Performance of Compressive Sensing Based Energy Detection

Eva Lagunas
Interdisciplinary Centre for
Security, Reliability and Trust (SnT)
University of Luxembourg, Luxembourg
Email: eva.lagunas@uni.lu

Luca Rugini
Department of Engineering
University of Perugia
Perugia, Italy
Email: luca.rugini@unipg.it

***Abstract*—This paper investigates closed-form expressions to evaluate the performance of the Compressive Sensing (CS) based Energy Detector (ED). The conventional way to approximate the probability density function of the ED test statistic invokes the central limit theorem and considers the decision variable as Gaussian. This approach, however, provides good approximation only if the number of samples is large enough. This is not usually the case in CS framework, where the goal is to keep the sample size low. Moreover, working with a reduced number of measurements is of practical interest for general spectrum sensing in cognitive radio applications, where the sensing time should be sufficiently short since any time spent for sensing cannot be used for data transmission on the detected idle channels. In this paper, we make use of low-complexity approximations based on algebraic transformations of the one-dimensional Gaussian Q-function. More precisely, this paper provides new closed-form expressions for accurate evaluation of the CS-based ED performance as a function of the compressive ratio and the Signal-to-Noise Ratio (SNR). Simulation results demonstrate the increased accuracy of the proposed equations compared to existing works.**

I. Introduction

Signal acquisition is a crucial processing step in signal processing. Compressive Sensing (CS) techniques have become increasingly prevalent in signal processing over the past decade for their ability to perform data acquisition and compression simultaneously [1], [2]. The CS paradigm relies on the assumption that the signal of interest has a sparse representation in some domain, which is an inherent characteristic of the licensed radio-frequency spectrum [3]. As a consequence, CS has been explored within the context of several topics related to signal processing and Cognitive Radio (CR), being spectrum sensing one of the most popular [4]–[6].

In this paper, we focus on compressive signal processing [7], where inference techniques are directly applied on the compressed domain. The main advantage of performing classical signal processing operations directly on the compressive measurements is that it avoids the computationally expensive reconstruction algorithms in order to obtain the original non-compressed signal. More specifically, in this paper we investigate the Energy Detector (ED) [8] in conjunction with CS-based signal acquisition, where only a small subset of measurements is employed for signal detection. ED is the optimal Neyman-Pearson (NP) detector (also known as the likelihood-ratio test) for stochastic signals in White Gaussian Noise (WGN) environments [9]. Herein we focus on the derivation of the test statistic for the CS-based ED and the formulation of the corresponding equations related to the detection performance.

Most of the existing contributions in the context of ED performance mainly focus on the non-compressed scenario, and only a few contributions address the CS-based counterpart. The performance evaluation of CS signal detection has been previously considered in [7], [10]–[12]. In [7], the primary signal is modeled as deterministic and therefore the analysis of [7] does not exploit the knowledge of the probability density function (pdf) of the primary signal. Differently from [7], our investigation assumes a stochastic signal model that is suitable to current digital transmissions. Indeed, as briefly summarized in [4], many primary signals to be sensed are practically both white (their power spectrum is almost constant) and Gaussian (they are obtained as linear combination of hundreds of variables), because many current standards employ Orthogonal Frequency-Division Multiplexing (OFDM). Stochastic signal models are used in [10] and [11], which both exploit a sparse signal model (in a subspace domain and in the frequency domain, respectively) to estimate the unknown signal covariance and the unknown primary signal, respectively. In [12], the central limit theorem is used to approximate the pdf of the CS-based ED test, thereby simplifing the expressions of the probability of detection and the probability of false alarm. However, the central limit theorem provides accurate estimates only when the number of samples is sufficiently large [13]. Therefore, the analysis of [12] is accurate only when the undersampling caused by CS does not produce a small number of observations.

In this paper, we extend the work in [12]–[14] by merging the approximation techniques proposed in [13], [14] with the CS framework presented in [12]. In particular, we propose novel analytical closed-form expressions for the probability of detection of the CS-based ED when the number of available measurements is too low

978-1-5386-3531-5/17/$31.00

to apply the central limit theorem. To this aim, we exploit the power transformations with generic exponents proposed in [13], [14], which have been shown to be a good alternative for low-complexity devices. Specifically, we formulate a family of equations to evaluate the CS-based ED performance, to easily quantify the effect of the reduced number of samples, and to determine the minimum number of samples needed to achieve certain detection performance. We show through numerical experiments that the power transformations accurately approximate the central chi-squared random variable resulting from the ED test statistic, also when the number of observations is low. Note that, due to different signal and noise power scenarios, a sensor may have to dynamically self-adjust its threshold: hence, the proposed expressions are very useful for low-complexity sensors that cannot numerically invert a chi-squared pdf due to their limited computation capabilities.

