Sensing Wireless Microphone with ESPRIT from Noise and Adjacent Channel Interference

Dan Zhang, Lijun Dong and Narayan Mandayam
WINLAB, Rutgers University
671 US-1 South, North Brunswick, NJ 08902
Email: {bacholic, lijdong, narayan}@winlab.rutgers.edu

***Abstract*—FCC requires that any white space device be able to sense wireless microphone (WM) signals at -114dBm typically corresponding to a SNR of -20dB, an extremely challenging task. This paper presents a novel WM detector based on the ESPRIT algorithm along with an auxiliary prewhitening filter that meets this requirement. Compared with existing detectors, the proposed WM detector not only successfully combats noise, but also provides a technique to mitigate the effect of adjacent channel interference caused by DTV transmission leakage, an issue much less studied.**

I. Introduction

As required by FCC, any white space device (WSD) must be capable of detecting wireless microphone (WM) signals at -114dBm, which usually translates into -20dB signal-to-interference-noise ratio (SINR) in practice. This requirement poses a big challenge to the design of spectrum sensing algorithms because most WM devices employ analog frequency modulation (FM) without preambles or outstanding carriers typical to digital transmissions that can be recognized as a signature of the particular signal. At very low SNR, WM signals are not only smeared by noise but are also possibly obscured by adjacent channel interference due to DTV transmission leakage. The latter problem is much less studied. The best result to the authors' knowledge was presented in [1] where an SINR of -28dBm is required to achieve an error rate of 0.01. Rest of the current literature on WM detection mainly focuses on reliably detecting WM at the designated Rx power of -114dBm in the background of white noise. While some of the detectors manage to handle such low Rx power, they do not handle interference equally well. Readers may refer to [2], [3] for a literature review of these detectors.

In principle, WM detection is a binary hypothesis test with N observations of the received signal $\{y_n\}_{n=0}^{N-1}$ uniformly sampled at F_s:

$$\begin{cases} H_0: & y_n = \nu_n + i_n \\ H_1: & y_n = s_n + \nu_n + i_n \end{cases}, \quad n = 0, 1, \ldots, N-1, \quad (1)$$

where s_n represents the WM signal, ν_n represents the noise and i_n represents the interference. s_n, ν_n and i_n are assumed to have zero mean. The statistics of ν_n and i_n are unknown but we usually assume they are wide sense stationary for the duration of sensing. Current treatments of this problem can be summarized as follows.

- *Energy Detector* H_1 is claimed if $\sum_{i=1}^{N} |y_n|^2$ is above a threshold. This method is very unreliable at low SNR.
- *Spectral Detector* This method estimates the power spectral density (PSD) or the autocovariance function (ACF) of y_n. H_1 is claimed if a high peak on the PSD or resonance between the ACF and certain sinusoid is observed [1] [4]. A variant works with the cyclic spectrum [5] instead of PSD.
- *Eigenvalue Detector* This method tries to separate the signal subspace from the noise subspace by exploring the eigenvalue patterns of the covariance matrix [6]. A variant works with singular values. They provide high spectral resolution, but they do not address the interference issue.
- *Filter Bank Detector* A filter bank detector divides the spectrum into subbands, after which detection is carried out in each subband where SNR has been improved [7]. Depending on the type of the filter bank employed, this method is either equivalent to the spectral detector, or the wavelet detector.
- *Wavelet Detector* A proper choice of wavelets can possibly reveal the feature of the WM signal [8]. The major challenge is to design the best wavelet for this job. Since WM signal is sinusoid-like, it is naturally difficult to match it with any temporally localized wavelets.

In this paper we present a novel WM detector based on the ESPRIT algorithm [9] along with an auxiliary prewhitening filter that meets the stringent low SNR requirement and also mitigates adjacent channel interference.

