

SPATIAL-SPECTRAL JOINT DETECTION FOR WIDEBAND SPECTRUM SENSING IN COGNITIVE RADIO NETWORKS

Zhi Quan[†], Shuguang Cui[‡], Ali H. Sayed[†], and H. Vincent Poor[§]

[†]Department of Electrical Engineering, University of California, Los Angeles, CA 90095

[‡]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843

[§]Department of Electrical Engineering, Princeton University, Princeton, NJ 08544

Email: {quan, sayed}@ee.ucla.edu, cui@tamu.edu, poor@princeton.edu

ABSTRACT

Spectrum sensing is an essential functionality that enables cognitive radios to detect spectral holes and opportunistically use under-utilized frequency bands without causing harmful interference to primary networks. Since individual cognitive radios might not be able to reliably detect weak primary signals due to channel fading/shadowing, this paper proposes a cooperative wideband spectrum sensing scheme, referred to as *spatial-spectral joint detection*, which is based on a linear combination of the local statistics from spatially distributed multiple cognitive radios. The cooperative sensing problem is formulated into an optimization problem, for which suboptimal but efficient solutions can be obtained through mathematical transformation under practical conditions.

Index Terms— Spectrum sensing, distributed detection, nonlinear optimization, and cognitive radio.

1. INTRODUCTION

As an essential functionality of cognitive radio (CR) networks [1], spectrum sensing needs to reliably detect weak primary radio signals of possibly-unknown formats. Generally, spectrum sensing techniques can be classified into three categories: energy detection, matched filter coherent detection [2], and cyclostationary feature detection. Since non-coherent energy detection is simple and able to generate the spectrum-occupancy information quickly, we adopt it as the building block for constructing the proposed wideband spectrum sensing schemes.

The literature on wideband spectrum sensing for CR networks is limited. An earlier approach is to use a tunable narrowband bandpass filter at the RF front-end to sense one narrow frequency band at a time, over which the existing narrowband spectrum sensing techniques can be applied. In order to search over multiple frequency bands at a time, the RF front-end needs a wideband architecture, and spectrum sensing usually operates over an estimate of the power spectral density (PSD) of the wideband signal. In [3], wavelet transformation

was used to estimate the PSD over a wide frequency range given its multi-resolution features. However, no prior work attempts to make decisions over multiple frequency bands jointly, which is essential for implementing efficient CR networks.

In this paper, we consider the situation in which spectrum sensing is compromised by destructive channel conditions between the target-under-detection and the detecting cognitive radios, which makes it hard to distinguish between a white spectrum and a weak signal. We propose a cooperative wideband spectrum sensing scheme that exploits the spatial diversity among cognitive radios to improve the sensing reliability. The cooperation is based on a linear combination of local statistics from spatially distributed cognitive radios [4] [5], where the signals are assigned different weights according to their positive contributions to joint sensing. In such a scenario, we treat the design of distributed wideband spectrum sensing as a *spatial-spectral joint detection* problem, which is further formulated into an optimization problem with the objective of maximizing the overall opportunistic throughput under constraints on the interference to primary users. Through mathematical reformulation, we derive a suboptimal but efficient solution for the optimization problem, which can considerably improve sensing performance.

2. SYSTEM MODEL

Consider a primary communication system (e.g., multicarrier based) over a wideband channel that is divided into K non-overlapping subchannels. At a particular time, some of the K subchannels might not be used by the primary users and are available for opportunistic spectrum access. Multiuser orthogonal frequency division multiplexing (OFDM) schemes are suitable candidates for such a scenario since they make it convenient to nullify or activate some portion of multiple narrow bands. We model the detection problem over the subband k as one to choose between hypothesis $\mathcal{H}_{0,k}$ (“0”), which represents the absence of primary signals, and hypothesis $\mathcal{H}_{1,k}$ (“1”), which represents the presence of primary signals. An illustration where only some of the K bands are occupied by primary users is illustrated in Fig. 1. The crucial task of spec-

This research was supported in part by NSF under Grants ANI-03-38807, CNS-06-25637, ECS-06-01266, ECS-07-25441, CNS-06-25637, and by DoD under Grant HOTRN-07-1-0037.

