COMPRESSIVE WIDEBAND SPECTRUM SENSING WITH SPECTRAL PRIOR INFORMATION

Daniel Romero[*†] *Roberto López-Valcarce*[*] *Geert Leus*[†]

[*]University of Vigo
Spain

[†]Delft University of Technology
The Netherlands

ABSTRACT

Wideband spectrum sensing provides a means to determine the occupancy of channels spanning a broad range of frequencies. Practical limitations impose that the acquisition should be accomplished at a low rate, much below the Nyquist lower bound. Dramatic rate reductions can be obtained by the observation that only a few parameters need to be estimated in typical spectrum sensing applications. This paper discusses the joint estimation of the power of a number of channels, whose power spectral density (PSD) is known up to a scale factor, using compressive measurements. First, relying on a Gaussian assumption, an efficient approximate maximum likelihood (ML) technique is presented. Next, a least-squares estimator is applied for the general non-Gaussian case.

Index Terms— Analog-to-Information Conversion, Spectrum Sensing, Maximum Likelihood Estimation.

1. INTRODUCTION

Spectrum sensing [1] refers to a collection of statistical inference procedures intended to determine the occupancy of a particular communication channel. This is of critical importance in certain applications such as those employing dynamic spectrum access [2]. When the bandwidth under analysis is large, limitations on the analog-to-digital converters (ADCs) together with computational issues impose restrictions on the techniques implemented thus leading to the concept of *wideband spectrum sensing*. Previous works include [3], where the goal is to optimize throughput subject to an interference constraint, but the setting is sensitive to the well-known noise uncertainty problem [4]. In [5], the noise power is not assumed known, but a certain number of channels need to be idle so that it can be easily estimated. A scheme that assumes neither noise power knowledge nor free channels was proposed in [6], but Nyquist sampling is required.

There has been a great interest in designing systems capable of acquiring frequency-sparse signals at a minimum rate. Works of this kind include, for example, [7] and [8]. With the boom of compressed sensing (CS) [9], this research line has been intensified (see e.g. [10]). Although most of these works deal with perfect reconstruction, this is not needed for spectrum sensing since typically only the second-order statistics are of interest. This observation has prompted recent schemes showing that a considerable reduction in the sampling rate is possible, even without the need for assuming sparsity [11,12].

This paper is pointed to exploit a further reduction in the sampling rate arising when certain prior information is used. Spectral masks and channelization parameters of the primary signals, such as central frequencies and bandwidths, are publicly available in many practical situations. Consequently, it is reasonable to assume that the received signal is a mixture of statistically independent signals, each one representing a channel received from a potentially different transmitter and whose power spectral density (PSD) is known up to a scaling factor related to the power of that transmission. Two Nyquist-rate sensing schemes exploiting this kind of prior information can be found in [6, 13]. We consider estimating these factors using the sub-Nyquist measurements provided by an analog-to-information converter (AIC) [14], whose operation principle can be thought of as projecting the analog signal onto a set of discrete sequences called the measurement vectors.

After formalizing the problem in Sec. 2, we review maximum likelihood (ML) estimation for Gaussian signals in Sec. 3, which is related to the classical problem referred to as covariance matching [15] or structured covariance estimation [16]. The numerical complexity required for the exact computation of the ML estimate is extremely high in moderately high dimensional settings, thus motivating the simple approximation presented in Sec. 3.1, which achieves a similar performance at a much more reasonable cost. This approximation also suggests a method to estimate the parameters when the Gaussian hypothesis does not hold true, leading to the weighted least-squares (WLS) estimator of Sec. 4. The estimation performance is then examined theoretically in Sec. 5 and by means of simulations in Sec. 6.

Two final remarks: first, although we do not consider signal detection, the proposed estimators may be used for decision making, e.g. by using the generalized likelihood ratio test (GLRT) [17]. Second, this paper focuses on algorithms, whereas a more fundamental analysis discussing minimum rates is the subject of [18] and subsequent publications.

Work supported by the European Regional Development Fund (ERDF) and the Spanish Government (TEC2010-21245-C02-02/TCM DYNACS, CONSOLIDER-INGENIO 2010 CSD2008-00010 COMONSENS), FPU Grant AP2010-0149, and the Galician Regional Government (CN 2012/260 AtlantTIC).

