WIDEBAND SPECTRUM SENSING IN COGNITIVE RADIO: JOINT ESTIMATION OF NOISE VARIANCE AND MULTIPLE SIGNAL LEVELS

Roberto López-Valcarce and Gonzalo Vazquez-Vilar

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

ABSTRACT

The cognitive radio paradigm is based on the ability to detect the presence of primary users in a given frequency band. In this scenario a spectrum monitor may estimate the signal power levels of all frequency channels in the band of interest, together with the background noise level. We address Maximum Likelihood estimation for this problem, exploiting *a priori* knowledge about the primary network, summarized in the spectral shape of primary transmissions. An iterative asymptotic ML estimate is proposed, which can be further simplified in order obtain a computationally more efficient Least Squares estimator with performance very close to the Cramer-Rao lower bound in several cases of interest.

1. INTRODUCTION

Cognitive Radio is receiving considerable interest as a means for wireless systems to improve spectral usage. The key idea of opportunistically accessing temporally and/or spatially unused licensed bands requires powerful spectrum monitoring techniques, since the interference produced to licensed (primary) users must be kept at sufficiently low levels. The energy detector is a popular choice due to its simplicity, but it is not robust to noise-level uncertainty [1]; the induced sensitivity to threshold setting is a characteristic common to many detectors [2]. Due to this, accurate estimation of noise variance becomes an important task in spectrum sensing, which is hampered by the fact that it is unknown whether the signal is present or not at the time of the measurement, and with which level. Thus, joint estimation of noise and signal levels must be performed. This was done in [3] assuming a single primary subchannel is monitored at a time. However, when the whole bandwidth to monitor is large, sequential individual sensing of many (say N) primary channels poses significant challenges to the design of the receiver's analog front-end oscillators [4]. A means to alleviate this problem is to break the bandwidth of interest into subbands comprising $M < N$ primary channels each. These subband signals are then downconverted, digitized and analyzed sequentially. Note that the resolution and speed requirements for the analog-to-digital converter (ADC) become more stringent as M increases, so this strategy allows a tradeoff between oscillator and ADC complexity by judiciously choosing the ratio N/M. In addition, one expects that, since the bandwidth available to the noise variance estimators is now M times larger than in the single-channel case, estimation performance should improve.

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

In the setting considered in this paper it is assumed that several primary network parameters, such as channelization, modulation type, etc., are available to the spectrum monitor, and this information is translated into knowledge about the spectral shape of primary transmissions. This is reasonable in many practical cases, in particular for broadcast-type primary networks. If in addition the network uses multicarrier modulation, then a mathematically tractable Gaussian model can be adopted. Under this assumption we derive a joint estimator for multicarrier signals based on a Maximum Likelihood (ML) derivation that simultaneously estimates the noise power and the signal level present in each of the channels.

Section 2 presents the system model. ML estimation is addressed in Section 3, whereas a Least Squares (LS) estimator and the derivation of the Cramer-Rao Lower Bound (CRLB) are derived in Sections 4 and 5. Numerical results and final conclusions are given in Sections 6 and 7, respectively.

2. SYSTEM MODEL

Primary users employ Frequency Division Multiplexing with fixed channelization, known to the spectrum monitor. Several primary channels are sensed simultaneously, selecting a subband containing M such channels, downconverting the wideband signal to baseband, and sampling this baseband signal at f_s samples/s, thus obtaining K complex-valued samples:

$$r_k = \sum_{m=1}^{M} \sigma_m x_k^{(m)} + \sigma w_k, \qquad 0 \le k \le K-1, \quad (1)$$

where r_k are the observations; $x_k^{(m)}$ are the (noiseless) samples of the primary signal in channel m, normalized to unit variance ($E\{|x_k^{(m)}|^2\} = 1$); w_k are samples of a zero mean, circular complex white Gaussian noise with unit variance; σ_m^2 is the power of the primary signal in the mth channel ($\sigma_m^2 = 0$ if the mth channel is vacant); and $\sigma^2 > 0$ is the background noise power. Assuming that the primary users employ multicarrier modulation, the processes $x_k^{(m)}$ can be taken as circular Gaussian. This is reasonable for the number of subcarriers typically used in practical multicarrier systems.

