WIDEBAND SPECTRAL ESTIMATION FROM COMPRESSED MEASUREMENTS EXPLOITING SPECTRAL *A PRIORI* INFORMATION IN COGNITIVE RADIO SYSTEMS

Gonzalo Vazquez-Vilar, Roberto López-Valcarce, Carlos Mosquera and Nuria González-Prelcic

Departamento de Teoría de la Señal y las Comunicaciones
Universidad de Vigo, 36310 Vigo, Spain
email: {gvazquez, valcarce, mosquera, nuria}@gts.tsc.uvigo.es

ABSTRACT

In Cognitive Radio scenarios channelization information from primary network may be available to the spectral monitor. Under this assumption we propose a spectral estimation algorithm from compressed measurements of a multichannel wideband signal. The analysis of the Cramer-Rao Lower Bound (CRLB) for this estimation problem shows the importance of detecting the underlaying sparsity pattern of the signal. To this end we describe a Bayesian based iterative algorithm that discovers the set of active signals conforming the band and simultaneously reconstructs the spectrum. This iterative spectral estimator is shown to perform close to a Genie-Aided CRLB that includes full knowledge about the sparsity pattern of the channels.

Index Terms— Spectral estimation, compressive sampling, cognitive radio

1. INTRODUCTION

Cognitive Radio (CR) is receiving considerable interest as a means for wireless systems to improve spectral efficiency. The key idea of opportunistically accessing temporally and/or spatially unused licensed bands implicitly assumes a powerful spectral monitor with the capability of scanning multiple channels in a wideband scenario. The large bandwidth involved makes Nyquist-rate wideband monitoring impractical, due to power consumption and analog implementation complexity constraints. Compressed sampling could result in significant savings in data rate (provided that one can find a domain in which the signal is sparse), hopefully maintaining the required spectral resolution.

Assuming a spectrum model consisting of several flat bandpass signals, and considering the edges between them, the observed signal is sparse in the "spectral edges domain". This fact is used in [1] to propose a spectrum reconstruction algorithm from compressed samples of the input autocorrelation estimate. This method was extended in [2] in order to work directly from compressed measurements of the received signal. These methods do not assume information about the primary network channelization, so that the spectral edges could assume any position within the frequency band.

We exploit *a priori* knowledge about the spectral shape (not necessarily flat) and location of the signals conforming the spectrum. This is reasonable in the CR context, since the channelization and modulation parameters of the primary system are often known (e.g. broadcast networks). Moreover, the CR paradigm is based on the infrautilization of spectral resources, and hence we can expect that only an (unknown) subset of the constituent signals will be simultaneously active at a given location and frequency band. Exploiting such sparsity in the "activity domain", we propose a spectrum estimator from compressed samples performing close to a Genie Aided Bound that assumes full knowledge on the sparsity pattern.

This paper is organized as follows. Section 2 formalizes the estimation problem, Section 3 gives the performance bounds, and Section 4 describes a Bayesian based spectral estimator. Numerical results and final conclusions are presented in Sections 5 and 6, respectively.

2. PROBLEM STATEMENT

Primary users employ Frequency Division Multiplexing with fixed channelization, known to the spectrum monitor. Several primary channels are sensed simultaneously, by selecting a wide band containing M of such channels, downconverting it to baseband and sampling the resulting analog signal through an Analog-to-Information (A2I) converter that produces samples at a rate below the Nyquist rate. The baseband analog signal at the receiver after wideband filtering is given by

$$r(t) = \sum_{m=1}^{M} \sigma_m x_m(t) + \sigma w(t), \tag{1}$$

where $w(t)$ is a zero mean, circular complex Gaussian noise with unit variance; σ^2 is the background noise power; $x_m(t)$ is the (noiseless) primary signal in channel m, normalized to have unit variance ($E\{|x_m(t)|^2\} = 1$); and σ_m^2 is the power of the primary signal in the mth channel. Assuming

This work was supported by the Spanish Government under projects SPROACTIVE (reference TEC2007-68094-C02-01/TCM) and COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010).

that primary transmissions employ Multicarrier Modulation, the signals $\{x_m(\cdot)\}$ can be modeled as wide-sense stationary, zero mean circular Gaussian proccesses. Since $\{x_m(\cdot)\}$ correspond to different primary transmissions, they are assumed statistically independent.