II. System Model and Methodology

A. Wireless Microphone Signals

A WM signal $y(t)$ is typically frequency modulated

$$y(t) = A_c \cos\left(2\pi f_c t + 2\pi\Delta f \int_0^t m(\tau)\mathrm{d}\tau\right) \quad (2)$$

where A_c, f_c, Δf and $m(t)$ represent the carrier magnitude, carrier frequency, frequency deviation and message (voice or audio) signals, respectively. According to [10], WM basically operates in three profiles which can be modeled as:

- *Silent Mode* System user is silent. $m(t)$ is a 32KHz sinusoid and $\Delta f = 5$KHz.
- *Soft Mode* System user is a soft speaker. $m(t)$ is a 3.9KHz sinusoid and $\Delta f = 15$KHz.

- *Loud Mode* System user is a loud speaker. $m(t)$ is a 13.4KHz sinusoid and $\Delta f = 32.6$KHz.

It can be seen from these models that WM signals have a very narrow bandwidth, resembling a pure sinusoid at low SNR. This observation is the starting point for practically all WM detectors. The problem then boils down to detecting low power (-114dBm) sinusoids in a 6MHz DTV channel containing strong noise and interference. One should note that the sinusoid resemblance is strong only when the SNR or SINR is fairly low. Limitations of this approximation can be discovered from simulations and will be discussed in Section III-D. Among the three profiles, the loud mode is the most difficult to detect because its large frequency deviation makes it least resemble a sinusoid.

B. Detecting WM Signals in White Noise with ESPRIT

To detect WM at low SNR with the sinusoid approximation assumption, we apply the ESPRIT [9] algorithm which is widely used in array processing as the core block of the detector. As a subspace method that works with eigenvalues of covariance matrices, ESPRIT has a lot in common with the algorithm proposed in [6]. However, one often cannot extract sufficient information by only looking at eigenvalues, which is in principle a refined energy detector that measures energy in the signal and noise subspaces separately. ESPRIT, on the contrary, simultaneously measures the frequency and power of the buried sinusoids, a crucial augmentation compared to pure eigenvalue detectors, especially when interference enters the picture. We briefly review ESPRIT in the following.

With ESPRIT one assumes $y_n = s_n + \nu_n$ with ν_n being wide sense stationary white noise with power of σ^2 and s_n a sum of sinusoids $s_n = \sum_{\ell=1}^{L} A_\ell e^{j\omega_\ell n}$, each with power $P_\ell = |A_\ell|^2$, $\ell = 1, 2, \ldots, L$. ESPRIT starts with the estimated order–M autocovariance matrix and delayed cross covariance matrix,

$$C = \frac{1}{N-M-1}\sum_{n=M}^{N-2}\mathbf{y}_n\mathbf{y}_n^T,\ C_1 = \frac{1}{N-M-1}\sum_{n=M}^{N-2}\mathbf{y}_n\mathbf{y}_{n+1}^T, \quad (3)$$

where $\mathbf{y}_n = [y_n, y_{n-1}, \ldots, y_{n-M}]^T$. Estimate of the noise power $\hat{\sigma}^2$ is identified as the smallest eigenvalue of C. Next ESPRIT calculates the roots of the matrix pencil

$$(C - \hat{\sigma}^2 I) - \lambda(C_1 - \hat{\sigma}^2 I_1) \quad (4)$$

where I is the identity matrix and I_1 is the zero matrix with the first upper diagonal filled with ones. Among the $M+1$ roots $\{z_m\}_{m=1}^{M+1}$, L of them, say $z_1, z_2, \ldots, z_L$, give the estimates of $e^{j\omega_1}, e^{j\omega_2}, \ldots, e^{j\omega_L}$. The rest are spurious roots. The estimates of $P_1, P_2, \ldots, P_L$ are calculated as diag$[S^\dagger(C - \hat{\sigma}^2 I)(S^H)^\dagger]$ where

$$S = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_L \\ \cdots & \cdots & \cdots & \cdots \\ z_1^M & z_2^M & \cdots & z_L^M \end{bmatrix} \quad (5)$$

is the (estimated) steering matrix. The constant M needs to be carefully chosen – larger M generally better separates the signal subspace from the noise subspace, yet at the same time adds to the computational cost.