The remainder of this paper is structured as follows. In Section II, we introduce the CS-based ED formulation and derive the corresponding test statistic. In Section III, we derive the exact performance of the CS-based ED, while in Section IV we propose the power transformations to approximate the complicated exact analysis. Finally, supporting numerical results are provided in Section V, and Section VI states the conclusion.

II. CS-Based Energy Detector

We consider the compressive acquisition of a received signal $\mathbf{x} \in \mathbb{R}^N$, assuming that this has a sparse representation in some domain. In this case, the information contained in $\mathbf{x}$ can be inferred from a small set of M ($M \ll N$) linear measurements $\mathbf{y} \in \mathbb{R}^M$, $\mathbf{y} = \mathbf{\Phi x}$, where $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ represents the sensing matrix. Depending on the hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$, the received signal is expressed by $\mathbf{x} = \mathbf{w}$ or $\mathbf{x} = \mathbf{s} + \mathbf{w}$, respectively, where $\mathbf{s} \in \mathbb{R}^N$ denotes the signal to be detected, which is assumed to be a Gaussian random process with zero mean and variance σ_s^2, i.e., $\mathbf{s} \sim \mathcal{N}\left(\mathbf{0}, \sigma_s^2 \mathbf{I}_N\right)$, and $\mathbf{w} \sim \mathcal{N}\left(\mathbf{0}, \sigma_w^2 \mathbf{I}_N\right)$ is the WGN independent of the signal $\mathbf{s}$. Thus, the CS-based ED can be formulated as

$$\begin{cases} \mathcal{H}_0: & \mathbf{y} = \mathbf{\Phi w}, \\ \mathcal{H}_1: & \mathbf{y} = \mathbf{\Phi}\left(\mathbf{s} + \mathbf{w}\right). \end{cases} \quad (1)$$

The resulting compression ratio is defined as $\rho = M/N$. For example, if $\rho = 0.25$, then $4\times$ undersampling is achieved.

A. Derivation of the Test Statistic

Since $\mathbf{y}$ in (1) is zero-mean Gaussian distributed with covariance $\sigma_w^2 \mathbf{\Phi\Phi}^T$ or $(\sigma_s^2 + \sigma_w^2)\mathbf{\Phi\Phi}^T$ for the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, the pdf of $\mathbf{y}$ is expressed by [12]

$$\begin{cases} \mathcal{H}_0: \quad f_0(\mathbf{y}) = \dfrac{\exp\left(-\frac{1}{2}\mathbf{y}^T\left(\sigma_w^2 \mathbf{\Phi\Phi}^T\right)^{-1}\mathbf{y}\right)}{\sqrt{\left|\sigma_w^2 \mathbf{\Phi\Phi}^T\right|(2\pi)^M}}, \\ \mathcal{H}_1: \quad f_1(\mathbf{y}) = \dfrac{\exp\left(-\frac{1}{2}\mathbf{y}^T\left((\sigma_s^2+\sigma_w^2)\mathbf{\Phi\Phi}^T\right)^{-1}\mathbf{y}\right)}{\sqrt{\left|(\sigma_s^2+\sigma_w^2)\mathbf{\Phi\Phi}^T\right|(2\pi)^M}}. \end{cases} \quad (2)$$

Note that the pdf expressions in (2) are different from those in [7], due to the different signal model.

In this work, we focus on sensing matrices that satisfy $\mathbf{\Phi\Phi}^T = \frac{1}{\rho}\mathbf{I}_M$. These are sensing matrices whose rows are orthogonal and whose columns have unit norm. Note that this is approximately satisfied by the popular matrices composed of zero-mean i.i.d. random variables with variance $1/M$, and also by the widely used sensing matrices employed in the multi-coset sampler [15], [16]. In Section V, we show simulation results for both cases, i.e., for sensing matrices that satisfy the assumption $\mathbf{\Phi\Phi}^T = \frac{1}{\rho}\mathbf{I}_M$ either exactly or approximately.