2. OBSERVATION MODEL

Suppose that a spectrum sensor receives a wideband signal $x(t)$ that is a linear combination of I independent zero-mean wide-sense stationary signals $x_i(t),\ i = 0, 1, \ldots, I-1$, some of them possibly representing noise or interference. The signal $x(t)$ can thus be written as $x(t) = \sum_i \sigma_i x_i(t)$, where the non-negative real coefficients σ_i are to be estimated. The conversion to the digital domain is carried out by an AIC that produces M sampling sequences $y_m[k],\ m = 0, 1, \ldots, M-1$ at the output by some kind of linear manipulation of $x(t)$. Although it may not have any physical existence, it is convenient to express these observations in terms of the Nyquist-sampled version of the received signal, $x[n] = x(nT)$, where T represents the reciprocal of the maximum frequency present in $x(t)$. Extending this idea to the components $x_i(t)$ we obtain $x[n] = \sum_i \sigma_i x_i[n]$.

The input samples $x[n]$ are arranged as N-tuples, each one giving rise to one sample at every output $y_m[k]$. Specifically, this relation is given by $y_m[k] = \boldsymbol{\phi}_m^H \boldsymbol{x}[k]$, where $\boldsymbol{\phi}_m \in \mathbb{C}^N$ and $\boldsymbol{x}[k] = [x[kN], x[kN+1], \ldots x[kN + (N-1)]]^T$. By stacking these M outputs in the vector $\boldsymbol{y}[k] = [y_0[k], y_1[k] \ldots, y_{M-1}[k]]^T$ we obtain the more compact form $\boldsymbol{y}[k] = \boldsymbol{\Phi}\boldsymbol{x}[k]$, where $\boldsymbol{\Phi} = [\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{M-1}]^H$ is referred to as the measurement matrix. Further, if we arrange all observations together, we can form the vector $\boldsymbol{y} = [\boldsymbol{y}^T[0], \boldsymbol{y}^T[1], \ldots, \boldsymbol{y}^T[K/N-1]]^T$ and write $\boldsymbol{y} = \bar{\boldsymbol{\Phi}}\boldsymbol{x}$, where KT is the acquisition time[1], $\bar{\boldsymbol{\Phi}} = \boldsymbol{I}_{K/N} \otimes \boldsymbol{\Phi}$ and $\boldsymbol{x} = [\boldsymbol{x}^T[0], \boldsymbol{x}^T[1], \ldots, \boldsymbol{x}^T[K/N-1]]^T$.

The second-order information of the signals $x_i[n]$ is collected in the autocorrelation sequence $r_i[n] = \mathrm{E}\{x_i[\nu+n]\, x_i^*[\nu]\}$, which is assumed known and normalized such that $r_i[0] = 1$). The Fourier transform of this sequence, when it exists, is called the PSD of the process $x_i[n]$. For commodity, let us arrange the coefficients in $r_i[n]$ as the elements of the Hermitian Toeplitz correlation (or covariance) matrix $\boldsymbol{\Sigma}_i = \mathrm{E}\left\{\boldsymbol{x}_i \boldsymbol{x}_i^H\right\}$, where $\boldsymbol{x}_i = [x_i[0], x_i[1], \ldots, x_i[K-1]]^T$. The set of matrices $\mathcal{S} = \{\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_{I-1}\}$ is assumed $\mathbb{R}$-linearly independent, in the sense that no two different linear combinations of these matrices using real coefficients can give the same matrix. Otherwise, the coefficients σ_i^2 are not identifiable [19]. The covariance matrix of the vector $\boldsymbol{x} = [x[0], x[1], \ldots, x[K-1]]^T$ containing the received signal, now written as $\boldsymbol{x} = \sum_i \sigma_i \boldsymbol{x}_i$, is given by $\boldsymbol{\Sigma} = \mathrm{E}\left\{\boldsymbol{x}\boldsymbol{x}^H\right\} = \sum_i \sigma_i^2 \boldsymbol{\Sigma}_i$. For convenience we also consider the decomposition of $\boldsymbol{\Sigma}$ into $N \times N$ blocks:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}[0] & \ldots & \boldsymbol{\Sigma}^H[\frac{K}{N}-1] \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}[\frac{K}{N}-1] & \ldots & \boldsymbol{\Sigma}[0] \end{bmatrix} \tag{1}$$