For ease of notation we define $x_k^{(0)} = w_k$ and $\sigma_0^2 \doteq \sigma^2$, such that we can compactly write (1) in vector notation as

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

with $\boldsymbol{r} = [r_0\, r_1\, \cdots\, r_{K-1}]^T$ and $\boldsymbol{x}_m = [x_0^{(m)}\, x_1^{(m)}\, \cdots\, x_{K-1}^{(m)}]^T$. Hence, each $\boldsymbol{x}^{(m)}$ is circular Gaussian with zero mean and covariance matrix $\boldsymbol{C}_m \doteq E\{\boldsymbol{x}_m \boldsymbol{x}_m^H\}$, which is assumed known. This is reasonable if the channelization and the modulation parameters of the primary system are fixed and public, as would be the case for broadcasting networks[1]. We model the asynchronously sampled discrete-time processes as wide-sense stationary; thus, $\boldsymbol{C}_m$ is Toeplitz with ones on the diagonal. Note that $\boldsymbol{C}_0 = \boldsymbol{I}$ since the noise is assumed white.

The signals $x_k^{(m)}$ with $m = 1, \ldots, M$ are assumed statistically independent, since they correspond to different primary transmissions; and in addition they are statistically independent of the background noise. Hence, the observation $\boldsymbol{r}$ is zero-mean circular Gaussian with covariance

$$\boldsymbol{R} \doteq E\{\boldsymbol{r}\boldsymbol{r}^H\} = \sum_{m=0}^{M} \sigma_m^2 \boldsymbol{C}_m, \tag{3}$$

and the power spectral density (psd) of the observed signal is

$$\mathsf{S}_r(e^{j\omega}) = \sum_{m=0}^{M} \sigma_m^2 \mathsf{S}_m(e^{j\omega}) \tag{4}$$

where $\mathsf{S}_m(e^{j\omega})$ are the psd's of the constituent signals $x_k^{(m)}$ for $m = 0, 1, \ldots, M$. We address the estimation of the signal and noise powers σ_m^2, $0 \le m \le M$ from the observations $\boldsymbol{r}$.

3. MAXIMUM LIKELIHOOD ESTIMATION

Let $\boldsymbol{\sigma} \doteq [\sigma_0^2\ \sigma_1^2\ \cdots\ \sigma_M^2]^T$ be the vector of unknown parameters. The probability density function (pdf) of the observations conditioned on the unknown parameters is

$$f(\boldsymbol{r}|\boldsymbol{\sigma}) = \frac{1}{\pi^K \det \boldsymbol{R}} \exp\left\{-\boldsymbol{r}^H \boldsymbol{R}^{-1} \boldsymbol{r}\right\} \tag{5}$$

[1]Knowledge of the $\boldsymbol{C}_m$'s implicitly assumes that the channel is frequency-flat. However, the resulting estimators turn out to be robust to *unknown* frequency selectivity effects, as in [3].

and the ML estimate should maximize $f(\boldsymbol{r}|\boldsymbol{\sigma})$, or equivalently minimize $-\ln f(\boldsymbol{r}|\boldsymbol{\sigma})$. Thus, the problem reduces to the minimization of

$$L(\boldsymbol{r};\boldsymbol{\sigma}) \doteq \ln\det \boldsymbol{R} + \boldsymbol{r}^H \boldsymbol{R}^{-1} \boldsymbol{r} \tag{6}$$

The partial derivatives of $L(\boldsymbol{r};\boldsymbol{\sigma})$ w.r.t. σ_m^2 are

$$\frac{\partial L(\boldsymbol{r};\boldsymbol{\sigma})}{\partial \sigma_m^2} = -\operatorname{Tr}\{\boldsymbol{R}^{-1}\boldsymbol{C}_m\} + \boldsymbol{r}^H \boldsymbol{R}^{-1} \boldsymbol{C}_m \boldsymbol{R}^{-1} \boldsymbol{r} \tag{7}$$

The ML estimate of $\boldsymbol{\sigma}$ satisfies $\partial L/\partial\sigma_m^2 = 0$ for $0 \le m \le M$. This yields a set of nonlinear equations with no closed-form solution, and one must resort to numerical schemes. We propose a fixed-point iterative method for the computation of the ML estimate in the asymptotic regime $K \to \infty$.