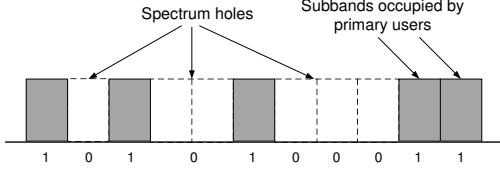


Fig. 1. Illustration of the occupancy of a multiband channel.

trum sensing is to sense the K frequency bands and identify spectral holes for opportunistic use. For simplicity, we assume that the high-layer protocols guarantee that all CRs keep quiet during the detection such that the main spectral power under detection is emitted by the primary users.

Consider a multi-path fading environment, where $h(l)$, $l = 0, 1, \dots, L-1$, denotes the discrete-time channel impulse response between the primary transmitter and a CR receiver with L equal to the number of resolvable paths. The received baseband signal at the CR front-end can be expressed as

$$r(n) = \sum_{l=0}^{L-1} h(l) s(n-l) + v(n), \quad n = 0, 1, \dots, N_0-1 \quad (1)$$

where $s(n)$ represents the primary transmitted signal with cyclic prefix at time n and $v(n)$ is additive complex white Gaussian noise with zero mean and variance σ_v^2 , i.e., $v(n) \sim \mathcal{CN}(0, \sigma_v^2)$. In a multi-path fading environment, the wide-band channel exhibits frequency-selective features and its discrete frequency responses are given by

$$H_k = \frac{1}{\sqrt{N_0}} \sum_{n=0}^{L-1} h(n) e^{-j2\pi nk/N_0}, \quad k = 0, 1, \dots, K-1 \quad (2)$$

where $L \leq N_0$. We assume that the channel is slowly varying and the channel frequency responses $\{H_k\}_{k=0}^{K-1}$ do not vary much during a detection interval. In the frequency domain, the received signal at each subchannel can be estimated by computing its discrete Fourier transform (DFT):

$$R_k = \frac{1}{\sqrt{N_0}} \sum_{n=0}^{N_0-1} r(n) e^{-j2\pi nk/N_0} = H_k S_k + V_k \quad (3)$$

where S_k is the primary signal at subchannel k and

$$V_k = \frac{1}{\sqrt{N_0}} \sum_{n=0}^{L-1} v(n) e^{-j2\pi nk/N_0} \quad (4)$$

is the received noise in the frequency domain. Note that $V_k \sim \mathcal{CN}(0, \sigma_v^2)$ since $v(n) \sim \mathcal{CN}(0, \sigma_v^2)$ and the DFT is a unitary linear operation. Without loss of generality, we assume that the transmitted signal S_k , channel gain H_k , and additive noise V_k are independent of each other.

To decide whether the k -th subchannel is occupied or not, we test the following binary hypotheses:

$$\begin{aligned} \mathcal{H}_{0,k} : R_k &= V_k, & k &= 0, 1, \dots, K-1 \\ \mathcal{H}_{1,k} : R_k &= H_k S_k + V_k, & k &= 0, 1, \dots, K-1 \end{aligned} \quad (5)$$

For each subchannel k , we compute the test statistic as the sum of received energy over an interval of M samples, i.e.,

$$Y_k = \sum_{m=0}^{M-1} |R_k(m)|^2 \quad k = 0, 1, \dots, K-1 \quad (6)$$

and the decision rule is given by

$$Y_k \underset{\mathcal{H}_{0,k}}{\overset{\mathcal{H}_{1,k}}{\geq}} \gamma_k \quad (7)$$

where γ_k is the corresponding decision threshold. For simplicity, we assume that the transmitted signal at each subchannel has unit power, i.e., $\mathbb{E}(|S_k|^2) = 1$; this assumption holds when the primary radios adopt uniform power transmission strategies given no channel knowledge at the transmitter side.

According to the central limit theorem for large M , Y_k is asymptotically normally distributed with mean

$$\mathbb{E}(Y_k) = \begin{cases} M\sigma_v^2 & \mathcal{H}_{0,k} \\ M(\sigma_v^2 + |H_k|^2) & \mathcal{H}_{1,k} \end{cases} \quad (8)$$

and variance

$$\text{Var}(Y_k) = \begin{cases} 2M\sigma_v^4 & \mathcal{H}_{0,k} \\ 2M(\sigma_v^2 + 2|H_k|^2)\sigma_v^2 & \mathcal{H}_{1,k} \end{cases} \quad (9)$$

for $k = 0, 1, \dots, K-1$. Thus, assuming large M , we have $Y_k \sim \mathcal{N}(\mathbb{E}(Y_k), \text{Var}(Y_k))$.