where $\boldsymbol{\Sigma}[k] = \mathrm{E}\left\{\boldsymbol{x}[\kappa+k]\boldsymbol{x}^H[\kappa]\right\}$. These blocks can be written as $\boldsymbol{\Sigma}[k] = \sum_i \sigma_i^2 \boldsymbol{\Sigma}_i[k]$, where $\boldsymbol{\Sigma}_i[k] = \mathrm{E}\{\boldsymbol{x}_i[\kappa+k]\, \boldsymbol{x}_i^H[\kappa]\}$ are the corresponding blocks in $\boldsymbol{\Sigma}_i$. The covariance matrix of $\boldsymbol{y}$ is clearly given by $\bar{\boldsymbol{\Sigma}} = \mathrm{E}\left\{\boldsymbol{y}\boldsymbol{y}^H\right\} = \bar{\boldsymbol{\Phi}}\boldsymbol{\Sigma}\bar{\boldsymbol{\Phi}}^H = \sum_i \sigma_i^2 \bar{\boldsymbol{\Sigma}}_i$, where $\bar{\boldsymbol{\Sigma}}_i = \bar{\boldsymbol{\Phi}}\boldsymbol{\Sigma}_i\bar{\boldsymbol{\Phi}}^H$. Thus, this transformation induces a new set of basis covariance matrices $\bar{\mathcal{S}} = \{\bar{\boldsymbol{\Sigma}}_0, \bar{\boldsymbol{\Sigma}}_1, \ldots, \bar{\boldsymbol{\Sigma}}_{I-1}\}$ whose blocks are given by $\bar{\boldsymbol{\Sigma}}[k] = \boldsymbol{\Phi}\boldsymbol{\Sigma}[k]\boldsymbol{\Phi}^H = \sum_i \sigma_i^2 \bar{\boldsymbol{\Sigma}}_i[k]$, where $\bar{\boldsymbol{\Sigma}}_i[k] = \boldsymbol{\Phi}\boldsymbol{\Sigma}_i[k]\boldsymbol{\Phi}^H$. Finally, note that although these blocks are not Toeplitz, the matrices $\bar{\boldsymbol{\Sigma}}$ and $\bar{\boldsymbol{\Sigma}}_i$ are block-wise Toeplitz, which means that the processes $y_i[k]$ are jointly stationary.

3. ESTIMATION FOR GAUSSIAN SIGNALS

The Gaussian assumption makes the statistical characterization of the observations $\boldsymbol{y}$ completely determined by the second-order statistics introduced in the previous section. The probability density function of the observations can thus be written as:

$$p\,(\boldsymbol{y}; \boldsymbol{\theta}) = \frac{\exp\left\{-\boldsymbol{y}^H \bar{\boldsymbol{\Sigma}}^{-1} \boldsymbol{y}\right\}}{\pi^{MK/N} |\bar{\boldsymbol{\Sigma}}|}, \tag{2}$$

where $\boldsymbol{\theta} = [\sigma_0^2, \sigma_1^2, \ldots, \sigma_{I-1}^2]^T$ is the vector of unknown parameters. Although $\bar{\boldsymbol{\Sigma}}$ depends on this vector, we dismiss the notation $\bar{\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ for clarity. The ML estimate of $\boldsymbol{\theta}$ given $\boldsymbol{y}$ is the maximizer of (2) seen as a likelihood function, that is, $\boldsymbol{\theta}_{ML} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{y}; \boldsymbol{\theta})$. This optimization problem has been widely analyzed and, up to now, no analytical solution is known [20].

One possible alternative is to search for a value of $\boldsymbol{\theta}$ that makes the gradient of the log-likelihood function equal to zero. Taking logarithms and disregarding constant terms in (2) yields

$$\mathcal{L}(\boldsymbol{\theta}) = \log |\bar{\boldsymbol{\Sigma}}| + \mathrm{Tr}\left(\bar{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{R}}\right), \tag{3}$$

where $\hat{\boldsymbol{R}} = \boldsymbol{y}\boldsymbol{y}^H$ is the sample covariance matrix. By noting that $\bar{\boldsymbol{\Sigma}} = \sum_i \sigma_i^2 \bar{\boldsymbol{\Sigma}}_i$, it is possible to write the partial derivatives as

$$\frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \sigma_i^2} = \mathrm{Tr}\left(\bar{\boldsymbol{\Sigma}}^{-1} \bar{\boldsymbol{\Sigma}}_i\right) - \mathrm{Tr}\left(\bar{\boldsymbol{\Sigma}}^{-1} \bar{\boldsymbol{\Sigma}}_i \bar{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{R}}\right). \tag{4}$$

