the fraction of interference codebooks for which the converse is true. We show that the probability of the set of all interference codebooks that are not worth knowing can be made arbitrarily close to 1 for sufficiently large blocklength.

***Theorem** 1:* Consider the scalar additive noise plus interference channel discussed in Section III. Fix $R_I > \frac{1}{2}\log\left(1 + \frac{I}{N}\right)$ and $R > \frac{1}{2}\log\left(1 + \frac{P}{I+N}\right)$. For every $0 < \delta \leq 1$, there exists some $0 < \epsilon \leq 1$ and $n_0 > 0$ such that, $\forall\ n \geq n_0$ we have

$$Pr\left\{\mathcal{S}_n \ : \ P_e^{(n)}(\mathcal{C}_n(\mathcal{S}_n), \mathcal{S}_n) \geq \epsilon\right\} \geq 1 - \delta \tag{4}$$

for any input codebook, $\mathcal{C}_n(\mathcal{S}_n)$. Here $P_e^{(n)}(\mathcal{C}_n(\mathcal{S}_n), \mathcal{S}_n)$ is the error probability (as given in (3)) when the encoder uses $\mathcal{C}_n(\mathcal{S}_n)$ for transmitting the message index[6]. Here we are assuming a product Gaussian measure on the interference codebooks.

Proof: We prove this result by contradiction. Suppose the theorem were false. Then, there would exist some $0 < \delta \leq 1$ such that $\forall\ \epsilon \in (0, 1]$, $\exists$ a sequence of $n_k \uparrow \infty$ such that

$$Pr\left\{\mathcal{S}_{n_k} \ : \ P_e^{(n_k)}(\mathcal{C}_{n_k}(\mathcal{S}_{n_k}), \mathcal{S}_{n_k}) < \epsilon\right\} \geq \delta \tag{5}$$

for some input codebook $\mathcal{C}_{n_k}(\mathcal{S}_{n_k})$.

Now, partition the messages within interference codebook $\mathcal{S}_{n_k}$ into bins of size $2^{n_k \tilde{R}_I}$, where $\tilde{R}_I = \frac{1}{2}\log\left(1 + \frac{I}{N}\right) - \gamma$ for some small $\gamma > 0$. Since we use a product Gaussian measure on the set of all power constrained $(2^{n_k R_I}, n_k)$ interference codebooks, the distribution of the codewords in each of the bins is also a product Gaussian distribution.

We now allow a genie to reveal the bin of the interference codebook containing the actual interference sequence to both the encoder and decoder. Call this thinned codebook $\tilde{\mathcal{S}}_{n_k}$. Since the total number of possible interference sequences is reduced, the genie-aided decoder performs at least as well as the original decoder. Therefore, we must have

$$Pr\left\{\tilde{\mathcal{S}}_{n_k} \ : \ P_e^{(n_k)}(\mathcal{C}_{n_k}(\tilde{\mathcal{S}}_{n_k}), \tilde{\mathcal{S}}_{n_k}) < \epsilon\right\} \geq \delta \tag{6}$$

Given the received channel output $\mathbf{Y}$ the genie-aided receiver first decodes the input codeword $\mathbf{X}$. We can then subtract this decoded codeword from $\mathbf{Y}$ to get an estimate of interference plus noise $\mathbf{S} + \mathbf{Z}$. Now, we can decode the interference codeword $\mathbf{S}$ from the estimate of $\mathbf{S} + \mathbf{Z}$ using nearest neighbor decoding. Hence, the genie-aided decoder can get an estimate of the true codeword pair $(\mathbf{X}, \mathbf{S})$. The error probability can be computed by union bounding the probability of the following two error events:

- E_1: The genie-aided decoder makes an error in decoding $\mathbf{X}$ from $\mathbf{Y}$.
- E_2: The thinned interference codebook, $\tilde{\mathcal{S}}_{n_k}$, is bad and hence the nearest neighbor decoder makes an error in decoding $\mathbf{S}$ from $\mathbf{S} + \mathbf{Z}$.

Let A_{n_k} denote the event $\left\{\tilde{\mathcal{S}}_{n_k} \ : \ P_e^{(n_k)}(\mathcal{C}_{n_k}(\tilde{\mathcal{S}}_{n_k}), \tilde{\mathcal{S}}_{n_k}) < \epsilon\right\}$. From (6) we know that $Pr(A_{n_k}) \geq \delta$. Now, let us focus on $\tilde{\mathcal{S}}_{n_k} \in A_{n_k}$. We know that $Pr(E_1) \leq \epsilon$ for all $0 < \epsilon \leq 1$.

Also, recall that $\tilde{R}_I < \frac{1}{2}\log\left(1 + \frac{I}{N}\right)$, and the distribution of the codewords in $\tilde{\mathcal{S}}_{n_k}$ is I.I.D. Gaussian with average power P. Therefore, for large enough n_k and $\forall\ 0 < \epsilon \leq 1$, A_{n_k} must contain at least one codebook $\tilde{\mathcal{S}}_{n_k}$ which approaches capacity for a point to point AWGN channel with noise power N, in the sense that, for this specific $\tilde{\mathcal{S}}_{n_k}$, $Pr(E_2) < \epsilon$.