3.1. Iterative solution for the asymptotic case

In view of (7), the natural approach to solving $\partial L/\partial\sigma_m^2 = 0$ seems to be the diagonalization of the matrices involved, as done in [3] for the case $M = 1$. However, when $M > 1$ channels are present, the covariance matrices $\boldsymbol{C}_m$ do not share common eigenvectors, and the eigenstructure of $\boldsymbol{R}$ in (3) is unclear. However, as $K \to \infty$, the eigenvalues of $\boldsymbol{C}_m$ are well approximated by the regularly spaced samples of $\mathsf{S}_m(e^{j\omega})$, whereas the eigenvectors approach the columns of the $K \times K$ orthonormal IDFT matrix $\boldsymbol{W}$ [5], i.e.,

$$\boldsymbol{C}_m \approx \boldsymbol{W}\boldsymbol{\Lambda}_m\boldsymbol{W}^H, \qquad m = 0, 1, \ldots, M \tag{8}$$

where $\boldsymbol{\Lambda}_m \doteq \operatorname{diag}(\boldsymbol{\lambda}_m)$, with $\boldsymbol{\lambda}_m \doteq [\lambda_0^{(m)}\ \lambda_1^{(m)}\ \cdots\ \lambda_{K-1}^{(m)}]^T$, and

$$\lambda_k^{(m)} \doteq \mathsf{S}_m(e^{j\frac{2\pi k}{K}}), \quad 0 \le k \le K-1. \tag{9}$$

From (3) and (8) it follows that

$$\boldsymbol{R} \approx \boldsymbol{W}\boldsymbol{\Delta}(\boldsymbol{\sigma})\boldsymbol{W}^H \quad \text{as } K \to \infty, \tag{10}$$

$$\text{with} \quad \boldsymbol{\Delta}(\boldsymbol{\sigma}) \doteq \sum_{m=0}^{M} \sigma_m^2 \boldsymbol{\Lambda}_m. \tag{11}$$

Substituting (10) back into (7) we obtain

$$\frac{\partial L(\boldsymbol{r};\boldsymbol{\sigma})}{\partial \sigma_m^2} \approx -\operatorname{Tr}\left\{\boldsymbol{\Delta}^{-1}(\boldsymbol{\sigma})\boldsymbol{\Lambda}_m\right\} + \boldsymbol{v}^H \boldsymbol{\Delta}^{-1}(\boldsymbol{\sigma})\boldsymbol{\Lambda}_m\boldsymbol{\Delta}^{-1}(\boldsymbol{\sigma})\boldsymbol{v}, \tag{12}$$

where $\boldsymbol{v} \doteq \boldsymbol{W}^H\boldsymbol{r}$ is the DFT of the observations. Using the definitions of $\boldsymbol{\Lambda}_m$ and $\boldsymbol{\Delta}(\boldsymbol{\sigma})$, we can rewrite (12) as

$$\frac{\partial L}{\partial \sigma_m^2} \approx -\sum_{k=0}^{K-1} \frac{\lambda_k^{(m)}}{\sum_{i=0}^{M} \sigma_i^2 \lambda_k^{(i)}} + \sum_{k=0}^{K-1} \frac{|v_k|^2 \lambda_k^{(m)}}{\left(\sum_{i=0}^{M} \sigma_i^2 \lambda_k^{(i)}\right)^2}. \tag{13}$$

Let us define

$$\Delta_k(\boldsymbol{\sigma}) \doteq \sum_{i=0}^{M} \sigma_i^2 \lambda_k^{(i)}, \qquad 0 \le k \le K-1, \tag{14}$$

so that

$$\mathbf{\Delta}(\boldsymbol{\sigma}) = \mathrm{diag}\{\ \Delta_0(\boldsymbol{\sigma}) \quad \Delta_1(\boldsymbol{\sigma}) \quad \cdots \quad \Delta_{K-1}(\boldsymbol{\sigma})\ \}. \quad (15)$$