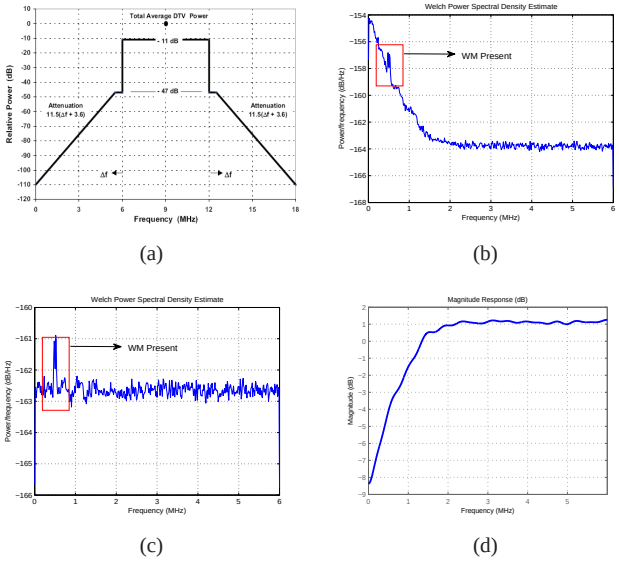


Fig. 1. (a)FCC DTV emission mask (based on measurement bandwidth of 500 kHz).(b)PSD of a 6MHz DTV channel with WM signal at 500KHz from the lower band edge and lower adjacent channel interference, SINR $= -20$dB.(c)The same signal prewhitened.(d)The magnitude response of the prewhitening filter.

Thermal noise generated in the receiver antenna within the 6MHz channel assumes a power of -106dBm at room temperature. With a 10dB detector noise figure, the noise has a power of -96dBm which is 18dB higher than the minimum detectable WM signal power as required by FCC.

C. Interference and Prewhitening

A major source of interference is the spectral leakage from the adjacent digital television (DTV) channel which carries ATSC (Advanced Television Systems Committee) transmission. High power DTV signal is measured at -28dBm [11], in which case the interference produced in the adjacent channel can be as high as -86dBm [1], equivalent to an SINR of -28dB for WM. This phenomenon is reflected in the DTV spectral mask shown in Fig. 1(a) (excerpted from [12]). In this paper, we mainly address how to mitigate the adverse effects of adjacent channel interference due to DTV transmission.

The consequence of adjacent channel interference is that strong colored noise will be present in the channel to be sensed as illustrated in Fig. 1(b). With -20dB SINR, it is possible that the interference almost completely masks out the WM signal. ESPRIT cannot directly handle such difficulty since the interference could induce too many spurious (and strong) harmonics, making identification of the true signal impossible.

To solve this problem, we adopt the strategy proposed in [13] and [14]: before the execution of ESPRIT, a prewhitening filter estimated from an auto-regressive (AR) model is applied to the signal to obtain a flat noise spectrum. Specifically, we may use the Yule-Walker method to extract the filter

coefficients from y_n. Let p denote the filter order and $\mathbf{a} = [1, a_1, a_2, \ldots a_p]^T$ be the coefficients, then $\mathbf{a}$ (as well as the prediction error power $\hat{\sigma}_\epsilon^2$) is obtained by solving

$$\widehat{R}\mathbf{a} = \begin{bmatrix}\hat{\sigma}_\epsilon^2 & 0 & \cdots & 0\end{bmatrix}^T \tag{6}$$

where $\widehat{R}$ is the estimated order-p autocovariance matrix (cf. (3)). Fig. 1(c) shows the same signal after prewhitening. Although the magnitude of the WM signal is equally suppressed with the interference, it now stands out spectrally and can be easily picked up by ESPRIT. Another problem with the prewhitening approach is that when WM is fairly strong, its presence can noticeably mislead the Yule-Walker method such that a spectral notch is created where WM signal sits. However, in this case prewhitening is most likely unnecessary and ESPRIT can be applied directly.

III. WM Detector Design with ESPRIT

With the building blocks of ESPRIT and the prewhitening filter, we can now design a complete WM detector whose structure is shown in Fig. 2.