Therefore, the genie-aided decoder decodes the pair $(\mathbf{X}, \mathbf{S})$ with error probability less than or equal to 2ϵ for all $0 < \epsilon \leq$ for some specific thinned interference codebook in A_{n_k}. This shows that the rate pair $(\tilde{R}_I, R)$ is achievable for a Gaussian MAC channel. Since $R > \frac{1}{2}\log\left(1 + \frac{I}{N}\right)$, by choosing γ sufficiently small enough, the point $(\tilde{R}_I, R)$ can be forced to lie outside the two user Gaussian MAC capacity region. This is a contradiction.

Therefore, the theorem is true. ■

Hence in this regime, knowing the interference codebook is not useful.

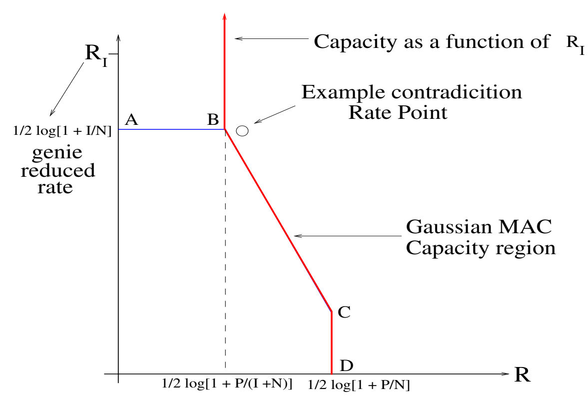

Fig. 4. The red line is the capacity of our noise plus interference channel as a function of the interference rate R_I. This is for the case of a random Gaussian interferer

A. What happens when interference is decodable?

For rates $R_I < \frac{1}{2}\log\left(1 + \frac{I}{N}\right)$ we can do better than treating interference as noise by exploiting the knowledge of the interference codebook. The red curve in Figure 4 gives the capacity of our channel as a function of the interference rate R_I. When $R_I < \frac{1}{2}\log\left(1 + \frac{I}{N}\right)$ we can appeal to the

[6]It is important to note that the input codebook can depend on the particular interference codebook.

converse of the Gaussian MAC capacity region and show that we cannot achieve rates higher than those given by the red curve in Figure 4. Here we don't need a genie to help us out.

To show achievability we consider the following cases:

- $R_I < \frac{1}{2}\log\left(1+\frac{I}{P+N}\right)$: In this case we can first decode the interference codeword by treating our message as noise and then decode our message by canceling out the interference codeword.
- $\frac{1}{2}\log\left(1+\frac{I}{P+N}\right) < R_I < \frac{1}{2}\log\left(1+\frac{I}{N}\right)$: This corresponds to the ramp in Figure 4. Points on the ramp can be achieved by using successive cancellation as in the achievability of the Gaussian MAC region. Specifically, this can be achieved by splitting our sender into two users (1 and 2) with power λ and $P-\lambda$ for some $0 \leq \lambda \leq P$. Then the decoder can first decode user 1 by treating everything else as noise, then decode the interference by treating user 2 as noise and finally we can decode user 2 seeing no interference. This scheme will achieve points strictly below the ramp in Figure 4. Hence we have proved the capacity region in Figure 4.

V. Extension to multiple interferers

The result in Theorem 1 states that under the product Gaussian measure, most interference codebooks are not worth knowing. However, there is no reason to restrict to product Gaussian measures. In this section we extend our result to measures formed by superposition of product Gaussian measures.

For instance, we can generate a $2^{nR_{I_i}}$ codebook, $\mathcal{S}_i$, according to the product Gaussian measure with average power constraint I_i for $i = 1, 2$. Here, $R_I = R_{I_1} + R_{I_2}$ and $I = I_1 + I_2$. Now, we can form the final codebook $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$. This imposes a measure on the set of all rate R_I with average power constraint I.

If each R_{I_i} is such that we cannot decode a part of the code treating the other as interference, then we can show that most interference codebooks are not worth knowing even under this measure. That is, if $R_{I_1} > \frac{1}{2}\log\left(1+\frac{I_1}{I_2+N}\right)$, $R_{I_1} > \frac{1}{2}\log\left(1+\frac{I_2}{I_1+N}\right)$ and the sum rate $R_I = R_{I_1} + R_{I_2} > \frac{1}{2}\log\left(1+\frac{I}{N}\right)$ then we are limited to treating interference as noise for essentially all interference codebooks. This result can be proved by the same argument as in the proof of Theorem 1 by use of a contradiction argument. We need to consider a genie, who thins both the codebooks $\mathcal{S}_i$ appropriately and reveals the thinned list to both the encoder and decoder. Using this genie-aided decoder we can show that, if $R > \frac{1}{2}\log\left(1+\frac{P}{I+N}\right)$ then a rate point outside the 3 user Gaussian MAC region is achievable, which leads to the desired contradiction.