Due to the regularity of this function, a maximum of $\mathcal{L}(\boldsymbol{\theta})$ is attained when the right hand side is zero for all i. After some algebraic manipulations, this condition becomes

$$\mathrm{Tr}\left((\bar{\boldsymbol{\Sigma}} - \hat{\boldsymbol{R}}) \bar{\boldsymbol{\Sigma}}^{-1} \bar{\boldsymbol{\Sigma}}_i \bar{\boldsymbol{\Sigma}}^{-1}\right) = 0 \quad \forall i. \tag{5}$$

Obtaining numerically the solution of this system of equations has been analyzed in [16], where an algorithm called the inverse iteration algorithm (IIA) is proposed. Unfortunately, the cost of this algorithm is considerably high since it is applicable to general non-stationary settings. Different approaches have been proposed in [21] and [22], but they only work for rank-one matrices.

[1] Throughout it will be assumed that K is an integer multiple of N.

3.1. Approximate ML solution: averaging and cropping

Note that the complexity of the IIA is a consequence of the relatively high dimension of the sample covariance matrix $\hat{\boldsymbol{R}}$. Although this is a sufficient statistic for the estimation of $\boldsymbol{\theta}$, it is clearly a non-consistent estimate of the true correlation matrix $\bar{\boldsymbol{\Sigma}}$. Thus, it makes sense to consider replacing $\hat{\boldsymbol{R}}$ in (2) by a consistent estimate of $\bar{\boldsymbol{\Sigma}}$ with a lower dimension. Although a formal motivation of this approach can be provided on the basis of the asymptotic theory of Toeplitz matrices, we omit such explanation in favor of a more heuristical one due to space limitations and leave the formalism to a subsequent publication.

Two observations apply. On the one hand, it is common practice in signal processing to estimate the autocorrelation coefficients by computing the traditional biased/unbiased estimates [23]. On the other hand, it is clear that not all the coefficients are estimated with equal *accuracy* since they rely on different numbers of samples. Both these remarks suggest estimating the correlation coefficients corresponding to lags $-L$ to L, for some choice of L satisfying $L+1 \leq K/N$, and compose the $M(L+1) \times M(L+1)$ sample covariance matrix $\hat{\boldsymbol{S}}$ whose k-th block is given by $\hat{\boldsymbol{S}}[k] = \frac{1}{K_k} \sum_\kappa \boldsymbol{y}[\kappa+k]\boldsymbol{y}^H[\kappa]$, where K_k is a constant depending on k that takes on the value $K_k = \frac{K}{N}$ for the biased estimate and $K_k = \frac{K}{N} - k$ for the unbiased estimate. This matrix can be thought of as an *averaged and cropped* version of the sample covariance matrix $\hat{\boldsymbol{R}}$, which presents the advantage of being consistent and having the same block structure as the true covariance matrix $\bar{\boldsymbol{\Sigma}}$. Note that $\hat{\boldsymbol{R}}$ lacks these two properties.

Repeating the process leading to (5) gives the same expression[2] with $\hat{\boldsymbol{S}}$ featuring in place of $\hat{\boldsymbol{R}}$. Expanding (5) for $\bar{\boldsymbol{\Sigma}}$ results in the following equivalent condition:

$$\sum_j \sigma_j^2 \operatorname{Tr}\left(\bar{\boldsymbol{\Sigma}}^{-1}\bar{\boldsymbol{\Sigma}}_i\bar{\boldsymbol{\Sigma}}^{-1}\bar{\boldsymbol{\Sigma}}_j\right) = \operatorname{Tr}\left(\bar{\boldsymbol{\Sigma}}^{-1}\bar{\boldsymbol{\Sigma}}_i\bar{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{S}}\right) \quad \forall i,$$

Note that this constitutes a linear system of equations in σ_j^2, with I unknowns and I equations. Following the same idea as the IIA, it is possible to iteratively attain a solution by replacing $\bar{\boldsymbol{\Sigma}}$ with that corresponding to the previous iterand. As an initialization, a good choice seems to be $\bar{\boldsymbol{\Sigma}} = \hat{\boldsymbol{S}}$. More details can be found in [16]. Although this algorithm, referred to as the simplified IIA (SIIA), does not achieve the exact ML solution, its complexity may be several orders of magnitude below the IIA depending on the L chosen.

