

Multi-Antenna Assisted Spectrum Sensing in Spatially Correlated Noise Environments[☆]

Ali Koochakzadeh^{a,*}, Mohammadreza Malek-Mohammadi^a, Massoud Babaie-Zadeh^a, Mikael Skoglund^b

^aElectrical Engineering Department, Sharif University of Technology, Tehran 1458889694

^bCommunication Theory Lab, KTH, Royal Institute of Technology, Stockholm, 10044, Sweden

Abstract

A significant challenge in spectrum sensing is to lessen the signal to noise ratio needed to detect the presence of primary users while the noise level may also be unknown. To meet this challenge, multi-antenna based techniques possess a greater efficiency compared to other algorithms. In a typical compact multi-antenna system, due to small interelement spacing, mutual coupling between thermal noises of adjacent receivers is significant. In this paper, unlike most of the spectrum sensing algorithms which assume spatially uncorrelated noise, the noises on the adjacent antennas can have arbitrary correlations. Also, in contrast to some other algorithms, no prior assumption is made on the temporal properties of the signals. We exploit low-rank/sparse matrix decomposition algorithms to obtain an estimate of noise and received source covariance matrices. Given these estimates, a Semi-Constant False Alarm Rate (S-CFAR) detector, in which the probability of false alarm is constant over the scaling of the noise covariance matrix, to examine the presence of primary users is proposed. In order to analyze the efficiency of our algorithm, we derive approximate probability of detection. Numerical simulations show that the proposed algorithm consistently and considerably outperforms state-of-the-art multi-antenna based spectrum sensing algorithms.

Keywords: Spectrum Sensing, Semi-Constant False Alarm Rate (S-CFAR), Cognitive Radio (CR), Generalized Likelihood Ratio Test (GLRT), Low-Rank/Sparse Matrix Decomposition, Spatial Correlation

1. Introduction

Cognitive radio (CR) [1], a novel technique emerged in the last decade, is intended to improve utilization of the radio spectrum. In CR networks, in addition to the typical primary users, working under licensed frequency bands, there are secondary unlicensed users who seek to opportunistically exploit the same spectrum resources when the primary users do not transmit any data. The key function for the secondary users is the ability to detect the occupancy of spectrum resources, which is known as spectrum sensing. To create the least possible interference to the primary users, a spectrum sensing algorithm should be able to detect the presence of primary signals, even in very low signal to noise ratios (SNR).

So far, there are many different algorithms proposed for spectrum sensing, namely energy detection [2–4], the Matched Filtering (MF) [4], cyclostationary detection [5, 6], and multi-antenna assisted detection [7–9], each of which has advantages and disadvantages. Matched filtering and cyclostationary detection need some prior knowledge about primary signal's modulation properties such as pulse shaping, carrier frequencies, timing, etc. However, in practice, they may not be available for secondary users. Energy detection, on the other hand, is a basic method that unlike the aforementioned methods is independent of signal properties, that makes it a desirable spectrum-sensing approach for CR. The main drawback of this approach is its reliance on an accurate noise power estimation, which

[☆]This work was supported in part by the Iran National Science Foundation under contract number 91004600.

*Corresponding author. Tel: +98 912 1867064.

Email addresses: ali_k@ee.sharif.edu (Ali Koochakzadeh), m.rezamm@ieee.org (Mohammadreza Malek-Mohammadi), mbzadeh@yahoo.com (Massoud Babaie-Zadeh), skoglund@ee.kth.se (Mikael Skoglund)

makes it less practical, since noise power usually fluctuates under circumstance changes such as aging of devices, humidity, and temperature.

Multi-antenna approaches generally do not need any *priori* assumptions about primary signal properties. Specially, methods based on the eigenvalues of the received sample covariance matrix have been center of interest in the last few years, since they outperform the popular energy detection methods [7]. For instance, the maximum-to-minimum ratio eigenvalue (MME) detector employs the ratio of the maximum eigenvalue to the minimum eigenvalue of the covariance matrix of the received signal [8]. Better detection statistics, like scaled largest eigenvalue (SLE), are obtained by deriving Generalized Likelihood Ratio Test (GLRT) for spatially uncorrelated noises with equal power in [10], [11]. Furthermore, in [12], authors derive the asymptotic GLRT for spatially uncorrelated noises with equal or different unknown power spectral densities.