Compression model. We restrict our study to linear A2I converters, that is, converters that can be represented in matrix form from an oversampled version of the analog signal $r(t)$. For compactness we define $x_0(t) \doteq w(t)$ and $\sigma_0^2 \doteq \sigma^2$. The finite discrete representation of (1) at Nyquist rate using the obvious vector notation can be written as

$$\boldsymbol{r} = \sum_{m=0}^{M} \sigma_m \boldsymbol{x}_m, \tag{2}$$

where $\boldsymbol{r}$ and $\boldsymbol{x}_m$ with $m = 0, \ldots, M$ are now $N \times 1$ circular Gaussian vectors, with zero mean and covariance matrix $\boldsymbol{C}_m \doteq E\{\boldsymbol{x}_m \boldsymbol{x}_m^H\}$. Due to the normalization and stationarity of the original processes $\{x_m(\cdot)\}$, $\boldsymbol{C}_m$ are Toeplitz with ones on the diagonal. If we define the $K \times N$ compression matrix $\boldsymbol{\Phi}$, with $K < N$, we can write the signal available to the digital spectrum monitor as

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{r} = \sum_{m=0}^{M} \sigma_m \tilde{\boldsymbol{x}}_m, \qquad \text{with} \quad \tilde{\boldsymbol{x}}_m \doteq \boldsymbol{\Phi}\boldsymbol{x}_m. \tag{3}$$

Thus $\boldsymbol{y}$ is zero-mean circular Gaussian with covariance

$$\tilde{\boldsymbol{R}} \doteq E\{\boldsymbol{y}\boldsymbol{y}^H\} = \sum_{m=0}^{M} \sigma_m^2 \tilde{\boldsymbol{C}}_m, \tag{4}$$

where $\tilde{\boldsymbol{C}}_m \doteq E\{\tilde{\boldsymbol{x}}_m \tilde{\boldsymbol{x}}_m^H\} = \boldsymbol{\Phi}\boldsymbol{C}_m\boldsymbol{\Phi}^H$. We say that the mth channel is vacant if $\sigma_m^2 = 0$. Then we can define the set of active channels as

$$\mathcal{S} = \left\{ m \mid \sigma_m^2 \neq 0,\ 0 \leq m \leq M \right\}. \tag{5}$$

It is assumed that the noise is always present and thus $0 \in \mathcal{S}$ always. For the signal channels, we model the sparsity of the system with each event $m \in \mathcal{S}$ following an independent Bernoulli distribution: $\text{Prob}\{m \in \mathcal{S}\} = p_1$ for $m = 1, \ldots, M$, with p_1 assumed known to the receiver (p_1 gives an indication of the average occupancy of the frequency band and can be estimated beforehand; in a CR context, it is expected that $p_1 \ll 1$). On the other hand no assumption is made on σ_m^2 given that channel m is active.

Problem statement. The goal is to estimate the power spectral density (psd) of $r(t)$, defined as $\mathsf{S}_r(e^{j\omega})$, from the compressed sampled vector $\boldsymbol{y}$. The digital processing unit is assumed to have perfect knowledge of the compression matrix $\boldsymbol{\Phi}$ and the normalized covariance matrices $\boldsymbol{C}_m$, $m = 0, \ldots, M$. Due to statistical independence, one has

$$\mathsf{S}_r(e^{j\omega}) = \sum_{m \in \mathcal{S}} \sigma_m^2 \mathsf{S}_m(e^{j\omega}), \tag{6}$$

with $\mathsf{S}_m(e^{j\omega})$ the psd of the signal $x_m(t)$. Since these normalized psds are assumed known, then under the parametric model adopted in this framework the problem reduces to estimating the sparse vector of power levels

$$\boldsymbol{\sigma} \doteq [\ \sigma_0^2 \quad \sigma_1^2 \quad \cdots \quad \sigma_M^2\ ]^T. \tag{7}$$