Using the decision rule in (7), the probabilities of false alarm and detection at subchannel k can be respectively calculated approximately as

$$P_f^{(k)}(\gamma_k) = \Pr(Y_k > \gamma_k | \mathcal{H}_0) = Q\left(\frac{\gamma_k - M\sigma_v^2}{\sigma_v^2 \sqrt{2M}}\right) \quad (10)$$

and

$$P_d^{(k)}(\gamma_k) = Q\left(\frac{\gamma_k - M(\sigma_v^2 + |H_k|^2)}{\sigma_v \sqrt{2M(\sigma_v^2 + 2|H_k|^2)}}\right) \quad (11)$$

where Q denotes the tail probability of the standard normal distribution. The choice of threshold γ_k leads to a tradeoff between the probabilities of false alarm and miss $P_m = 1 - P_d$. Specifically, a higher threshold will result in a smaller probability of false alarm but a larger probability of miss, and vice versa.

3. SPATIAL-SPECTRAL JOINT DETECTION

Suppose that N spatially distributed cognitive radios collaboratively sense a wide frequency band. By combining the local statistics from individual cognitive radios at the fusion center, which can be one of the CRs, the network can make a better decision on the presence or absence of primary signals on each of the K subchannels. The cooperation assumes a separate control channel, through which the statistics of individual CRs are transmitted to the fusion center. Let $Y_k(n)$ denote the received energy in the k -th subchannel at cognitive radio n . For each subchannel, these statistics can be written in a vector as $\mathbf{Y}_k = [Y_k(0), Y_k(1), \dots, Y_k(N-1)]^T$.

To exploit the spatial diversity, we linearly combine the summary statistics from spatially distributed cognitive radios at each subchannel k to obtain a final test statistic:

$$z_k = \sum_{n=0}^{N-1} w_k(n) Y_k(n) = \mathbf{w}_k^T \mathbf{Y}_k \quad (12)$$

where $\mathbf{w}_k = [w_k(0), w_k(1), \dots, w_k(N-1)]^T$ are the combining coefficients for subchannel k , which can be compactly written in a matrix as $\mathbf{W} = [\mathbf{w}_0 \ \mathbf{w}_1 \ \dots \ \mathbf{w}_{K-1}]$. Note that $w_k(n) \geq 0$, for every k and n .

Since the entries in \mathbf{Y}_k are normally distributed, the test statistics $\{z_k\}_{k=0}^{K-1}$ are also normally distributed with means

$$\mathbb{E}(z_k) = \begin{cases} M\sigma_v^2 \mathbf{w}_k^T \mathbf{1} & \mathcal{H}_{0,k} \\ M\mathbf{w}_k^T (\sigma_v^2 \mathbf{1} + \mathbf{G}_k) & \mathcal{H}_{1,k} \end{cases} \quad (13)$$

where $\mathbf{1}$ is an all-one vector, and variances

$$\text{Var}(z_k) = \begin{cases} 2M\sigma_v^4 \mathbf{w}_k^T \mathbf{w}_k & \mathcal{H}_{0,k} \\ 2M\sigma_v^2 \mathbf{w}_k^T [\sigma_v^2 \mathbf{I} + 2\text{diag}(\mathbf{G}_k)] \mathbf{w}_k & \mathcal{H}_{1,k} \end{cases} \quad (14)$$

where $\mathbf{G}_k = [|H_k(0)|^2, |H_k(1)|^2, \dots, |H_k(N-1)|^2]^T$ are the squared magnitudes of the channel gains between the primary transmitter and the N CR receivers for subchannel k .

In order to decide the presence or absence of the primary signal in subchannel k , we use the following binary test

$$z_k \underset{\mathcal{H}_{0,k}}{\overset{\mathcal{H}_{1,k}}{\geq}} \gamma_k, \quad k = 0, 1, \dots, K-1. \quad (15)$$

Accordingly, the detection performance in terms of the probabilities of false alarm and detection are given by

$$P_f^{(k)}(\mathbf{w}_k, \gamma_k) = Q\left(\frac{\gamma_k - M\mathbf{w}_k^T \mathbf{1}}{\sigma_v \sqrt{2M\mathbf{w}_k^T \mathbf{w}_k}}\right) \quad (16)$$

and

$$P_d^{(k)}(\mathbf{w}_k, \gamma_k) = Q\left(\frac{\gamma_k - M\mathbf{w}_k^T (\sigma_v^2 \mathbf{1} + \mathbf{G}_k)}{\sigma_v \sqrt{2M\mathbf{w}_k^T [\sigma_v^2 \mathbf{I} + 2\text{diag}(\mathbf{G}_k)] \mathbf{w}_k}}\right) \quad (17)$$