Current demands for compact multi-antenna receivers lead to appearance of antenna arrays with interelement spacing much smaller than half a wavelength. As a side effect, neglecting the mutual coupling between closely spaced antennas seems to be no longer practical (cf. [13–15]). Mutual coupling can cause strong correlation between noises of neighboring antennas [16]. However, there are a few algorithms that handle *spatially correlated noise* environments which also make additional assumptions on the temporal properties of the signals. For instance, [17] assumes that the temporal correlation of the primary user’s signal is known up to a scalar factor, and [18] exploits cyclostationary features of the primary user’s signal.

In this paper, without making any assumption on the temporal properties of the primary users’ signals, we propose a novel detection algorithm for an unknown spatially correlated noise environment where there are arbitrary correlations between the noise of adjacent sensors. This method is efficient for custom spatial correlations between noises as far as the noise spatial covariance matrix is sparse, meaning that the number of nonzero entries is much smaller than the total number of entries. For example, in sparse arrays of sensors, consisting of several largely spaced subarrays, due to the inter-subarray noise coupling, the noise covariance matrix may have a block diagonal structure [19, 20], which is, in effect, an sparse matrix.

This detection scheme has two steps. First, using a low-rank/sparse decomposition algorithm [21], the received signal sample covariance matrix is decomposed to the source and noise covariance matrices. Next, detection is performed using a heuristic detection statistics obtained from the decomposed matrices. Based on this detection statistics, we derive a semi-constant false alarm rate detector and approximate its probability of detection and false alarm. According to experimental results in a correlated noise environment, the proposed detector performs significantly better than usual GLRT detectors of [10], [12] which are designed for uncorrelated noise case. Specially, the gap between the performance of the proposed method and other detectors increases when the estimation of covariance matrix becomes more accurate; i.e., when the number of samples increases.

Spatially colored noise has been considered as a more practical assumption in other communications and signal processing problems too. For example, in [22], the effects of spatially correlated noise in the performance of direction-of-arrival estimation methods were investigated, in [23], based on this assumption, the outage probability of an optimal diversity receiver was studied, and, in [24], channel capacity of the multiple-input multiple-output (MIMO) system in the presence of spatially colored noise was derived.

This paper is organized as follows. After presenting the problem setting in Section 2, the proposed detection method is introduced in Section 3. Section 4 is devoted to the performance analysis of our algorithm. Section 5 provides some numerical simulations showing efficiency of the proposed method, and Section 6 concludes the paper.

2. System Model

We consider the problem of detecting the presence of primary users based on the signals received by an array of M sensors in a cognitive radio node. Each sensor collects N samples during the sensing time. Using these observed samples, the secondary user decides between two hypotheses \mathcal{H}_0 , \mathcal{H}_1 denoting the absence and the presence of the primary signal, respectively.

Let $\mathbf{y}(t) = (y_1(t), \dots, y_M(t))^T \in \mathbb{C}^{M \times 1}$ represent the vector of received samples at time instant t , where $y_i(t)$ is the sample of the i -th sensor. Furthermore, assume $\mathbf{s}(t) = (s_1(t), \dots, s_K(t))^T \in \mathbb{C}^{K \times 1}$ is the vector of transmitted signals from the sources at time t , in which $K < M$ denotes the number of active users. Assuming that the channel is flat fading and the propagation time of the received signals across the array is much smaller than the inverse of the signals

bandwidth, the received signals can be modeled as

$$\begin{cases} \mathcal{H}_0 : \mathbf{y}(t) = \mathbf{w}(t), & t = 1, \dots, N \\ \mathcal{H}_1 : \mathbf{y}(t) = \mathbf{H}\mathbf{s}(t) + \mathbf{w}(t), & t = 1, \dots, N \end{cases} \quad (1)$$