Then, equating (13) to zero, we find that the ML estimate $\hat{\boldsymbol{\sigma}}_{\mathrm{ML}}$ must satisfy

$$\sum_{k=0}^{K-1} \frac{\lambda_k^{(m)}}{\Delta_k(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}})} = \sum_{k=0}^{K-1} \frac{|v_k|^2 \lambda_k^{(m)}}{\Delta_k^2(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}})}, \quad 0 \le m \le M, \quad (16)$$

as $K \to \infty$. The left hand side of (16) can be rewritten as

$$\sum_{k=0}^{K-1} \frac{\lambda_k^{(m)}}{\Delta_k} = \sum_{k=0}^{K-1} \frac{\Delta_k \lambda_k^{(m)}}{\Delta_k^2} \quad (17)$$

$$= \sum_{i=0}^{M} \sigma_i^2 \left(\sum_{k=0}^{K-1} \frac{\lambda_k^{(i)} \lambda_k^{(m)}}{\Delta_k^2} \right). \quad (18)$$

Substituting (18) into (16) and writing the result in matrix-vector form,

$$\boldsymbol{B}(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}})\hat{\boldsymbol{\sigma}}_{\mathrm{ML}} = \boldsymbol{b}(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}}), \quad (19)$$

where $\boldsymbol{B}(\boldsymbol{\sigma})$ and $\boldsymbol{b}(\boldsymbol{\sigma})$ are defined elementwise as

$$[\boldsymbol{B}(\boldsymbol{\sigma})]_{ij} \doteq \sum_{k=0}^{K-1} \frac{\lambda_k^{(i)} \lambda_k^{(j)}}{\Delta_k^2(\boldsymbol{\sigma})}, \quad 0 \le i, j \le M, \quad (20)$$

$$[\boldsymbol{b}(\boldsymbol{\sigma})]_{i} \doteq \sum_{k=0}^{K-1} \frac{|v_k|^2 \lambda_k^{(i)}}{\Delta_k^2(\boldsymbol{\sigma})}, \quad 0 \le i \le M. \quad (21)$$

Eqn. (19) suggests the following fixed point algorithm to obtain the asymptotic ML estimate: starting with some initial guess $\hat{\boldsymbol{\sigma}}_1$, compute

$$\hat{\boldsymbol{\sigma}}_{n+1} = \boldsymbol{B}^{-1}(\hat{\boldsymbol{\sigma}}_n)\boldsymbol{b}(\hat{\boldsymbol{\sigma}}_n), \qquad n = 1, 2, \ldots \quad (22)$$

This requires an initial K-point FFT, and solving a linear system of $(M+1)$ equations at each iteration. Fast convergence has been observed within a few iterations (<10) in all cases tested. In practice, $M \ll K$ and thus the computational complexity of this method is $O(K \log_2 K)$.

Remark: If we define the $K \times (M+1)$ matrix $\boldsymbol{L}$ and the $K \times 1$ vector $\boldsymbol{p}$ (the periodogram) respectively as

$$\boldsymbol{L} \doteq [\ \boldsymbol{\lambda}_0 \quad \boldsymbol{\lambda}_1 \quad \cdots \quad \boldsymbol{\lambda}_M\ ], \quad (23)$$

$$\boldsymbol{p} \doteq [\ |v_0|^2 \quad |v_1|^2 \quad \cdots \quad |v_{K-1}|^2\ ]^T, \quad (24)$$

then we can write (20)-(21) compactly as

$$\boldsymbol{B}(\boldsymbol{\sigma}) = \boldsymbol{L}^H \mathbf{\Delta}^{-2}(\boldsymbol{\sigma})\boldsymbol{L}, \quad (25)$$

$$\boldsymbol{b}(\boldsymbol{\sigma}) = \boldsymbol{L}^H \mathbf{\Delta}^{-2}(\boldsymbol{\sigma})\boldsymbol{p}. \quad (26)$$