For compactness of notation, we collect the probabilities of false alarm and detection over the K subchannels into vectors $\mathbf{P}_f(\mathbf{W}, \boldsymbol{\gamma})$ and $\mathbf{P}_d(\mathbf{W}, \boldsymbol{\gamma})$. Thus, the probabilities of miss can be represented as $\mathbf{P}_m(\mathbf{W}, \boldsymbol{\gamma}) = \mathbf{1} - \mathbf{P}_d(\mathbf{W}, \boldsymbol{\gamma})$.

Our goal is to maximize the opportunistic rate while meeting some constraints on the interference to the primary communication system. Let r_k denote the throughput achievable over the k -th subchannel if used by cognitive radios, and $\mathbf{r} = [r_0, r_1, \dots, r_{K-1}]^T$. Since $1 - P_f^{(k)}$ measures the opportunistic spectrum utilization of subchannel k , we define the aggregate opportunistic throughput capacity as

$$R(\mathbf{W}, \boldsymbol{\gamma}) = \mathbf{r}^T [\mathbf{1} - \mathbf{P}_f(\mathbf{W}, \boldsymbol{\gamma})]. \quad (18)$$

For a wideband primary system, the impact of interference induced by cognitive devices can be characterized by a relative priority vector over the K subchannels, i.e., $\mathbf{c} = [c_0, c_1, \dots, c_{K-1}]^T$, where c_k indicates the cost incurred if the primary user at subchannel k is interfered with. As such, we define the aggregate interference to the primary user as

$\mathbf{c}^T \mathbf{P}_m(\mathbf{W}, \boldsymbol{\gamma})$. Consequently, the spatial-spectral joint detection problem is formulated as

$$\max_{\mathbf{W}, \boldsymbol{\gamma}} R(\mathbf{W}, \boldsymbol{\gamma}) \quad (\text{P1})$$

$$\text{s.t.} \quad \mathbf{c}^T \mathbf{P}_m(\mathbf{W}, \boldsymbol{\gamma}) \leq \varepsilon \quad (19)$$

$$\mathbf{P}_m(\mathbf{W}, \boldsymbol{\gamma}) \preceq \boldsymbol{\alpha} \quad (20)$$

$$\mathbf{P}_f(\mathbf{W}, \boldsymbol{\gamma}) \preceq \boldsymbol{\beta} \quad (21)$$

where $\boldsymbol{\alpha} = [\alpha_0, \dots, \alpha_{K-1}]^T$ and $\boldsymbol{\beta} = [\beta_0, \dots, \beta_{K-1}]^T$.

Finding the exact solution for the above problem is difficult since for any k , $P_f^{(k)}(\mathbf{w}_k, \gamma_k)$ and $P_d^{(k)}(\mathbf{w}_k, \gamma_k)$ are neither convex nor concave functions according to (16) and (17). To jointly optimize \mathbf{W} and $\boldsymbol{\gamma}$, we can show that (P1) can be reformulated into an equivalent form with convex constraints and an objective function lower bounded by a concave function under the following conditions:

$$0 < \alpha_k \leq \frac{1}{2} \quad \text{and} \quad 0 < \beta_k \leq \frac{1}{2}, \quad k = 0, \dots, K-1. \quad (22)$$

Through maximizing the lower bound of the objective function, we are able to obtain a good approximation to the optimal solution of the original problem.