4. ESTIMATION FOR GENERAL SIGNALS

When the signals are not Gaussian, applying the IIA or SIIA algorithms may make little sense. However, the averaging/cropping procedure proposed for the SIIA provides us with some guidelines to perform estimation in the general non-Gaussian case. Note that when the number of samples approaches infinity, $\hat{\boldsymbol{S}}$ converges in probability to $\bar{\boldsymbol{\Sigma}}$ (see e.g. [19]). In the absence of any other prior information about the received signal, but the second-order statistics, this remark suggests finding the vector $\boldsymbol{\theta}$ that minimizes some metric between $\hat{\boldsymbol{S}}$ and $\bar{\boldsymbol{\Sigma}}$, for example the Frobenius distance $||\hat{\boldsymbol{S}} - \bar{\boldsymbol{\Sigma}}||_F^2$. Since these matrices are block-Toeplitz, this can be simplified by taking just one representative of each block-diagonal and weighting it appropriately. If we define

$$\hat{\boldsymbol{s}} = \operatorname{vec}\begin{bmatrix} L_L\hat{\boldsymbol{S}}^H[L] \\ \vdots \\ L_1\hat{\boldsymbol{S}}^H[1] \\ L_0\hat{\boldsymbol{S}}[0] \\ L_1\hat{\boldsymbol{S}}[1] \\ \vdots \\ L_L\hat{\boldsymbol{S}}[L] \end{bmatrix} \quad \text{and} \quad \boldsymbol{v}_i = \operatorname{vec}\begin{bmatrix} L_L\bar{\boldsymbol{\Sigma}}_i^H[L] \\ \vdots \\ L_1\bar{\boldsymbol{\Sigma}}_i^H[1] \\ L_0\bar{\boldsymbol{\Sigma}}_i[0] \\ L_1\bar{\boldsymbol{\Sigma}}_i[1] \\ \vdots \\ L_L\bar{\boldsymbol{\Sigma}}_i[L] \end{bmatrix},$$

where $L_l = L + 1 - l$ accounts for the number of times the l-th block is present in $\hat{\boldsymbol{S}}$ or $\bar{\boldsymbol{\Sigma}}_i$, then the problem can be formulated as a LS program:

$$\boldsymbol{\theta}_{LS} = \arg\min_{\boldsymbol{\theta}} ||\hat{\boldsymbol{s}} - \boldsymbol{V}\boldsymbol{\theta}||^2 \tag{6}$$

where $\boldsymbol{V} = [\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{I-1}]$. The reason for considering the blocks $\hat{\boldsymbol{S}}[k]$ along with their Hermitian versions $\hat{\boldsymbol{S}}^H[k]$ is to naturally impose $\sigma_i^2 \in \mathbb{R}$. Another alternative would be to just consider the real and imaginary parts of the covariance matrix blocks and operate directly on $\mathbb{R}$.

Due to the scaling factors L_l, we may prefer to term this algorithm the WLS algorithm. In principle, since the coefficients σ_i^2 are non-negative, the solution must be found using constrained WLS (CWLS), although, for simplicity, one may also consider using unconstrained WLS. Note that what we actually do in the latter case is to project the sample correlation $\hat{\boldsymbol{S}}$ onto the space spanned by the I basis correlation matrices in $\bar{\mathcal{S}}$. Unfortunately, in this case, some of the estimated coefficients may be negative.

5. PERFORMANCE ANALYSIS

The asymptotic performance of the IIA is clearly determined for the cases where this algorithm converges to the global maximum of the likelihood function. In those cases, it provides the ML estimate, which is asymptotically unbiased and efficient [24]. By construction, the simplified IIA is expected to share the same asymptotic properties provided that L is high enough so that the truncated (cropped) covariance matrices in $\bar{\mathcal{S}}$ are still linearly independent. Unfortunately, evaluating the performance of these two algorithms for finite-length data records seems not tractable, so that we are forced to resort to Monte Carlo simulation in Sec. 6.