Iteration (22) is well defined provided $\boldsymbol{B}$ remains non-singular along the iterations. Note from (25) that $\boldsymbol{B}$ is the Gram matrix of $\mathbf{\Delta}^{-1}\boldsymbol{L}$. Assuming that the diagonal matrix $\mathbf{\Delta}$ stays invertible, the invertibility of $\boldsymbol{B}(\boldsymbol{\sigma})$ amounts to linear independence of the columns of $\boldsymbol{L}$, i.e. the vectors $\boldsymbol{\lambda}^{(m)}$, $0 \le m \le M$. This condition, which can be checked *a priori*, states that the constituent reference psd's $\mathsf{S}_m(e^{j\omega})$, $0 \le m \le M$, must be linearly independent. This is intuitively satisfying: since the observed process is a mixture of circular Gaussian processes, the only means available to estimate the relative powers is by exploiting spectral diversity. When this linear independence condition is violated, the parameter vector $\boldsymbol{\sigma}$ is not identifiable.

4. LEAST SQUARES ESTIMATION

Substituting (25)-(26) back into (19), one obtains

$$\boldsymbol{L}^H \mathbf{\Delta}^{-2}(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}})\boldsymbol{L}\hat{\boldsymbol{\sigma}}_{\mathrm{ML}} = \boldsymbol{L}^H \mathbf{\Delta}^{-2}(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}})\boldsymbol{p}, \quad (27)$$

revealing certain similarity between the left and right terms, which can be exploited in order to obtain a simplified estimator as follows. First, note that the periodogram $\boldsymbol{p}$ is an asymptotically unbiased estimate of the received psd [6], and therefore $\boldsymbol{p}_\star \doteq \lim_{K\to\infty} E\{\boldsymbol{p}\} = \boldsymbol{L}\boldsymbol{\sigma}_\star$ with $\boldsymbol{\sigma}_\star$ the vector of true parameters. Thus, asymptotically, the expected value of $\boldsymbol{p}$ lies in the subspace spanned by the columns of $\boldsymbol{L}$. By assumption, $\boldsymbol{L}$ has full column rank, and $\boldsymbol{L}^\dagger \boldsymbol{L} = \boldsymbol{I}_{M+1}$, with $\boldsymbol{L}^\dagger$ denoting the pseudoinverse of $\boldsymbol{L}$. Then it holds that

$$\boldsymbol{L}\boldsymbol{L}^\dagger \boldsymbol{p}_\star = \boldsymbol{L}\boldsymbol{L}^\dagger \boldsymbol{L}\boldsymbol{\sigma}_\star = \boldsymbol{L}\boldsymbol{\sigma}_\star = \boldsymbol{p}_\star, \quad (28)$$

which suggests the approximation $\boldsymbol{p} \approx \boldsymbol{L}\boldsymbol{L}^\dagger \boldsymbol{p}$ (asymptotically exact in expected value). Substituting this in (27),

$$\hat{\boldsymbol{\sigma}}_{\mathrm{ML}} \approx \left[\boldsymbol{L}^H \mathbf{\Delta}^{-2}(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}})\boldsymbol{L}\right]^{-1} \boldsymbol{L}^H \mathbf{\Delta}^{-2}(\hat{\boldsymbol{\sigma}}_{\mathrm{ML}})\boldsymbol{L}\boldsymbol{L}^\dagger \boldsymbol{p}$$

$$= \boldsymbol{L}^\dagger \boldsymbol{p} \doteq \hat{\boldsymbol{\sigma}}_{\mathrm{LS}} \quad (29)$$

The subscript LS refers to the fact that this estimate is the solution to the Least Squares problem $\min_{\hat{\boldsymbol{\sigma}}} \|\boldsymbol{L}\hat{\boldsymbol{\sigma}} - \boldsymbol{p}\|^2$. The appeal of $\hat{\boldsymbol{\sigma}}_{\mathrm{LS}}$ resides in its one-shot nature, as opposed to the iterative scheme (22) for the computation of the ML estimate. Note that the iterations required by the ML method may pose a problem in practical implementations in which the subband to be analyzed contains a large number of channels.