According to the previous discussion, the corresponding NP decision statistic is given by,

$$L(\mathbf{y}) = \left[\frac{\sigma_w^2}{\sigma_s^2+\sigma_w^2}\right]^{M/2} \exp\left(\left(\frac{\sigma_s^2}{2\frac{1}{\rho}\sigma_w^2(\sigma_s^2+\sigma_w^2)}\right)\|\mathbf{y}\|_2^2\right). \quad (3)$$

Clearly, the log-likelihood ratio $L(\mathbf{y})$ in (3) can be replaced by the simpler ED test statistic

$$T(\mathbf{y}) = \|\mathbf{y}\|_2^2 = \sum_{i=1}^{M} y_i^2. \quad (4)$$

To decide the signal presence or absence, the test in (4) is compared to a threshold η': when $T(\mathbf{y}) \geq \eta'$, the CS-based ED decides that $\mathbf{s}$ is present, otherwise $\mathbf{s}$ is assumed as absent.

III. Exact Performance Analysis

Herein we derive the Receiver Operating Characteristic (ROC) of the CS-based ED. By exploiting $\mathbf{\Phi\Phi}^T = \frac{1}{\rho}\mathbf{I}_M$ in (2), the test $T(\mathbf{y})$ in (4) is a central chi-squared random variable with $M = \rho N$ degrees of freedom. Hence, the probability of detection P_D is expressed by

$$P_D = \Pr\{T(\mathbf{y}) > \eta'; \mathcal{H}_1\} = 1 - F_{\rho N}\left(\frac{\eta' \rho}{\sigma_s^2 + \sigma_w^2}\right), \quad (5)$$

where $F_{\rho N}(x)$ is the cumulative distribution function (cdf) of a central chi-squared random variable with ρN degrees of freedom, expressed by

$$F_{\rho N}(x) = \frac{\int_0^{x/2} \nu^{\rho N/2-1}\exp(-\nu)d\nu}{\Gamma\left(\rho N/2\right)}. \quad (6)$$

Similarly to (5), the probability of false alarm P_{FA} can be expressed as

$$P_{FA} = \Pr\{T(\mathbf{y}) > \eta'; \mathcal{H}_0\} = 1 - F_{\rho N}\left(\frac{\eta'\rho}{\sigma_w^2}\right). \quad (7)$$

By inverting (7), we obtain an analytical expression for the threshold η', expressed by

$$\eta' = \frac{\sigma_w^2}{\rho} F_{\rho N}^{-1}\left(1 - P_{FA}\right), \quad (8)$$

where $F_{\rho N}^{-1}(x)$ is the inverse of $F_{\rho N}(x)$ with respect to the argument x. By inserting (8) into (5), we obtain the exact ROC expression, given by

$$P_D = 1 - F_{\rho N}\left(\frac{\sigma_w^2}{\sigma_s^2 + \sigma_w^2} F_{\rho N}^{-1}\left(1 - P_{FA}\right)\right). \quad (9)$$

Since the number of degrees of freedom ρN in (9) is proportional to the compression ratio ρ, clearly the probability of detection P_D increases with ρ.

Note that the threshold η' in (8) and the probability of detection P_D in (9) can be evaluated using mathematical software toolboxes (e.g., MATLAB, or Mathematica). One possible approach is to calculate η' and P_D using a powerful computing machine, and then load the precomputed threshold η' onto the low-complexity sensor. However, in many cases, this approach is impractical. For instance, when a constant false alarm rate is required, the threshold must be updated taking into account the noise power σ_w^2, which can be different depending on many conditions (e.g., temperature). In addition, η' heavily depends on the number of degrees of freedom ρN, and therefore, the threshold has to be updated also when the sample size changes. In other words, a low-complexity sensor would need a two-dimensional lookup table with large memory to store all the possible precomputed thresholds. Hence, a more clever approach is to let the sensor calculate or update its threshold η'. Unfortunately, the computation of the exact threshold can be hardly performed by a low-complexity low-memory sensor. Indeed, (8) requires the inverse function $F_{\rho N}^{-1}(x)$ of the cdf $F_{\rho N}(x)$, which is a two-dimensional function with one continuous parameter x, and one discrete parameter ρN. Basically, for each value of ρN, a different one-dimensional inverse function $F_{\rho N}^{-1}(x)$ is necessary. Therefore, in order to reduce the complexity and the memory requirements of the sensor, in the next section we propose accurate approximations of (9) and (8), aiming at avoiding the use of $F_{\rho N}^{-1}(x)$ and $F_{\rho N}(x)$.