We can also extend the same argument to measures formed by superposition of K product Gaussian measures. In this case, the interference codebook is formed by a sum of I.I.D. random Gaussian codebooks with rate R_{I_k} and power I_k, where $R_I = \sum_{k=1}^{K} R_{I_k}$ and $I = \sum_{k=1}^{K} I_k$. Here we consider the case when the individual rates are such that we cannot decode any subset of the interference codewords by treating the rest as noise, even when our transmitter is off. In this case, if $R > \frac{1}{2}\log\left(1+\frac{P}{I+N}\right)$, then most of the interference codebooks are not worth knowing.

Considering these superposition measures is practical when the interference is coming from the uplink of a MAC channel, i.e., from multiple Gaussian interferers. Therefore, the converse result can be directly extended to the case of multiple interferers.

VI. General discussion of "good" codes

Consider and interferer with two receivers, receiver 1 with Gaussian noise power N_1 and receiver 2 with Gaussian noise power N_2. Let us assume that $N_1 \ll N_2$. We assume that the interferer uses superposition coding to transmit at rates $R_{I_1} = \frac{1}{2}\log\left(1+\frac{\alpha I}{N_1}\right)$ and $R_{I_2} = \frac{1}{2}\log\left(1+\frac{(1-\alpha)I}{\alpha I+N_2}\right)$ for some $0 \leq \alpha \leq 1$. To send the message index pair (i, j) to its receivers, the interferer takes codeword $\mathbf{s}_1(i)$ from the first codebook and codeword $\mathbf{s}_2(j)$ from the second codebook and computes the sum. He then sends this sum over the channel.

Given this interferer setup, the channel seen by our receiver can be written as $\mathbf{Y} = \mathbf{X} + \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{Z}$. If $R_{I_1} + R_{I_2} > \frac{1}{2}\log\left(1+\frac{I}{N}\right)$ then it is clear that our receiver cannot decode all the interference even when our sender is off. In this case if $N + P \leq N_2$, then our receiver is less noisy than the worst receiver of the interferer's channel, i.e., receiver 2 in the above setup. Hence our receiver can decode $\mathbf{S}_2$ treating the remaining part of $\mathbf{Y}$ as noise. Then by subtracting $\mathbf{S}_2$ from $\mathbf{Y}$ our receiver faces interference from $\mathbf{S}_1$ only, which has power αI. Therefore our sender can transmit at $R < \frac{1}{2}\log\left(1+\frac{P}{\alpha I+N}\right)$ which is strictly higher than $\frac{1}{2}\log\left(1+\frac{P}{I+N}\right)$. Hence this is an example where we can do better than treating interference as noise. This example shows that the result in Theorem 1 is not true for all interference codebooks.

Let's re-examine this code from a point to point channel view for the interferer. For the given power constraint I, this code achieves a much lower rate than the capacity $\frac{1}{2}\log\left(1+\frac{I}{N}\right)$. Hence, this is a bad point to point channel code, in that it uses a lot more power than is necessary for the given rate. This badness is the reason why we can decode a part of the codeword. From our proof point of view this codebook is bad because it has very few cloud centers and dense clouds around each cloud center. Therefore, we cannot thin this codebook as required by the proof of Theorem 1.

VII. Concluding remarks

In this paper we tried to address the important question – "Is interference like noise when its codebook is known?". We have shown that, in the case of a high rate (undecodable) non-interactive interferer, we cannot do better than treating the interference as additional power constrained noise, even when its codebook is known. This result holds for essentially all interference codebooks under the product Gaussian measures.

We also verified our result for other measures, like superposition of product Gaussian codebooks.

We have also shown by the example in Section VI that there exists some interference codebooks that are good for the secondary, i.e., the secondary can achieve higher rates when the interferer uses these codebooks.

We conclude by proposing the following conjecture

***Conjecture** 1:* If a interference codebook is "good" for the secondary user from a rate point of view, then it must be a bad rate code from the interferer's point of view.

REFERENCES

[1] I. J. Mitola, "Software radios: Survey, critical evaluation and future directions," *IEEE Aerosp. Electron. Syst. Mag*, vol. 8, pp. 25–36, Apr. 1993.

[2] FCC, "FCC 03-322," Dec. 2003. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-03-322A1.pdf

[3] ——, "FCC 03-289," Nov. 2003. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-03-289A1.pdf

[4] A. Lapidoth, "Nearest neighor decoding for additive non-gaussian noise channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1520–1529, September 1996.

[5] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1813–1827, May 2006.

[6] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[7] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Tran. Inform. Theory*, Submitted April 2006.

[8] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Forty-second Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2004.

[9] R. Tandra, "Fundamental limits on detection in low SNR," Master's thesis, University of California, Berkeley, Berkeley, 2005.

[10] A. Sahai, N. Hoven, S. M. Mishra, and R. Tandra, "Fundamental tradeoffs in spectrum sensing for opportunistic frequency reuse," *Journal on Selected Areas in Communications*, submitted February 2006.