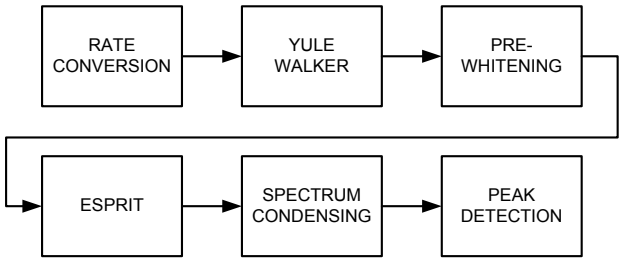


Fig. 2. Structure of the WM detector.

A. Rate Conversion

In practice, the typical sampling rate and IF combinations for DTV tuners are 33.333MHz/8MHz or 21.52MHz/5.38MHz. To fully explore the ability of ESPRIT and to gain maximum frequency resolution, it is advisable to further down shift the IF signal to the baseband and convert the sampling rate to the Nyquist limit, i.e., we assume in this paper that eventually the TV channel signal is centered at 3MHz with a sampling frequency of 12MHz. Practically this is not absolutely necessary but it is helpful in demonstrating the design in this paper.

B. Yule-Walker Estimation

This module estimates the prewhitening filter coefficients with (6). The only parameter we need to choose for this function is the filter order p, which can be determined by the Minimum Description Length (MDL) [15] criterion or by simply picking the smallest p that yields satisfactory results.

C. Prewhitening and ESPRIT

With the estimated filter $\mathbf{a}$, we obtain the whitened signal $x_n = a_n * y_n$ to which we apply ESPRIT. In practice, the sensing time is usually 5ms for signals buried in thermal noise only and 50ms when interference is present. We therefore assume a sensing time of 50ms corresponding to $N = 60000$. With this choice, we can determine the parameter M for ESPRIT. To have a satisfactory performance, we generally require $M \ll N/100$ so that the estimation in (3) is reliable. On the contrary, M must be sufficiently large (e.g., larger than the signal subspace dimension). Although in Section IV we verify the detector's performance with a single potential WM user in the whole channel, it is well possible that multiple WM users exist in a practical scenario. Besides, a real harmonic will be detected as two conjugate harmonics. M should be greater than the largest possible number of true harmonics. However, larger M also implies higher computational complexity. Trade-off has to be made when selecting proper M.

The ESPRIT algorithm yields $M+1$ harmonics $e^{j\omega_\ell n}$ with power P_ℓ, $\ell = 1, 2, \ldots, M+1$ (typically all the roots of (4) are close to the unit circle). We will use all the estimated harmonics to form a pseudo-spectrum for detection.

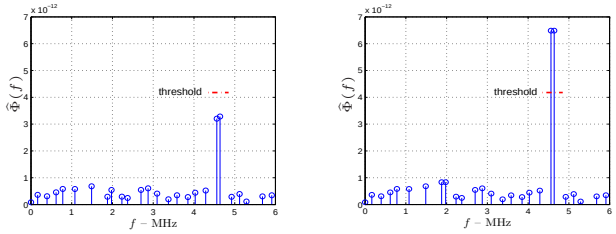
D. Pseudo-Spectrum Condensing

Subspace methods for harmonics estimation often make use pseudo-spectrum (e.g., the MUSIC, Eigenvector and Minimum Norm methods, cf. [16]) which is a function defined on the spectrum where harmonics are shown as peaks. For the WM detector, we form a pseudo-spectrum with ESPRIT, where the peaks represent the power of the harmonics. We define the function $\widehat{\Phi} : \{f_\ell\}_{f_\ell \geq 0} \to \mathbb{R}^+$ as the pseudo-spectrum by combining the conjugate harmonics:

$$\begin{aligned}\widehat{\Phi}(f_\ell) &= 2\Phi(f_\ell), \quad f_\ell > 0,\\ \widehat{\Phi}(0) &= \sum_{\substack{\ell \\ f_\ell = 0}} \Phi(f_\ell).\end{aligned} \tag{7}$$