IV. Approximated Performance Analysis

Several approximations of (9) are possible, including the popular Gaussian approximation [12]. Clearly, Gaussian approximations are accurate at large sample size and therefore, can be inaccurate when either N or the compression ratio ρ is small. Here, we make use of the power transformation approach [13], [14]. Basically, the power transformation approach approximates a central chi-squared random variable with the rth power of a Gaussian variable with appropriate mean and variance. The statistical relations between chi-squared random variables and power-transformed Gaussian variables are well known in the statistical literature [17]–[19], but have been investigated only recently for CR applications [13], [14]. Specifically, $F_{\rho N}(x)$ is approximated as

$$\hat{F}_{\rho N}(x) = 1 - Q\left(\frac{[x/(\rho N)]^{1/r} - m_r(\rho, N)}{[V_r(\rho, N)]^{1/2}}\right), \quad (10)$$

where $Q(x) = (2\pi)^{-1/2}\int_x^{+\infty} \exp\left(-\nu^2/2\right) d\nu$, and the mean $m_r(\rho, N)$ and the variance $V_r(\rho, N)$ in (10) are expressed by

$$m_r(\rho, N) = 1 - \frac{r-1}{r^2 \rho N}, \quad V_r(\rho, N) = \frac{2}{r^2 \rho N}. \quad (11)$$

Note that $r = 1$ corresponds to a Gaussian approximation [12], while $r = 3$ corresponds to a cube-of-Gaussian approximation [13]. By inverting (10), we obtain

$$\hat{F}_{\rho N}^{-1}(x) = \rho N \left\{[V_r(\rho, N)]^{1/2} Q^{-1}(1-x) + m_r(\rho, N)\right\}^r. \quad (12)$$

An approximated expression for the threshold η' can be obtained by replacing $F_{\rho N}^{-1}(x)$ with (12), and inserting it into (8), thereby obtaining

$$\hat{\eta}' = \sigma_w^2 N \left\{[V_r(\rho, N)]^{1/2} Q^{-1}(P_{FA}) + m_r(\rho, N)\right\}^r. \quad (13)$$

Differently from (8), the approximation (13) depends on the one-dimensional function $Q^{-1}(x)$, and consequently $\hat{F}_{\rho N}^{-1}(x)$ is easier to calculate than $F_{\rho N}^{-1}(x)$ and easier to store in a lookup table.

The equations (10) and (12) can be exploited in order to obtain an analytical approximation for the ROC. By replacing $F_{\rho N}(x)$ and $F_{\rho N}^{-1}(x)$ with (10) and (12), respectively, and by inserting them into (9), we obtain

$$\hat{P}_D = Q\left(S_r(\sigma_s^2, \sigma_w^2) Q^{-1}(P_{FA}) - \left[1 - S_r(\sigma_s^2, \sigma_w^2)\right] R_r(\rho, N)\right), \quad (14)$$

$$S_r(\sigma_s^2, \sigma_w^2) = \left(\frac{\sigma_w^2}{\sigma_s^2 + \sigma_w^2}\right)^{1/r}, \quad R_r(\rho, N) = \frac{m_r(\rho, N)}{[V_r(\rho, N)]^{1/2}}, \quad (15)$$

where $m_r(\rho, N)$ and $V_r(\rho, N)$ are expressed by (11). Like the approximated threshold (13), the approximated ROC (14) depends on one-dimensional functions only, such as $Q(x)$ and $Q^{-1}(x)$. Using (11), the ratio $R_r(\rho, N)$ in (15) can be expressed as

$$R_r(\rho, N) = r\left(\frac{\rho N}{2}\right)^{1/2} - \frac{r-1}{r(2\rho N)^{1/2}}, \quad (16)$$

which depends on the product $M = \rho N$, rather than on ρ and N separately.