where $\mathbf{w}(t) = (w_1(t), \dots, w_M(t))^T \in \mathbb{C}^{M \times 1}$ is the vector of noise at different sensors at time instant t and $\mathbf{H} \in \mathbb{C}^{M \times K}$ is the channel matrix. Consequently, if the signals and noises are uncorrelated and zero-mean wide-sense stationary processes, we have

$$\begin{cases} \mathcal{H}_0 : \mathbf{R}_y = \mathbf{R}_w \\ \mathcal{H}_1 : \mathbf{R}_y = \mathbf{R}_x + \mathbf{R}_w \end{cases} \quad (2)$$

where $\mathbf{R}_x = \mathbf{H}\mathbf{R}_s\mathbf{H}^H$ and $\mathbf{R}_y, \mathbf{R}_s, \mathbf{R}_w$ are the covariance matrices of received signals, transmitted signals, and noises, respectively. Let $\hat{\mathbf{R}}_y \triangleq \frac{1}{N} \sum_{t=1}^N \mathbf{y}(t)\mathbf{y}^H(t)$, $\hat{\mathbf{R}}_s \triangleq \frac{1}{N} \sum_{t=1}^N \mathbf{s}(t)\mathbf{s}^H(t)$, $\hat{\mathbf{R}}_w \triangleq \frac{1}{N} \sum_{t=1}^N \mathbf{w}(t)\mathbf{w}^H(t)$ be the sample covariance matrices for \mathbf{y} , \mathbf{s} , \mathbf{w} , respectively. Similarly, $\hat{\mathbf{R}}_y = \hat{\mathbf{R}}_w$ and $\hat{\mathbf{R}}_y = \hat{\mathbf{R}}_x + \hat{\mathbf{R}}_w$ under \mathcal{H}_0 and \mathcal{H}_1 hypotheses, respectively, in which $\hat{\mathbf{R}}_x = \mathbf{H}\hat{\mathbf{R}}_s\mathbf{H}^H$. Assuming \mathbf{w} is zero mean Gaussian with covariance matrix \mathbf{R}_w , the sample covariance matrix $\hat{\mathbf{R}}_w$ would have central Wishart distribution with N degrees of freedom [25]. Define $\mathbf{Q} = \hat{\mathbf{R}}_w - \mathbf{R}_w$ as the disturbance of sample covariance matrix due to finite samples, N . Obviously, as $N \rightarrow \infty$, $\mathbf{Q} \rightarrow \mathbf{0}$.

In most of previous work, it is assumed that \mathbf{R}_w is diagonal; that is, noise is spatially white [7, 8]. In this paper, a more general and practical assumption is made on correlation between the noises. Mutual coupling between antennas of an array is usually a source of this correlation [26] which is a function of interelement spacing. Roughly speaking, the larger the distance between elements, the less mutual coupling, and, consequently, the less correlation between thermal noises. In this fashion, we consider an antenna array whose adjacent sensors have correlated noises, while the noises on non-adjacent sensors are uncorrelated or weakly correlated due to their longer distances. Therefore, \mathbf{R}_w has many zero elements, and, hence, is a sparse matrix with large values on or close to the diagonal. Furthermore, it is easy to show that $\text{rank}(\mathbf{R}_x) \leq K$, so if $K \ll M$, \mathbf{R}_x is a low-rank matrix. In this paper, we consider the case where \mathbf{R}_w is unknown with some non-zero off-diagonal entries.

3. The proposed method

A common solution for obtaining an asymptotically optimal detection statistics is to implement the GLRT [27]. However, for the spatially correlated noise case, the GLRT criterion leads to solving a non-convex optimization problem that, in general, cannot be easily solved (See Appendix A). To overcome this difficulty, we use a different approach to solve the detection problem. Prior to stating our method, we first make the following assumptions that will be used in the sequel.