When WM signal is present in a band where interference is weak, the sinusoid approximation may no longer hold. Especially with loud mode, ESPRIT may produce multiple harmonics to represent the WM signal with the actual signal power distributed among them. On the pseudo-spectrum, this is shown as multiple dwarfed adjacent peaks as illustrated in Fig. 3(a). Spectral splitting severely degrades the performance of the detector based on scanning the peaks of the pseudo-spectrum. To remedy this problem, one may perform spectrum condensing to recombine the signal power. The first approach is to use the ESPRIT outcome to form a refined pseudo-spectrum. Specifically, for each root z_ℓ of (4) and the associated power P_ℓ, a signal model of the form $s_n^\ell = \sqrt{P_\ell} z_\ell^n$, $n = 0, 1, 2, \ldots$ is assumed, which has a spectrum of

$$S^\ell(f) = \frac{\sqrt{P_\ell}}{|1 - z_\ell e^{-j2\pi f/F_s}|}. \tag{8}$$



(a) (b)

Fig. 3. (a)Pseudo-spectrum with spectral splitting at 4.6MHz causing a miss detection.(b)Condensed pseudo-spectrum and successful detection.

A condensed (refined) pseudo-spectrum is then obtained as

$$\overline{\Phi}(f) = \sum_{\ell=1}^{L} S^{\ell}(f). \tag{9}$$

Not only this method recovers the true WM power as the height of the peak, the location of the peak also serves as an estimate of the center frequency of WM. This method works well especially when M is small. If the precise center frequency of the WM signal is not the primary concern, a simpler condensed pseudo-spectrum can be computed instead. As a prior knowledge, we know that WM signal has a bandwidth of no more than 200KHz, therefore, for each harmonic with a frequency of f_ℓ in the pseudo-spectrum, we can recalculate its power as the sum of all harmonics within $\pm$100KHz as identified by ESPRIT,

$$\overline{\Phi}(f_\ell) = \sum_{\substack{1\le j\le M+1 \\ |f_j - f_\ell|<100\text{KHz}}} P_j. \tag{10}$$

It turns out that the simplified method delivers quite good performance when M is large and it is used for the simulation presented in this paper. Fig. 3(b) shows the pseudo-spectrum condensed from Fig. 3(a).

E. Peak Detection and Thresholding

The condensed pseudo-spectrum's peak value $\rho = \max \overline{\Phi}$ serves as the decision statistic. The decision rule is

$$\rho \underset{H_0}{\overset{H_1}{\gtrless}} \rho_{\text{th}}. \tag{11}$$

Usually ρ_{th} is set to produce the desired miss detection probability P_{MD}. As the distribution of ρ is unknown, we can either determine ρ_{th} by generating a large sample of ρ and set ρ_{th} to yield $P_{\text{MD}} \times 100\%$ miss rate for the sample [1], or we can assume an approximate distribution (e.g. by checking the histogram of ρ) and determine the distribution parameters with the Maximum-Likelihood Estimator (MLE). The latter approach is adopted in this paper because, if the distribution is chosen properly, it requires only a small sample size and exhibits little variance. Fig. 4(a) shows the histogram of a sample of ρ with a size of 50, from which we make the following observations:

- The support of the distribution of ρ is approximately $[0, \infty)$;
- The distribution is unimodal, i.e., it has a single local maximum on the support.

There are many distributions sharing the same properties, among which the Gamma distribution is the simplest one. For this reason, we assume the distribution function $q(\rho|k,\theta)$ to be Gamma, specified by k (shape parameter) and θ (scale parameter):

$$q(\rho|k,\theta) = \frac{\rho^{k-1}e^{-\rho/\theta}}{\Gamma(k)\theta^k}. \tag{12}$$

However, since we applied the prewhitening filter whose magnitude response is not flat (cf. Fig. 1(d)), ρ_{th} must be frequency dependent, i.e., we need to set $\rho_{\text{th}}(f)$ for all f between 0 and 6MHz. Similarly, parameter k and θ are also frequency dependent. The only practical way to evaluate $\rho_{\text{th}}(f)$ is by interpolation: we first evaluate $\rho_{\text{th}}(f_i)$ with MLE for a selected collection of $\{f_i\}$, then additional $\rho_{\text{th}}(f)$ is obtained by linear interpolation or extrapolation.