3. ON THE CRAMER-RAO LOWER BOUND

Suppose that one attempts to estimate $\boldsymbol{\sigma}$ from $\boldsymbol{y}$ without making any assumption on its sparsity. Given $\boldsymbol{\sigma}$, the observation $\boldsymbol{y}$ is zero mean circular Gaussian with covariance $\tilde{\boldsymbol{R}}(\boldsymbol{\sigma})$ as in (4). The elements of the Fisher information matrix (FIM) $\boldsymbol{F}(\boldsymbol{\sigma})$, of size $(M+1) \times (M+1)$, are given by (see e.g. [3]):

$$[\boldsymbol{F}(\boldsymbol{\sigma})]_{ij} = \text{Tr}\left\{ \tilde{\boldsymbol{R}}^{-1}(\boldsymbol{\sigma})\tilde{\boldsymbol{C}}_i \tilde{\boldsymbol{R}}^{-1}(\boldsymbol{\sigma})\tilde{\boldsymbol{C}}_j \right\}, \tag{8}$$

where we used that in our model $\partial\tilde{\boldsymbol{R}}(\boldsymbol{\sigma})/\partial\sigma_m^2 = \tilde{\boldsymbol{C}}_m$. The variance of any unbiased estimator $\hat{\boldsymbol{\sigma}}$ (ignoring sparsity) of the parameters of interest $\boldsymbol{\sigma}$ is lower bounded by the Cramer-Rao Lower Bound (CRLB),

$$\text{var}(\sigma_m^2) \geq \gamma_{\text{CRLB}}^{(m)} \doteq [\boldsymbol{F}(\boldsymbol{\sigma})^{-1}]_{mm}. \tag{9}$$

However, when only a subset $\mathcal{S}$ of the M signals is active we expect that this bound can be beaten. To show this imagine a genie that provides the receiver with *a priori* information about the set of active channels $\mathcal{S}$. Then, the spectral monitor only has to estimate the powers for the $|\mathcal{S}|$ active channels. The FIM for this genie aided estimation problem $\tilde{\boldsymbol{F}}$ can be derived from $\boldsymbol{F}$ by eliminating the rows and columns corresponding to the (known) inactive channels. That is, $\tilde{\boldsymbol{F}}$ keeps only the rows and columns with indices in the set $\mathcal{S} = \{m_1, m_2, \ldots, m_{|\mathcal{S}|}\}$:

$$\text{var}(\sigma_{m_i}^2) \geq \gamma_{\text{GACRLB}}^{(m_i)} \doteq [\tilde{\boldsymbol{F}}(\boldsymbol{\sigma})^{-1}]_{ii}, \tag{10}$$

and since the presence of nuisance parameters can only hurt [4], we must have $\gamma_{\text{GACRLB}}^{(m_i)} \leq \gamma_{\text{CRLB}}^{(m_i)}$. We refer to (10) as Genie Aided CRLB (GACRLB). Its significance resides in that it can be asymptotically achieved in some simplified estimation problems, as it was shown in [5] for estimators based on asymptotic typicality. Next we present practical estimation algorithms that perform close to the GACRLB in several cases of interest for the problem at hand.

4. ESTIMATION FROM COMPRESSED DATA

The previous discussion shows the importance of exploiting the *a priori* information available about the problem in order to approach the GACRLB. Therefore we pose here the problem of estimating the sparsity pattern $\mathcal{S}$ together with the power vector $\boldsymbol{\sigma}$. Using Bayes' rule we can state the Maximum