5. CRAMER-RAO LOWER BOUND

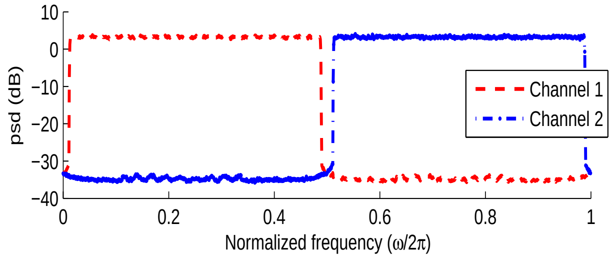
Since the observations are (complex valued) circular Gaussian with zero mean and covariance $\boldsymbol{R}(\boldsymbol{\sigma})$, the elements of the Fisher information matrix $\boldsymbol{F}(\boldsymbol{\sigma})$ are given by (see e.g. [7]):

$$[\boldsymbol{F}(\boldsymbol{\sigma})]_{ij} = \mathrm{Tr}\left\{ \boldsymbol{R}^{-1}(\boldsymbol{\sigma}) \frac{\partial \boldsymbol{R}(\boldsymbol{\sigma})}{\partial \sigma_i^2} \boldsymbol{R}^{-1}(\boldsymbol{\sigma}) \frac{\partial \boldsymbol{R}(\boldsymbol{\sigma})}{\partial \sigma_j^2} \right\}. \quad (30)$$

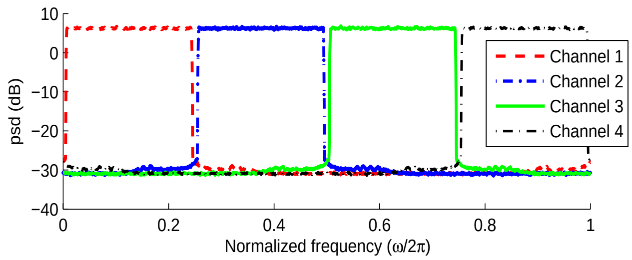
In our case, $\partial \boldsymbol{R}(\boldsymbol{\sigma})/\partial \sigma_m^2 = \boldsymbol{C}_m$. Using the asymptotic approximations (8) and (10),

$$[\boldsymbol{F}(\boldsymbol{\sigma})]_{ij} \approx \mathrm{Tr}\left\{ \mathbf{\Delta}^{-1}(\boldsymbol{\sigma})\mathbf{\Lambda}_i \mathbf{\Delta}^{-1}(\boldsymbol{\sigma})\mathbf{\Lambda}_j \right\} \quad (31)$$

$$= \sum_{k=0}^{K-1} \frac{\lambda_k^{(i)} \lambda_k^{(j)}}{\Delta_k^2(\boldsymbol{\sigma})} = [\boldsymbol{B}(\boldsymbol{\sigma})]_{ij} \quad (32)$$

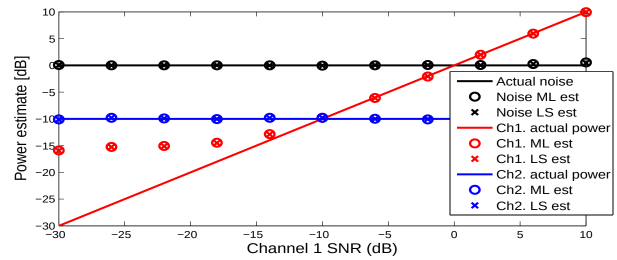


(a) Scenario 1: $M = 2$ channels.

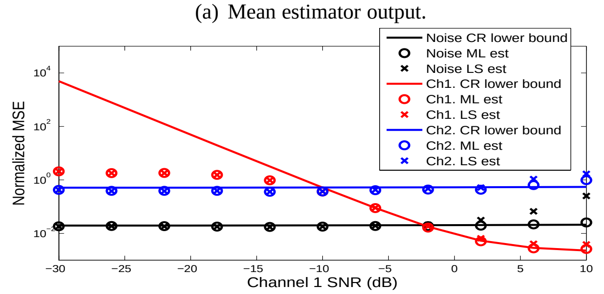


(b) Scenario 2: $M = 4$ channels.

Fig. 1. PSD of the 8 MHz DVB-T reference signals.



(a) Mean estimator output.



(b) Normalized MSE.

Fig. 2. Estimator performance in Scenario 1 ($M = 2$ channels). $K = 1024$ samples, $\text{SNR}_2 = -10$ dB.