The MLE of k and θ (i.e., $\hat{k}(f_i)$ and $\hat{\theta}(f_i)$) can be obtained, assuming we have formed a sample $\{\rho_j(f_i)\}_{j=1}^{J}$ at frequency f_i, by solving the following equations

$$\begin{aligned} \sum_{j=1}^{J} \log \rho_j(f_i) - J\psi_0(\hat{k}(f_i)) - J\log\hat{\theta}(f_i) = 0, \\ \frac{\sum_{j=1}^{J}\rho_j(f_i)}{\hat{\theta}(f_i)} - J\hat{k}(f_i) = 0, \end{aligned} \tag{13}$$

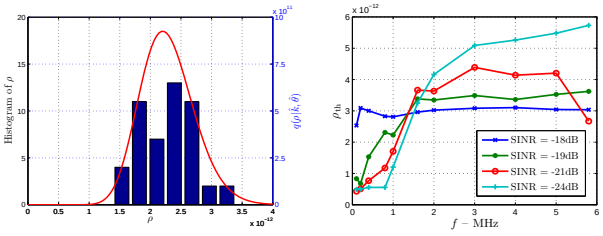
where ψ_0 is the digamma function defined as

$$\psi_0(k) = \frac{\mathrm{d}}{\mathrm{d}k}\log\Gamma(k). \tag{14}$$

Finally, the threshold is determined by

$$\int_0^{\rho_{\text{th}}(f_i)} q(\rho(f_i)|\hat{k}(f_i),\hat{\theta}(f_i))\mathrm{d}\rho(f_i) = P_{\text{MD}}. \tag{15}$$

Fig. 4(b) shows the frequency dependency of thresholds at different SINR's, assuming lower adjacent channel interference exists. We can observe that near the lower boundary where the interference is strong, the threshold is smaller as the prewhitening filter has more suppression in this region. As SINR goes up, the interference becomes weaker since we fix the noise power at -96dBm, hence the curve levels off.



(a) (b)

Fig. 4. With noise and lower adjacent channel interference: (a)histogram of ρ at $f = 800$KHz; (b)threshold–frequency curves, set to achieve $P_{\text{MD}} = 0.01$.

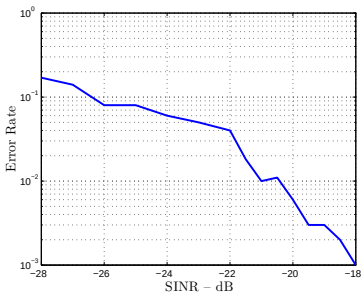


Fig. 5. Error rate performance of the ESPRIT detector at different SINR's.

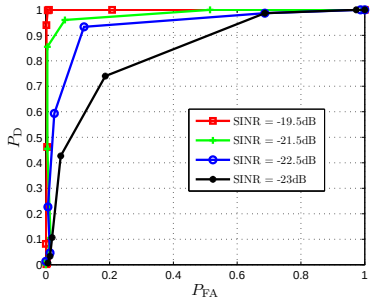


Fig. 6. ROC of the ESPRIT detector.

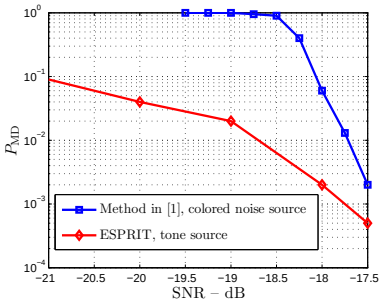


Fig. 7. Comparison between ESPRIT and the detector in [1].