A Posteriori (MAP) estimation of $\{\boldsymbol{\sigma}, \mathcal{S}\}$ as

$$\{\hat{\boldsymbol{\sigma}}, \hat{\mathcal{S}}\} = \arg\max_{\boldsymbol{\sigma},\mathcal{S}} f(\mathcal{S}, \boldsymbol{\sigma}|\boldsymbol{y}) \tag{11}$$

$$= \arg\max_{\boldsymbol{\sigma},\mathcal{S}} f(\boldsymbol{y}|\mathcal{S}, \boldsymbol{\sigma}) f(\boldsymbol{\sigma}|\mathcal{S}) f(\mathcal{S}) \tag{12}$$

$$= \arg\max_{\boldsymbol{\sigma}(\mathcal{S}),\mathcal{S}} f(\boldsymbol{y}|\mathcal{S}, \boldsymbol{\sigma}(\mathcal{S})) f(\mathcal{S}). \tag{13}$$

where in the last step we made use of the fact that the *a priori* distribution $f(\boldsymbol{\sigma}|\mathcal{S})$ is modeled as non-informative for those active components of $\boldsymbol{\sigma}$ with the sparsity pattern imposed by $\mathcal{S}$ (denoted here as $\boldsymbol{\sigma}(\mathcal{S})$). That is, $\boldsymbol{\sigma}(\mathcal{S})$ has zeros at the positions specified by $\{0, \ldots, M\} - \mathcal{S}$, but no prior is assumed for the remaining components.

Note that (13) is a mixed discrete/continuous maximization problem, i.e., $\mathcal{S}$ can only take one out of 2^M values, whereas for fixed $\mathcal{S}$ the maximization is performed over the *continuous* parameter vector $\boldsymbol{\sigma}(\mathcal{S})$. Given $\mathcal{S}$, the problem reduces to one of Maximum Likelihood (ML) estimation:

$$\hat{\boldsymbol{\sigma}}_{\mathrm{ML}}(\mathcal{S}) = \arg\max_{\boldsymbol{\sigma}(\mathcal{S})} \ln f(\boldsymbol{y}|\boldsymbol{\sigma}(\mathcal{S}), \mathcal{S}). \tag{14}$$

where $f(\boldsymbol{y}|\boldsymbol{\sigma}, \mathcal{S}) = \exp\{-\boldsymbol{y}^H \tilde{\boldsymbol{R}}^{-1}(\boldsymbol{\sigma})\boldsymbol{y}\}/(\pi^K \det \tilde{\boldsymbol{R}}(\boldsymbol{\sigma}))$. This problem amounts to one of structured covariance matrix estimation from Gaussian observations [6], where the structure of $\tilde{\boldsymbol{R}}(\boldsymbol{\sigma})$ is defined by (4). This estimation problem has unfortunately no closed-form solution, though a fixed point iteration exists [6] that converges to the ML estimate $\hat{\boldsymbol{\sigma}}_{\mathrm{ML}}(\mathcal{S})$. Substituting the ML estimate $\hat{\boldsymbol{R}}_{\mathcal{S}} \doteq \tilde{\boldsymbol{R}}(\hat{\boldsymbol{\sigma}}_{ML}(\mathcal{S}))$ in (13) and disregarding constant additive terms, we define an equivalent log-likelihood function as

$$\mu(\mathcal{S}) \doteq \ln\{f(\boldsymbol{y}|\mathcal{S}, \hat{\boldsymbol{\sigma}}_{ML}(\mathcal{S})) f(\mathcal{S})\} - \text{constant terms}$$
$$= -\ln\det(\hat{\boldsymbol{R}}_{\mathcal{S}}) - \boldsymbol{y}^H \hat{\boldsymbol{R}}_{\mathcal{S}}^{-1}\boldsymbol{y} + |\mathcal{S}| \ln \frac{p_1}{1-p_1}. \tag{15}$$

In principle, the metric $\mu(\mathcal{S})$ must be maximized with respect to $\mathcal{S}$ by exhaustive search. This is impractical since the number of possible sets is 2^M. Instead we propose a Bayesian matching pursuit [7] algorithm that iteratively estimates the set of active channels, as described in Table 1. This suboptimal greedy solution finds the right set of active channels with high probability, as shown in the simulations section. The idea is to construct the active set estimate $\hat{\mathcal{S}}$ sequentially: starting with the "only noise" set $\hat{\mathcal{S}} = \{0\}$, at each step a new active channel is added to $\hat{\mathcal{S}}$ in order to maximize the corresponding metric $\mu(\hat{\mathcal{S}})$. This procedure is repeated until we find a set with $M_{\max}$ active channels, where $M_{\max}$ is user-selected. The final estimate $\hat{\mathcal{S}}$ is given by the partial solution $\hat{\mathcal{S}}_n$ with maximum posterior loglikelihood function, obtaining as byproduct the corresponding estimate $\hat{\boldsymbol{\sigma}}_{\mathrm{ML}}(\hat{\mathcal{S}})$.