We now make use of (14) to calculate an analytical expression for the the number of samples $\hat{N}$ required for a desired performance (P_{FA}, P_D). We assume that the primary signal power σ_s^2, the noise power σ_w^2, and the compression ratio ρ, are fixed. By defining

$$b_r(P_{FA}, P_D, \sigma_s^2, \sigma_w^2) = \frac{S_r(\sigma_s^2, \sigma_w^2) Q^{-1}(P_{FA}) - Q^{-1}(P_D)}{1 - S_r(\sigma_s^2, \sigma_w^2)}, \quad (17)$$

and by inverting (14), we obtain

$$b_r(P_{FA}, P_D, \sigma_s^2, \sigma_w^2) = R_r(\rho, \hat{N}). \quad (18)$$

Equation (18), together with (16), leads to a quadratic equation in $\hat{N}^{1/2}$, whose positive solution is

$$\hat{N} = \frac{1}{\rho}\left(\frac{b_r}{2^{1/2} r} + \sqrt{\frac{b_r^2}{2r^2} + \frac{r-1}{r^2}}\right)^2, \quad (19)$$

where $b_r = b_r(P_{FA}, P_D, \sigma_s^2, \sigma_w^2)$ is expressed by (17) and (15). Equation (19) shows that the required number of samples $\hat{N}$ increases when the compression ratio ρ is reduced, and when b_r in (17) increases. From (19), the required compression ratio $\hat{\rho}$ for a desired

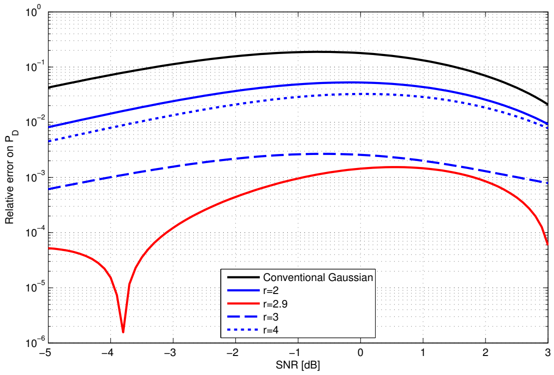


Fig. 1: Relative error on P_D versus SNR for $P_{FA} = 10^{-3}$, $N = 200$ and $\rho = 0.15$.

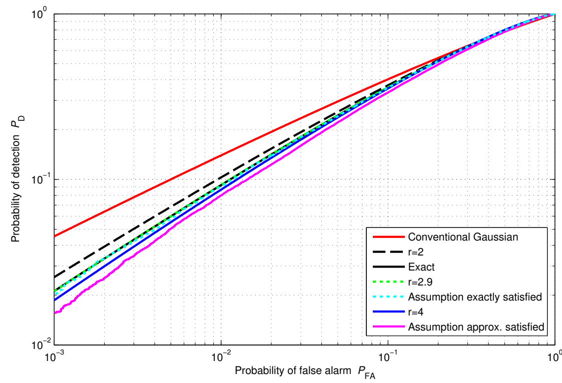


Fig. 2: ROC (P_D versus P_{FA}) for SNR$= -6$dB, $N = 200$ and $\rho = 0.15$.

performance (P_{FA}, P_D) with a fixed number of samples N is expressed by

$$\hat{\rho} = \frac{1}{N}\left(\frac{b_r}{2^{1/2}r} + \sqrt{\frac{b_r^2}{2r^2} + \frac{r-1}{r^2}}\right)^2. \qquad (20)$$

Note that (20) is valid only when the solution $\hat{\rho} \leq 1$.

Clearly, a different choice for the power exponent r leads to a different approximations (13), (14), (19), and (20). In the next section, we make use of simulation results to check the accuracy of different choices for r in the context of CS-based ED.

V. Simulation results

To illustrate the accuracy of the proposed CS-based ED performance when dealing with very low number of measurements, we consider a total number of samples of $N = 200$ and a very ambitious compression ratio of $\rho = 0.15$. First, we assume that the sensing matrix $\mathbf{\Phi}$ exactly satisfies $\mathbf{\Phi}\mathbf{\Phi}^T = \frac{1}{\rho}\mathbf{I}_M$ and, thus, the equations derived in Section IV exactly apply.