Assumption 1. *For the detection problem given in (1) and simplified in (2), it is assumed that*

- (i) *The number of sources is unknown and much smaller than the number of sensors; i.e., $K \ll M$. Furthermore, the sources can be correlated or even coherent.*
- (ii) *The noises on the sensors are zero-mean Gaussian with cross-correlations on a few sensors. In other words, \mathbf{R}_w is non-zero on the diagonal and a few off-diagonal entries. Hence, \mathbf{R}_w is a sparse matrix, and we assume that its support, the set of non-zero entries, is known. In practice, using the information from the array geometry, we can assume that the noises on the adjacent sensors are correlated, while others are uncorrelated.*
- (iii) *The signals $\mathbf{s}(t)$ and noises $\mathbf{w}(t)$ are uncorrelated and zero-mean.*

Taking the above assumptions into account, under \mathcal{H}_1 hypothesis, the covariance matrix of the received signals is formed as the sum of a low-rank matrix \mathbf{R}_x and a sparse matrix \mathbf{R}_w . Consequently, we can exploit low-rank/sparse matrix decomposition algorithms [21] to decompose \mathbf{R}_x and \mathbf{R}_w from the \mathbf{R}_y matrix. Since we know the support of \mathbf{R}_w , this assumption reduces the decomposition problem to a matrix completion problem [28]. In matrix completion problem, the goal is to recover a low-rank matrix by observing a set of its entries. Using \mathbf{R}_y , we wish to recover the

low-rank matrix \mathbf{R}_x , whose entries are equal to \mathbf{R}_y outside of the support of \mathbf{R}_w . Let Ω be the set of all (i, j) entries with $[\mathbf{R}_w]_{ij} = 0$. Mathematically speaking, the rank minimization problem is formulated as

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \text{ s.t. } \mathcal{P}_\Omega(\mathbf{X}) = \mathcal{P}_\Omega(\mathbf{R}_y), \quad (3)$$

where \mathcal{P}_Ω is the projection onto the set Ω ; that is,

$$[\mathcal{P}_\Omega(\mathbf{X})]_{ij} = \begin{cases} [\mathbf{X}]_{ij} & (i, j) \in \Omega \\ 0 & o.w. \end{cases}$$

In general, the rank minimization problem (3) is NP-hard and nonconvex [29]. A well-known convex relaxation for this problem is replacing the $\text{rank}(\mathbf{X})$ with the nuclear norm, $\|\mathbf{X}\|_* = \sum_{i=1}^M \sigma_i(\mathbf{X})$, leading to [30]:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \text{ s.t. } \mathcal{P}_\Omega(\mathbf{X}) = \mathcal{P}_\Omega(\mathbf{R}_y). \quad (4)$$

Under some mild conditions, problems (3) and (4) will meet the same solutions with overwhelming probability [30].¹ However, in practice, only an estimation of \mathbf{R}_y , i.e. $\hat{\mathbf{R}}_y$, is available, where $\hat{\mathbf{R}}_y = \mathbf{R}_y + \mathbf{Q}$, and \mathbf{Q} is the disturbance term due to the finite number of samples. To alleviate the effect of finite number of samples, we use the following program:

$$\tilde{\mathbf{R}}_x = \arg \min_{\mathbf{X}} \left\{ \mu \|\mathbf{X}\|_* + \|\mathcal{P}_\Omega(\mathbf{X}) - \mathcal{P}_\Omega(\hat{\mathbf{R}}_y)\|_F \mid \mathbf{X} \succeq \mathbf{0} \right\}, \quad (5)$$