Thus, the Fisher information matrix converges to $\boldsymbol{B}(\boldsymbol{\sigma})$ asymptotically. The CRLB is given by

$$\text{var}(\hat{\sigma}_i^2) \geq [\boldsymbol{F}^{-1}(\boldsymbol{\sigma})]_{ii} \overset{K\to\infty}{\longrightarrow} [\boldsymbol{B}^{-1}(\boldsymbol{\sigma})]_{ii}. \qquad (33)$$

6. SIMULATION RESULTS

The performance of the proposed estimation schemes is tested via Monte Carlo simulations, in which the primary system is a DVB-T broadcast network with 8 MHz channel spacing. Each DVB-T signal[2] has bandwidth $B = 7.61$ MHz and was quantized to 9-bit precision.

Two different scenarios are presented. In the first, $M = 2$ channels (centered at ± 4 MHz) are sampled at $f_s = 16$ MHz. This multichannel setting is the simplest possible, and illustrates the mutual effect of adjacent channels with possibly disparate powers. In the second scenario the number of channels is $M = 4$, with center frequencies at ± 4 and ± 12 MHz, and the sampling frequency is $f_s = 32$ MHz.

The psd's[3] of the reference signals for both scenarios are shown in Fig. 1. This psd estimates are used to approximate the actual psd's $\mathsf{S}_m(e^{j\omega})$. In the following, the SNR at channel m is defined as $\text{SNR}_m \doteq \sigma_m^2/\sigma_0^2$.

6.1. Scenario 1: $M = 2$ channels

The SNR of channel 1 is swept between -30 and 10 dB, while the SNR of channel 2 remains fixed. These low SNR values are expected to be typical operation regions for a Cognitive Radio spectrum monitor, which must detect primary activity even under strong fading conditions. Both the ML[4] and LS estimators were applied on $K = 1024$ data samples.

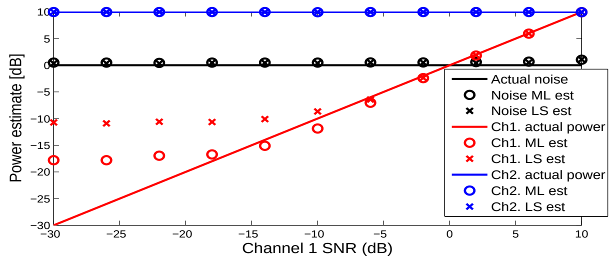
Fig. 2 shows the results obtained when $\text{SNR}_2 = -10$ dB. Regarding noise power estimation, both the ML and LS approaches are unbiased and perform close to the CRLB in the range shown, although the LS estimator starts to deviate for $\text{SNR}_1 > 0$ dB. Similar statements can be made about the estimates of the signal power in channel 2. On the other hand, regarding the estimation of σ_1^2, both methods become severely biased for $\text{SNR}_1 < -15$ dB. Above this threshold value, both are unbiased and achieve the CRLB.

Fig. 3 shows the results obtained when $\text{SNR}_2 = 10$ dB. Now we can see noticeable differences between both methods. The ML estimate offers a performance very similar to that in the previous case ($\text{SNR}_2 = -10$ dB). The LS estimate of σ_1^2 becomes now biased for $\text{SNR}_1 < 10$ dB, and even above this new threshold value, it presents a gap to the CRLB. Similarly, although the LS estimate of the noise power remains unbiased in the range shown, it also suffers from a gap to the corresponding CRLB.

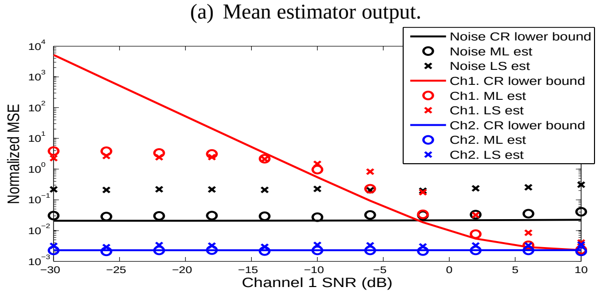
6.2. Scenario 2: $M = 4$ channels

Here we study the effect of sample size K on the mean value of the estimates. The SNRs of channels 1 through 4 were fixed at 10, 5 -5 and -10 dB respectively, and then K was swept between 256 and 4096. The results are shown in Fig. 4.