Fig. 1 shows the relative error ϵ on the probability of detection obtained with the proposed approximations, as a function of the SNR, when $P_{FA} = 10^{-3}$. The relative error is defined as $\epsilon = |\hat{P}_D - P_D|/P_D$, while the SNR is defined as SNR$= \frac{\sigma_s^2}{\sigma_w^2}$. The conventional Gaussian approximation [12] is also depicted in Fig. 1 for comparison purposes. Note that the conventional Gaussian approximation provides the highest error, being the weakest approximation of the plot. This is because the central limit theorem does not hold anymore for very low number of samples, in this case for $M = \rho N = 30$. Specifically, in Fig. 1, *the relative error of the conventional approximation is always between 2% and 20%.* Differently, the proposed power approximations produce a significantly reduced relative error on the probability of detection. We evaluated several values of the exponent r, and the most relevant results were included in Fig. 1: *the best approximation is obtained when $r = 2.9$, which provides a relative error always below 0.2%.*

Fig. 2 compares the relationship between P_{FA} and P_D provided by the proposed approximations for SNR$= -6$dB, $N = 200$ and $\rho = 0.15$. In Fig. 2, the exact relationship (9) has been included, as well as the conventional Gaussian approximation, for comparison purposes. We also plot in Fig. 2 the empirical results obtained from $100,000$ Monte Carlo runs with random and deterministic $\mathbf{\Phi}$ matrix. Note that for a CS-based acquisition system, $\mathbf{\Phi}$ can be either composed of random variables or deterministic. For the random case, we assumed $\mathbf{\Phi}$ with entries being a zero-mean i.i.d. Gaussian random variables with variance $1/M$, so that $\mathbf{\Phi}\mathbf{\Phi}^T \approx \frac{1}{\rho}\mathbf{I}_M$. For the deterministic case, we used $\mathbf{\Phi} = \frac{1}{\sqrt{\rho}}\left[\mathbf{I}_M \quad \mathbf{0}_{M\times(N-M)}\right]$ so that the assumption $\mathbf{\Phi}\mathbf{\Phi}^T = \frac{1}{\rho}\mathbf{I}_M$ is exactly satisfied. Clearly, the curve corresponding to the approximation with $r = 2.9$, and the empirical curve that perfectly satisfies the assumption, both overlap with the curve of the exact performance. On the other hand, for the random $\mathbf{\Phi}$, where the assumption does not perfectly hold true, we observe a slight loss of accuracy.

Fig. 3 compares the P_D versus SNR curve obtained with the conventional Gaussian approximation, the one obtained with the proposed power approximation with $r = 2.9$ and the empirical results obtained from $100,000$ Monte Carlo runs with random and deterministic $\mathbf{\Phi}$ matrix as described above. In Fig. 3, we fixed $P_{FA} = 10^{-3}$, $N = 200$ and $\rho = 0.15$. It can be observed from Fig. 3 that, although the proposed curve does not perfectly match with the empirical curve when the assumption is approximately satisfied, it provides a much better approximation to real world CS-based ED system than that obtained by the conventional Gaussian approximation. The discrepancy between the empirical random case and the proposed theoretical curve is due to the weakness of the sensing matrix assumption. How to amended this error based on the knowledge of $\mathbf{\Phi}$ is left for future works.

Finally, Fig. 4 illustrates the usefulness of (20) for determining the required compression ratio ρ to achieve a certain detection performance (P_{FA}, P_D). The results in Fig. 4 assume $N = 200$, $r = 2.9$ and SNR$= 2$dB.

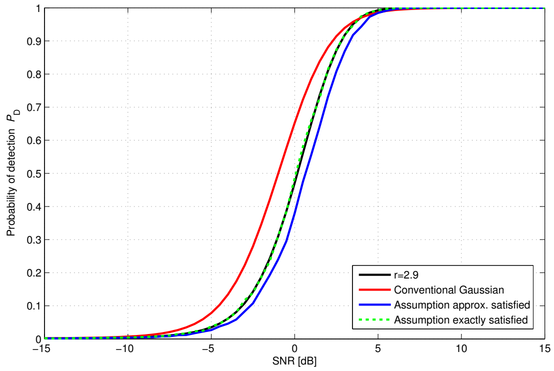
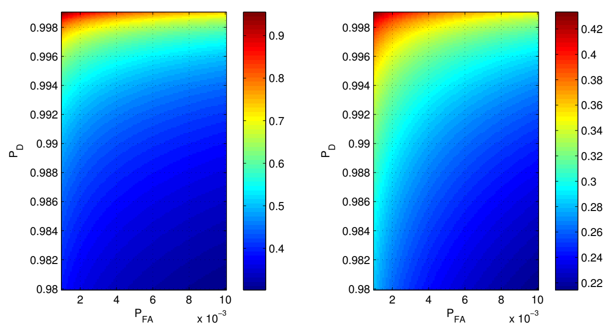


Fig. 3: P_D versus SNR for $P_{FA} = 10^{-3}$, $N = 200$ and $\rho = 0.15$.