in which $\mathbf{X} \succeq \mathbf{0}$ indicates that \mathbf{X} is a positive semidefinite matrix, $\|\cdot\|_F$ denotes the Frobenius norm, and $\mu > 0$ is a regularization parameter. The decomposed noise covariance matrix $\tilde{\mathbf{R}}_w$ is then obtained by $\tilde{\mathbf{R}}_w = \hat{\mathbf{R}}_y - \tilde{\mathbf{R}}_x$. Notice that in (5), contrary to the common formulation (the regular LASSO approach [31]), the data fidelity term ($\|\mathcal{P}_\Omega(\mathbf{X}) - \mathcal{P}_\Omega(\hat{\mathbf{R}}_y)\|_F$) is not squared. This allows us to select the regularization parameter, similar to the square-root LASSO approach [32], independent from the scaling of the covariance of \mathbf{Q} . Although the square-root LASSO approach has a slightly lower performance than the regular LASSO [32, 33], having a fixed regularization parameter makes the matrix decomposition step less vulnerable to the disturbance term \mathbf{Q} . As a result, the decomposition step becomes less dependent on the SNR value, which is desirable for our spectrum sensing algorithm, as the SNR value is assumed to be unknown. The optimization problem (5) can be converted to a semidefinite programming (SDP) as shown in Appendix C; thus, it can be solved using efficient solvers like SDPT3 [34] and SeDuMi [35].

In the case that the support of \mathbf{R}_w is unknown to us, the following program is applicable:

$$(\tilde{\mathbf{R}}_x, \tilde{\mathbf{R}}_w) = \arg \min_{(\mathbf{L}, \mathbf{S})} \left\{ \mu_1 \|\mathbf{L}\|_* + \mu_2 \|\mathbf{S}\|_1 + \|\hat{\mathbf{R}}_y - \mathbf{L} - \mathbf{S}\|_F \mid \mathbf{L} \succeq \mathbf{0}, \mathbf{S} \succ \mathbf{0} \right\}, \quad (6)$$

in which $\mu_1 > 0$ and $\mu_2 > 0$ are regularization parameters, $\mathbf{S} \succ \mathbf{0}$ means that \mathbf{S} is positive definite, and $\|\mathbf{S}\|_1 = \sum_{i,j} |[\mathbf{S}]_{ij}|$. Following the same argument as in Appendix C, the program (6) can be formulated as an SDP, as well. Since the programs (5) and (6) are convex, they can be solved using conventional convex optimization tools like CVX [36].

Now, using the decomposed covariance matrices $\tilde{\mathbf{R}}_x$ and $\tilde{\mathbf{R}}_w$, we can detect the presence of the primary signal. If the signal is absent, $\mathbf{R}_x = \mathbf{0}$; thus, the received power on the sensors is completely due to the noise term. In this case, the diagonal entries of \mathbf{R}_y correspond to the noise power on each sensor. On the other hand, if the signal is present, \mathbf{R}_x is a non-zero matrix with positive entries on the diagonal, representing the received power of the sources on each sensor. Taking these facts into account, a heuristic detection rule can be formulated by comparing the ratio of the total received power of the sources to the total received power. More specifically, we perform the following test

$$T = \frac{\text{tr}(\tilde{\mathbf{R}}_x)}{\text{tr}(\hat{\mathbf{R}}_y)} \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\geq}} \gamma, \quad (7)$$

¹Though \mathcal{P}_Ω does not satisfy the conditions in [30, Theorem 7], our numerical simulations confirm that (3) and (4) share the same solutions.

where $\text{tr}(\cdot)$ designates the trace of a matrix and γ is the detection threshold computed approximately in Section 4.1 or can be computed using numerical simulations. If matrix decomposition could be done perfectly, i.e. $\tilde{\mathbf{R}}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}$, T would become zero under \mathcal{H}_0 hypothesis, and a non-zero value under \mathcal{H}_1 . Therefore, we could detect the primary signal perfectly in almost any SNR value. However, in practice, the disturbance matrix \mathbf{Q} , caused by the finite number of samples, acts as an additive noise and degrades the performance of the low-rank/sparse matrix decomposition algorithm. For instance, when the primary signal is not present, the output of the decomposition algorithm, $\tilde{\mathbf{R}}_{\mathbf{x}}$, is not equal to $\mathbf{0}$ and has many nonzero but small entries.