[2] 8K mode, 64-QAM, guard interval 1/4, inner code rate 2/3.
[3] via Welch's method (2^{10}-point Hamming windowing and 50% overlap).
[4] implemented as in (22) initialized at $\hat{\sigma}_i^2 = 10^{-2}$, $i = 0, 1, 2$, and stopping at iteration 10.



(a) Mean estimator output.



(b) Normalized MSE.

Fig. 3. Estimator performance in Scenario 1 ($M = 2$ channels). $K = 1024$ samples, $\text{SNR}_2 = 10$ dB.

For small K both methods present a strong bias in the estimation of the noise power, as well as in the estimation of the signal power in those channels with low SNR. This is not surprising, since the estimators developed in the previous sections rely on asymptotic results. To reduce the bias in this regime, it becomes necessary to increase the sample size K. It is also seen that the bias of the LS scheme goes to zero more slowly as K increases than that of the ML estimate. Note that the reduced space between contiguous DVB-T signals, as seen in Fig. 1(b), poses a significant challenge to noise variance estimators, and thus a relatively large value of K is required to estimate the noise level correctly.

7. CONCLUSIONS

We have considered Maximum Likelihood estimation of the noise and signal power levels from samples of a wideband signal comprising multiple multicarrier channels. Although no closed-form solution is available for the ML estimate in the finite data case, a recursive method was developed based on asymptotic approximations. The resulting iterative scheme can be further simplified in order to obtain a one-shot Least Squares estimator.

For both methods, signal level estimates suffer from a bias in those channels with low SNR. This bias is exacerbated for the LS estimator when a strong adjacent channel is present, and can be reduced by increasing the number of samples to process. Regarding noise variance estimation, strong channels also seem to affect the LS method more adversely. This would favor the use of the recursive ML scheme in Cognitive Radio applications, in which an accurate estimate of the background noise level is required for adequate detector performance [1].

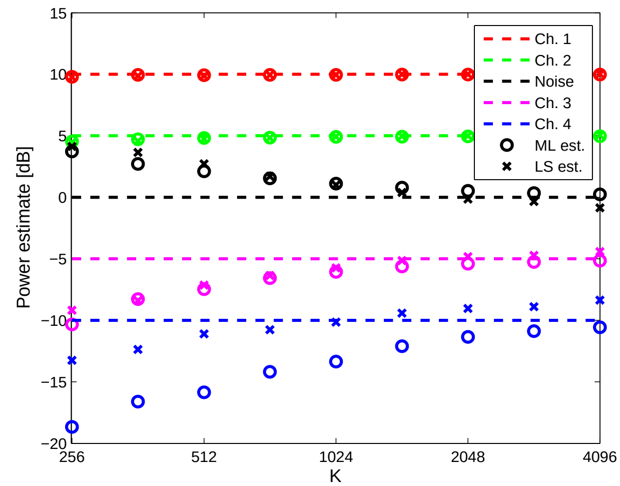


Fig. 4. Mean Estimator Output in Scenario 2 ($M = 4$ channels).

8. REFERENCES

[1] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Top. Signal Process.*, vol. 2, pp. 4–17, Feb 2008.

[2] A. Sonnenschein and P. M. Fishman, "Radiometric detection of spread-spectrum signals in noise of uncertain power," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, pp. 654–660, July 1992.

[3] R. López-Valcarce and C. Mosquera, "Maximum Likelihood SNR estimation for asynchronously oversampled OFDM signals," *Proc. 9th IEEE Workshop on Signal Proc. Adv. in Wireless Commun.*, pp. 26–30, July 2008.

[4] H. Zamat and B. Natarajan, "Use of dedicated broadband sensing receiver in cognitive radio," *Proc. 9th IEEE Conf. Commun. Workshops*, pp. 508–512, May 2008.

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

[6] P. Stoica and R. L. Moses, *Spectral analysis of signals*, Prentice-Hall, Englewood Cliffs, NJ, 2005.

[7] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Prentice-Hall, Englewood Cliffs, NJ, 1993.