(a) (b)

Fig. 4: $\hat{\rho}$ versus (P_{FA}, P_D) for SNR= 2dB, $N = 200$ and $r = 2.9$: (a) Gaussian, (b) Proposed.

As expected, the compression ratio increases as P_D increases and P_{FA} decreases. It can be observed that, for a particular detection performance point (P_{FA}, P_D), the accuracy of the proposed expression allows to use a lower number of samples than that estimated with the conventional Gaussian approximation.

VI. Conclusion

We have provided accurate approximations to evaluate the performance of the CS-based ED for small number of measurements. Essentially, we have made use of low-complexity power transformations to estimate the cdf of the central chi-squared random variable resulting from the CS-based ED test statistic. Simulation results have demonstrated the effectiveness of the proposed approach when dealing with significantly reduced data measurements and have proven its superiority versus a conventional Gaussian approximation.

Acknowledgment

This work was partially supported the National Research Fund, Luxembourg, under CORE project SeMIGod and SATSENT.

References

[1] E.J. Candes and M.B. Wakin, “An introduction to compressed sampling,” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

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

[3] B. Wang and K.J.R. Liu, “Advances in cognitive radio networks: A survey,” *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 5–23, Feb. 2011.

[4] 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.

[5] S.K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, “Applications of compressive sensing in cognitive radio communications: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1838–1860, 2016.

[6] E. Lagunas and M. Najar, “Spectral feature detection with sub-Nyquist sampling for wideband spectrum sensing,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3978–3990, Jul. 2015.

[7] M.A. Davenport, P.T. Boufounos, M.B. Wakin, and R.G. Baraniuk, “Signal processing with compressive measurements,” *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 3, pp. 445–460, Apr. 2010.

[8] H. Urkowitz, “Energy detection of unknown deterministic signals,” *Proc. of the IEEE*, vol. 55, pp. 523–531, Apr. 1967.

[9] S.M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*, Prentice Hall, New Jersey, USA, 1998.

[10] B. Shankar, S. Chatterjee, and B. Ottersten, “Detection of sparse random signals using compressive measurements,” *Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP), Kyoto, Japan*, Mar. 2012.

[11] M. Başaran, S. Erküçük, and H.A. Çırpan, “Bayesian compressive sensing for primary user detection,” *IET Signal Process.*, vol. 10, no. 5, pp. 514–523, 2016.

[12] E. Lagunas, S.K. Sharma, S. Chatzinotas, and B. Ottersten, “Compressive sensing based energy detector,” *European Signal Process. Conf. (EUSIPCO), Budapest, Hungary*, pp. 1678–1682, Aug. 2016.

[13] L. Rugini, P. Banelli, and G. Leus, “Small sample size performance of the energy detector,” *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1814–1817, Sep. 2013.

[14] L. Rugini, P. Banelli, and G. Leus, “Spectrum sensing using energy detectors with performance computation capabilities,” *European Signal Process. Conf. (EUSIPCO), Budapest, Hungary*, pp. 1608–1612, Aug. 2016.

[15] P. Feng, *Universal Minimum-Rate Sampling and Spectrum-Blind Reconstruction for Multiband Signals*, PhD Thesis, Univ. of Illinois, Urbana-Champaign, USA, 1997.

[16] R. Venkataramani and Y. Bresler, “Optimal sub-Nyquist nonuniform sampling and reconstruction for multiband signals,” *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2301–2313, Oct. 2001.

[17] E.B. Wilson and M.M. Hilferty, “The distribution of chi-square,” *Proc. Nat. Acad. Sci.*, vol. 17, no. 4, pp. 684–688, Nov. 1931.

[18] D.M. Hawkins and R.A.J. Wixley, “A note on the transformation of chi-squared variables to normality,” *Amer. Statist.*, vol. 40, pp. 296–298, Nov. 1986.

[19] M.N. Goria, “On the fourth root transformation of chi-square,” *Austral. J. Statist.*, vol. 34, no. 1, pp. 55–64, 1992.