On the contrary, the performance of the unconstrained WLS algorithm may be analyzed for finite data sets. If L is chosen high enough such that the matrices in $\bar{\mathcal{S}}$ are linearly

[2] Of course the dimensions of the covariance matrices should be modified accordingly.

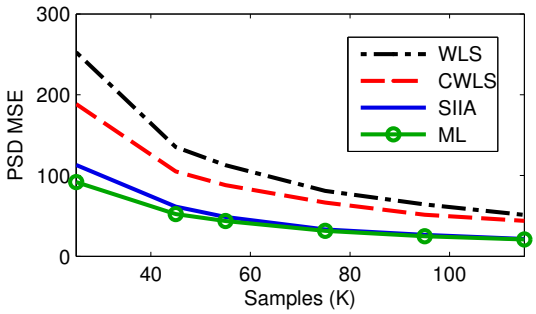

Fig. 1: Estimation error for an increasing number of samples when $N = 5$ and $M = 3$.

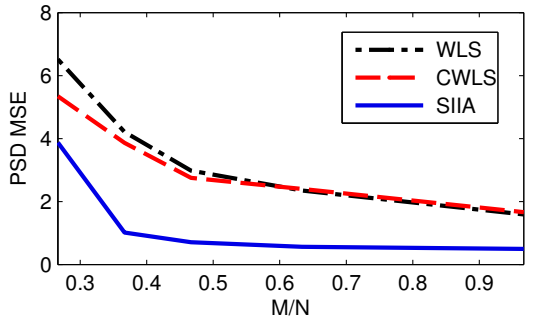

Fig. 2: Influence of the compression ratio when N is fixed to 30 and $K = 900$.

independent, then the columns of $\boldsymbol{V}$ in (6) are linearly independent and we can write $\boldsymbol{\theta}_{LS} = \boldsymbol{V}^{\dagger}\hat{\boldsymbol{s}}$, where the superscript † means pseudo-inverse. Thus, the estimate is linear with the sample covariance $\hat{\boldsymbol{S}}$. This means that $\boldsymbol{\theta}_{ML}$ is unbiased in case that $\hat{\boldsymbol{S}}$ is so and *vice versa*. Moreover, although it is out of the scope of the present paper, the second-order statistics of this estimate may be written in terms of the fourth-order statistics of $x[n]$. Finally note that the asymptotic performance of the CWLS estimate coincides with that for the unconstrained estimate whenever $\sigma_i^2 > 0\ \forall i$, since the positivity constraints will automatically hold in the limit.

6. SIMULATION RESULTS

This section provides a comparison by Monte Carlo (MC) simulation of the performance of the estimation schemes presented above. The first $I-1$ signals $x_i[n]$ are generated by passing white Gaussian noise of some specified power through a low-pass prototype FIR filter of order 30 and bandpass bandwidth $0.25/T$ for the first two figures and $0.05/T$ for the last one. Every signal is then frequency shifted to occupy a different band. The last signal $x_{I-1}[n]$ is simply white Gaussian noise. The value of L used is 4 in all cases.

The IIA algorithm was implemented by making use of several extensions suggested in [16] since the raw iteration is numerically quite unstable. The non-negativity constraint is

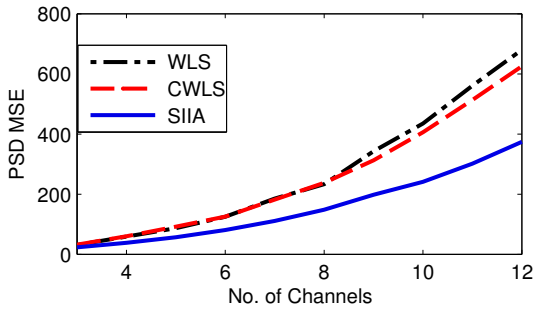

Fig. 3: Effect of the number of channels when $K = 500$, $N = 5$ and $M = 3$. All the signals have power 4 except for the noise, which has power 9.

imposed by dividing the stepsize by two whenever the new $\boldsymbol{\theta}$ becomes negative. This also avoids problems with the conditioning of the covariance matrix in the next iteration by keeping $\bar{\boldsymbol{\Sigma}}$ away from the boundary of the positive definite cone. This same updating rule was also used for SIIA.