5. NUMERICAL RESULTS

We analyze the performance of the proposed spectral estimator via Monte Carlo simulations, focusing on a scenario where the primary system is a TV broadcast network and the noise is assumed white. The compression matrix $\boldsymbol{\Phi}$ is a *random pinning* matrix, i. e. it is comprised by K randomly selected rows of the $N \times N$ identity matrix.

$\hat{\mathcal{S}}_0 = \{0\}$
for $n = 1$ to $M_{\max}$ do
  begin
    $m^\star = \arg\max_{m \notin \hat{\mathcal{S}}_{n-1}} \mu(\hat{\mathcal{S}}_{n-1} \cup \{m\})$
    $\hat{\mathcal{S}}_n = \hat{\mathcal{S}}_{n-1} \cup \{m^\star\}$
  end;
$\hat{\mathcal{S}} = \arg\max_{\hat{\mathcal{S}}_n} \mu(\hat{\mathcal{S}}_n)$

Table 1. Pseudocode for the proposed greedy approach.

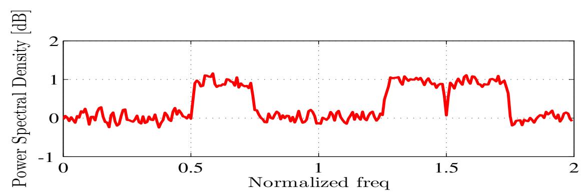

Fig. 1. One realization of a scenario with 8 DVB-T channels.

Synthetic scenario with 8 DVB-T channels. Channels are 8 MHz wide and can be occupied by DVB-T signal of bandwidth $B = 7.61$ MHz. Each of this channels is active with probability p_1 and, for simplicity in the presentation of results, they are all received with the same power when active. Fig. 1 shows one realization of this scenario.

Fig. 2 shows the behavior of the Normalized Mean Squared Error (NMSE) of the proposed scheme for $p_1 = 0.25$, $M_{max} = 4$, $N = 2048$, and $K = 512$. We sweep the power of active channels between -15 and 0 dB, while the noise power is kept constant to 0 dB. The NMSE of both the noise level and of one of the active channels is shown, together with the bounds presented in Section 3, averaged over the active channel set realizations. For low SNR values close to -15 dB the active channel power estimate becomes biased, which explains the apparent violation of the GACRLB. This bias can be reduced by increasing the observation time. It is interesting to note that (i) a gap exists between the CRLB and the corresponding GACRLB, which is particularly significant for the estimation of the noise variance, and (ii) the noise variance estimate performs close to the GACRLB over this SNR range. Providing the receiver with accurate noise variance estimates is of great importance to CR systems, in order to determine the correct threshold for the target probabilities of detection and false alarm [8]. Fig. 3 shows the variation of the NMSE with the compression ratio. For the same p_1 we fixed here the observation time to $N = 2^{10}$, and let K vary

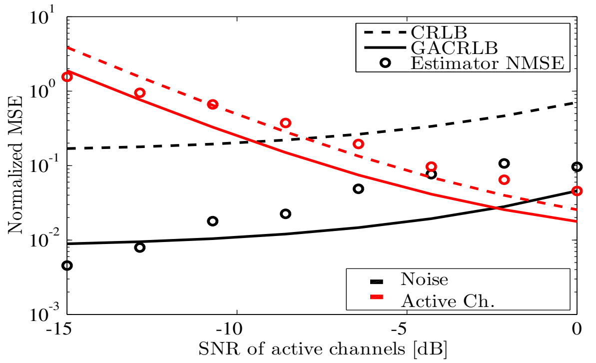

Fig. 2. NMSE for varying SNR. $N = 2048$, $K = 512$.