IV. Simulation

Since the loud mode of WM presents the biggest challenge to any detector that uses the sinusoid approximations. Simulation in this section will be carried out solely with the loud mode. We choose $p = 29$ and $M = 51$ based on numerical experiments. They can be set to (moderately) different values without much impact on the results. Sensing time is fixed to 50ms. To evaluate error rate, we fix the signal power at -114dBm and noise power at -96dBm and vary the interference power. WM signal is randomly superimposed on the noise plus interference and the center frequency is also randomly chosen in the 6MHz bandwidth. During the training phase, the threshold is set to achieve $P_{\rm MD} = 0.01$, while we count both misses and false alarms in the testing phase. Fig. 5 shows the error rates against different SINR's, which clearly shows impressive performance of the ESPRIT detector. Even when adjacent channel carries strong DTV signal (SINR $= -28$dB), the error rate is as low as 0.17. It achieves 0.01 error rate at the SINR of -21dB.

The receiver operating characteristics (ROC) of the ESPRIT detector are illustrated in Fig. 6. Since the detector threshold is frequency dependent, the ROC have to be drawn with regard to a single frequency. We fix the WM signal 500KHz from the lower band edge where interference is moderately strong and show the ROC for different SINR's. It can be seen that, impressively, the detector becomes very reliable as soon as SINR is above -21.5dB.

For comparative evaluation, in Fig. 7 we show the miss detection probability P_{MD} of our detector along with a detector recently proposed in [1] where a different colored noise model is assumed for the WM source. As in [1] which uses a spectral analysis based technique, the results are shown for a fixed interference level of -86dBm, noise level of -96dBm and false alarm probability of 0.01. The results are shown as a function of the SNR obtained by varying the signal power and it can be seen that our detector outperforms that in [1] even at very low SNR's.

V. Conclusion

In this paper, we presented a novel WM detector based on the ESPRIT algorithm and auxiliary prewhitening, which shows impressive performance even in the case of strong adjacent channel DTV interference, achieving an error rate of 0.01 at -21dB SINR. We also introduced the concept of spectral condensation to improve the performance when we have a medium detector SINR.

References

[1] H. C. Chen and W. Gao, "Spectrum sensing for FM wireless microphone signals," in *IEEE DySPAN 2010*, Singapore, April 2010.

[2] S. J. Shellhammer, "Spectrum sensing in IEEE 802.22," in *CIP 2008*, 2008.

[3] Y. H. Zeng, Y. C. Liang, A. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, 2010.

[4] H. S. Chen, W. Gao, and D. Daut, "Spectrum sensing for wireless microphone signals," in *SECON Workshops '08*, San Francesco, CA, July 2008.

[5] A. M. Mossa and V. Jeoti, "Cyclostationarity-based spectrum sensing for analog TV and wireless microphone signals," in *First International Conference on Computational Intelligence, Communication Systems and Networks*, 2009, pp. 380–385.

[6] Y. H. Zeng and Y. C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, June 2009.

[7] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Trans. Signal Processing*, vol. 56, no. 5, pp. 1801–1811, April 2008.

[8] Z. Tian, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *CROWNCOM*, 2006.

[9] R. Roy and T. Kailath, "ESPRIT – estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 37, no. 7, pp. 984–995, July 1989.

[10] C. Clanton, M. Kenkel, and Y. Tang, "Wireless microphone signal simulation method," IEEE 802.22-07/0124r0, March 2007.

[11] FCC, "Evaluation of the performance of prototype TV-band white space devices phase II," OET 08-TR-1005, October 2008.

[12] A. T. S. Committee, "ATSC recommended practice: Transmission measurement and compliance for digital television," Document A/64B, May 2008.

[13] S. Kay and S. Salisbury, "Improved active sonar detection using autoregressive prewhiteners," *J. Acoust. Soc. Am.*, vol. 87, pp. 1603–1611, 1990.

[14] V. Carmillet, P.-O. Amblard, and G. Jourdain, "Detection of phase- or frequency-modulated signals in reverberation noise," *J. Acoust. Soc. Am.*, vol. 105, no. 6, pp. 3375–3389, June 1999.

[15] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1978.

[16] S. Orfanidis, "Optimum signal processing," http://www.ece.rutgers.edu/~orfanidi/osp2e/osp2e.pdf, 2007.