To evaluate the performance, the components of $\boldsymbol{\theta}$ are estimated, the PSD is reconstructed and the Euclidean distance to the true one is computed to obtain the squared error. The mean squared error (MSE) follows by averaging this quantity over all realizations. The entries in the measurement matrices $\boldsymbol{\Phi}$ are drawn from independent Gaussian distributions with zero mean and unit standard deviation. For the sake of generality, a different matrix is generated at every MC iteration, thus making the results independent of any particular choice.

In Fig. 1, the MSE is represented *vs.* the number of samples K. We observe that all of them are effectively consistent and that the SIIA performs asymptotically like the IIA. Next, in Fig. 2, the effect of the compression ratio on the performance of the estimators is illustrated. Clearly, the higher the quotient M/N, the better the performance. ML is not shown in this figure because of the high computational time required, but it must be assumed to perform quite similarly to the SIIA. Note that $M/N = 1$ corresponds to the Nyquist rate whereas $M/N = 0.5$ corresponds to the Landau rate in absence of noise [25]. Finally, the influence of the number of channels I is shown in Fig. 3. As intuition predicts, the higher I, the higher the uncertainty and the higher the error.

7. CONCLUSIONS

Observing that a typical spectrum sensing application only requires the estimation of a few parameters, such as signal/noise power, an important reduction in the sampling rate may be accomplished by using compressive acquisition techniques. Several low-complexity estimation techniques have been applied to this setting under a multichannel model. Among other advantages, the proposed schemes are not sensitive to the well-known noise uncertainty problem. Future work is pointed to the analysis of the fundamental principles lying behind these techniques.

8. REFERENCES

[1] E. Axell, G. Leus, E.G. Larsson, and H.V. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101 –116, May 2012.

[2] Q. Zhao and B.M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, 2007.

[3] Z. Quan, S. Cui, A.H. Sayed, and H.V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128–1140, 2009.

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

[5] A. Taherpour, S. Gazor, and M. Nasiri-Kenari, "Wideband spectrum sensing in unknown white gaussian noise," *IET Communications*, vol. 2, no. 6, pp. 763–771, Jul. 2008.

[6] G. Vazquez-Vilar and R. Lopez-Valcarce, "Spectrum sensing exploiting guard bands and weak channels," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6045–6057, 2011.

[7] Ping Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *1996 IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 1996, vol. 3, pp. 1688 –1691 vol. 3.

[8] Yuan-Pei Lin and P.P. Vaidyanathan, "Periodically nonuniform sampling of bandpass signals," *IEEE Trans. Circuits and Syst. II: Analog Digit. Signal Process.*, vol. 45, no. 3, pp. 340 –351, Mar. 1998.

[9] D.L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289 –1306, Apr. 2006.

[10] M. Mishali and Y.C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993 – 1009, Mar. 2009.

[11] D.D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4775–4789, 2012.

[12] G. Leus and D.D. Ariananda, "Power spectrum blind sampling," *IEEE Signal Process. Lett.*, vol. 18, no. 8, pp. 443–446, 2011.

[13] A.I. Perez-Neira, Miguel-Angel Lagunas, M.A. Rojas, and P. Stoica, "Correlation matching approach for spectrum sensing in open spectrum communications," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4823–4836, Dec. 2009.

[14] J.N. Laska, S. Kirolos, M.F. Duarte, T.S. Ragheb, R.G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *2007 IEEE Int. Symp. Circuits Syst.*, May 2007, pp. 1959 –1962.

[15] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185–210, 1998.

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

[17] S.M. Kay, *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory*, Prentice-Hall, 1998.

[18] D. Romero and G. Leus, "Compressive Covariance Sampling," in *Inform. Theory Appl. Workshop*, Feb. 2013.

[19] E.L. Lehmann and G. Casella, *Theory of point estimation*, vol. 31, Springer, 1998.

[20] D.R. Fuhrmann, "Application of Toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Trans. Signal Process.*, vol. 39, no. 10, pp. 2194 –2198, Oct. 1991.

[21] P. Stoica and P. Babu, "Maximum-likelihood nonparametric estimation of smooth spectra from irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5746 –5758, Dec. 2011.

[22] P. Stoica, P. Babu, and Jian Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35 –47, Jan. 2011.

[23] P. Stoica and R.L. Moses, *Spectral analysis of signals*, Pearson/Prentice Hall, 2005.

[24] S.M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory*, Prentice-Hall, 1993.

[25] HJ Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, no. 1, pp. 37–52, 1967.