First, we show how to transform the nonconvex constraints in (20) and (21) into convex constraints by exploiting the monotonicity of the Q -function. Substituting (16) into the constraint (21), we have

$$Q^{-1}(\beta_k) \sqrt{2M\mathbf{w}_k^T \mathbf{w}_k} \leq \frac{\gamma_k}{\sigma_v} - M\mathbf{w}_k^T \mathbf{1} \quad (23)$$

where $Q^{-1}(\beta_k) \geq 0$ given $\beta_k \leq 1/2$. From (17), the constraint (20) can be expressed as

$$\sqrt{2M\mathbf{w}_k^T [\sigma_v^2 \mathbf{I} + 2\text{diag}(\mathbf{G}_k)] \mathbf{w}_k} \leq \frac{\gamma_k - M\mathbf{w}_k^T (\sigma_v^2 \mathbf{1} + \mathbf{G}_k)}{\sigma_v Q^{-1}(1 - \alpha_k)} \quad (24)$$

given $\alpha_k \leq 1/2$ and $Q^{-1}(1 - \alpha_k) \leq 0$. Since the left-hand side on the constraint (23) is convex and the right hand side is linear in (γ_k, \mathbf{w}_k) , (23) defines a convex set for (γ_k, \mathbf{w}_k) . Similarly, (24) is also a convex constraint.

Then, we reformulate (P1) by introducing a new variable

$$\mu_k = \sigma_v \sqrt{2M\mathbf{w}_k^T [\sigma_v^2 \mathbf{I} + 2\text{diag}(\mathbf{G}_k)] \mathbf{w}_k}. \quad (25)$$

By defining $\gamma'_k = \gamma_k/\mu_k$ and $\mathbf{w}'_k = \mathbf{w}_k/\mu_k$, the constraints (23) and (24) can be further written as

$$Q^{-1}(\beta_k) \sqrt{2M\mathbf{w}'_k{}^T \mathbf{w}'_k} \leq \frac{\gamma'_k}{\sigma_v} - M\mathbf{1}^T \mathbf{w}'_k \quad (26)$$

and

$$\gamma'_k - M(\sigma_v^2 \mathbf{1} + \mathbf{G}_k)^T \mathbf{w}'_k \leq \sigma_v Q^{-1}(1 - \alpha_k). \quad (27)$$

Note that (27) is actually a linear constraint in $(\gamma'_k, \mathbf{w}'_k)$. The constraint (19) now becomes

$$\mathbf{1}^T \mathbf{c} - \sum_{k=0}^{K-1} c_k Q\left(\gamma'_k - M(\sigma_v^2 \mathbf{1} + \mathbf{G}_k)^T \mathbf{w}'_k\right) \leq \varepsilon, \quad (28)$$

which can be shown to be convex by the following result.

Lemma 1 If $\gamma'_k \leq M(\sigma_v^2 \mathbf{1} + \mathbf{G}_k)^T \mathbf{w}'_k$, then the function $Q(\gamma'_k - M(\sigma_v^2 \mathbf{1} + \mathbf{G}_k)^T \mathbf{w}'_k)$ is concave in $\{\gamma'_k, \mathbf{w}'_k\}$.

By changing the variables \mathbf{W}' and γ' , $P_f^{(k)}(\mathbf{w}_k, \gamma_k)$ can be expressed as

$$Q\left[\left(\frac{\gamma'_k}{\sigma_v^2} - M\mathbf{1}^T \mathbf{w}'_k\right) \sqrt{\sigma_v^2 + \frac{2\mathbf{w}'_k{}^T \text{diag}(\mathbf{G}_k) \mathbf{w}'_k}{\mathbf{w}'_k{}^T \mathbf{w}'_k}}\right] \quad (29)$$

From the the Rayleigh-Ritz theorem [6], we have

$$\min_n |H_k(n)|^2 \leq \frac{\mathbf{w}'_k{}^T \text{diag}(\mathbf{G}_k) \mathbf{w}'_k}{\mathbf{w}'_k{}^T \mathbf{w}'_k} \leq \max_n |H_k(n)|^2 \quad (30)$$

Define a new function

$$g_k(\gamma'_k, \mathbf{w}'_k) \triangleq Q\left[\left(\frac{\gamma'_k}{\sigma_v^2} - M\mathbf{1}^T \mathbf{w}'_k\right) \sqrt{\sigma_v^2 + 2 \min_n |H_k(n)|^2}\right] \quad (31)$$

which can be shown to be convex by the following result.

Lemma 2 If $\gamma'_k \geq \sigma_v^2 M\mathbf{1}^T \mathbf{w}'_k$, then the function $g_k(\gamma'_k, \mathbf{w}'_k)$ is convex in $\{\gamma'_k, \mathbf{w}'_k\}$.