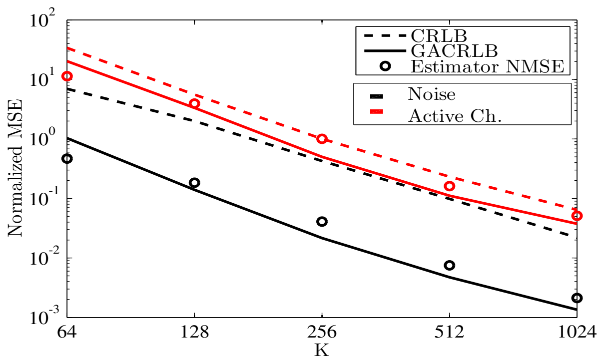

Fig. 3. NMSE for varying compression ratios. $N = 1024$, SNR $= -10$ dB.

from 2^6 to 2^{10}, for SNR $= -10$ dB. For $K \geq 128$ the noise variance estimate is unbiased with performance close to the GACRLB.

Captured TV band. In this (more realistic) scenario we captured part of the Spanish TV broadcast band (112 MHz bandwidth, comprising 14 channels with PAL/DVB-T signals). The *a priori* covariance matrices were generated using the channelization information of the PAL/DVB-T broadcast network, while the occupacy probability was considered $p_1 = 0.3$. The compression procedure has been simulated in Matlab using a 512×2048 random pinning matrix. No knowledge is fed to the reconstruction algorithm about the particular modulation (PAL or DVB-T) encountered at a given channel. Fig. 4 shows the psd of the band (obtained using a large number of uncompressed samples) together with the reconstruction obtained by the proposed method using just $K = 512$ samples. Even this reduced number of samples allows the estimation of 29 power levels needed for the reconstruction (14 DVB-T + 14 PAL + noise level).

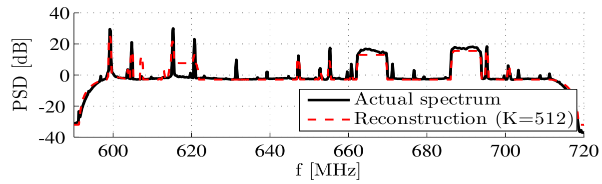

Fig. 4. Example of reconstruction of a mixed analog/digital broadcasting television band.

6. CONCLUSIONS

In CR applications, the observed signal is expected to be sparse in the "activity domain". Capitalizing on recent compressed sensing ideas, we proposed an iterative greedy approach which simultaneously estimates the set of active channels and their power levels. Significant improvement is observed in terms of noise variance estimation, an important parameter for detection threshold design. In practice, the complexity of the iteration described in [6], which is applied at the parameter estimation stage of the proposed method, can become too costly. Research is underway in order to devise alternative methods with lower computational cost.

7. REFERENCES

[1] Zhi Tian and G.B. Giannakis, "Compressed sensing for wideband cognitive radios," in *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, 2007.

[2] Y. L. Polo, Ying Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, 2009.

[3] S. M. Kay, *Fundamentals of statistical signal processing: detection theory*, Prentice-Hall, 1998.

[4] L.L. Scharf and L.T. McWhorter, "Geometry of the Cramer-Rao bound," in *IEEE Workshop Stat. Signal Array Process.,*, 1992.

[5] B. Babadi, N. Kalouptsidis, and V. Tarokh, "Asymptotic achievability of the Cramer-Rao bound for noisy compressive sampling," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 1233–1236, March 2009.

[6] J.P. Burg, D.G. Luenberger, and D.L. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, no. 9, pp. 963–974, Sept. 1982.

[7] P. Schniter, L. C. Potter, and J. Ziniel, "Fast bayesian matching pursuit," in *Inf. Theory and Applications Workshop*, 2008.

[8] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 4–17, Feb. 2008.