Since $P_f^{(k)}(\mathbf{w}_k, \gamma_k) \leq g_k(\gamma'_k, \mathbf{w}'_k)$, the objective function in (P1) can be lower bounded by $\sum_{k=0}^{N-1} r_k [1 - g_k(\gamma'_k, \mathbf{w}'_k)]$, which is a concave function. Thus, an efficient suboptimal method to solve (P1) is to maximize the lower bound of its objective function, i.e.,

$$\begin{aligned} \max_{\mathbf{W}', \gamma'} \quad & \sum_{k=0}^{N-1} r_k [1 - g_k(\gamma'_k, \mathbf{w}'_k)] \quad (P2) \\ \text{st.} \quad & - \sum_{k=0}^{K-1} c_k Q\left[\gamma'_k - M(\sigma_v^2 \mathbf{1} + \mathbf{G}_k)^T \mathbf{w}'_k\right] \leq \varepsilon - \mathbf{1}^T \mathbf{c} \\ & Q^{-1}(\beta_k) \sqrt{2M\mathbf{w}'_k{}^T \mathbf{w}'_k} \leq \frac{\gamma'_k}{\sigma_v^2} - M\mathbf{1}^T \mathbf{w}'_k \\ & \gamma'_k - M(\sigma_v^2 \mathbf{1} + \mathbf{G}_k)^T \mathbf{w}'_k \leq \sigma_v Q^{-1}(1 - \alpha_k). \end{aligned}$$

Implied by the practical conditions in (22), this problem is a convex optimization problem and can be solved efficiently.

4. SIMULATIONS

Suppose that two CRs cooperatively sense a multiband OFDM system with 8 subbands. For each subband, it is expected that the opportunistic spectrum utilization is at least 50%, i.e., $\beta_k = 0.5$, and the probability that the primary user is interfered is at most $\alpha_k = 0.1$. It is assumed that $\sigma_v^2 = 1$ and $M = 100$. Other parameters are given in Table 1. Fig. 2 shows result of solving (P2), which maximizes the opportunistic throughput subject to the constraints on the interference. We observe that the joint detection results in much higher opportunistic throughput than the algorithms without cooperation. Note that the increase in the throughput of the joint optimization scheme becomes rather slow as we relax the interference constraint because the interaction between γ and \mathbf{W} pushes the system to an operating point at which the throughput is more limited by β than by ε .

Table 1. Parameters used in simulations

$\mathbf{G}(0)$.17	.21	.27	.14	.37	.38	.49	.33
$\mathbf{G}(1)$.21	.17	.21	.21	.17	.43	.15	.35
r	356	327	972	806	755	68	720	15
c	.71	5.95	3.91	4.21	.44	2.03	.58	2.85

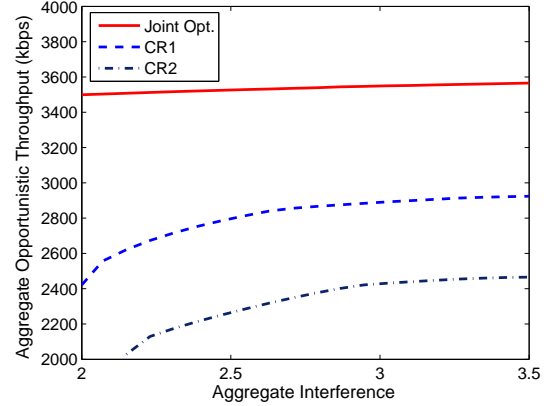


Fig. 2. Aggregate opportunistic throughput capacity vs. the constraint on the aggregate induced interference.

5. CONCLUSION

In this paper, we have proposed a spatial-spectral joint detection framework for distributed wideband spectrum sensing in cognitive radio networks, within which the cooperation among spatially distributed cognitive radios is optimized over multiple frequency bands. By exploiting the inherent structure of the formulation, we have developed suboptimal but efficient solutions for the non-convex optimization problem. This paper establishes important principles for the design of distributed wideband spectrum sensing algorithms in cognitive radio networks.

6. REFERENCES

- [1] J. Mitola III, "The software radio architecture," *IEEE Trans. Comm.*, vol. 33, no. 5, pp. 26–38, May 1995.
- [2] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1994.
- [3] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proc. CROWNCOM*, Mykonos Island, Greece, June 2006, pp. 1–5.
- [4] Z. Quan, S. Cui, and A. H. Sayed, "An optimal strategy for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE GLOBECOM*, Washington D.C., Nov. 2007, pp. 2947–2951.
- [5] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, June 2008.
- [6] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, Hoboken, NJ, 2003.