It is also possible to introduce other detection tests on the $\tilde{\mathbf{R}}_{\mathbf{x}}$ matrix, which can perform similar to (7). For instance, one can use $\|\tilde{\mathbf{R}}_{\mathbf{x}}\|_F/\|\hat{\mathbf{R}}_{\mathbf{y}}\|_F$ or $\lambda_{\max}(\tilde{\mathbf{R}}_{\mathbf{x}})/\lambda_{\min}(\hat{\mathbf{R}}_{\mathbf{y}})$ as detection statistics, where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ represent the largest and smallest eigenvalues of a matrix. However, numerical simulations reveal that the mentioned ratio detectors have similar performances (see Section 5). In the sequel, we derive an approximate performance analysis for the trace ratio test (7).

4. Performance Analysis

The performance of the sensing algorithm depends on the selected threshold value γ . In this section, we compute the threshold value in a way that the probability of false alarm is constant with the scaling of the noise covariance matrix. This criterion is referred to as Semi-Constant False Alarm Rate (S-CFAR) in the sequel. Let $p_d \triangleq P(T > \gamma|\mathcal{H}_1)$ be the probability of detection and $p_f \triangleq P(T > \gamma|\mathcal{H}_0)$ be the probability of false alarm. We determine p_f as a function of γ .

Determining the distribution of T under both \mathcal{H}_0 and \mathcal{H}_1 hypotheses seems to be very tricky because $\tilde{\mathbf{R}}_{\mathbf{w}}$ and $\tilde{\mathbf{R}}_{\mathbf{x}}$ are the outputs of a highly nonlinear matrix decomposition algorithm. Nevertheless, a naive approach is to simply run a large number of simulations to extract the best γ value to gain the given p_f . To ease the burden of computations, we derive some approximations for the threshold value using some simplifying assumptions about the matrix decomposition procedure.

In the following subsections, we approximate a threshold for the S-CFAR. Then, to characterize the effectiveness of the detector, we study its performance in two cases. First, considering the imperfection of the decomposition algorithm, we approximate the probability of detection of the proposed detector. Second, neglecting this imperfection (i.e., assuming we are using an ideal decomposition algorithm), we derive the exact probability of detection which gives an insight to the effect of using an actual decomposition algorithm.

Since $\mathbf{R}_{\mathbf{w}}$ is unknown to the secondary users but is needed to determine the threshold value, in the sequel, we assume that the sensing algorithm estimates $\mathbf{R}_{\mathbf{w}}$. $\mathbf{R}_{\mathbf{w}}$ can be estimated using the results from the matrix decomposition. Furthermore, a simple but more accurate approach is to average $\tilde{\mathbf{R}}_{\mathbf{w}}$ in the few recent times that the spectrum sensing algorithm has been executed, assuming that the noise properties do not change rapidly.

4.1. An approximation for the threshold

Under \mathcal{H}_0 hypothesis, $\mathbf{R}_{\mathbf{x}} = \mathbf{0}$ and $\hat{\mathbf{R}}_{\mathbf{y}} = \mathbf{R}_{\mathbf{w}} + \mathbf{Q}$. Throughout numerical simulations, we observed that due to absence of the low-rank matrix, \mathbf{Q} leaks to the low-rank output of the matrix decomposition algorithm. Accordingly, a simple approximation would be to replace $\tilde{\mathbf{R}}_{\mathbf{x}}$ with \mathbf{Q} and analyze its distribution instead. In this case, we assume that the sparse matrix $\tilde{\mathbf{R}}_{\mathbf{w}}$ is exactly recovered. Although it is not theoretically provable, numerical simulations show that the simplified distribution function looks similar to the original one. Moreover, the error caused by approximation of the threshold value γ is negligible. In summary, we analyze the following random variable

$$T' = \frac{|\text{tr}(\mathbf{Q})|}{\text{tr}(\hat{\mathbf{R}}_{\mathbf{y}})},$$

as an approximation for $T|\mathcal{H}_0$ to determine the probability of false alarm, p_f . Figure 1 shows that $T|\mathcal{H}_0$ is closely approximated by T' , under simulation conditions defined in Section 5, with $\mathbf{R}_{\mathbf{w}} = \mathbf{R}_{\mathbf{w}}^{(1)}$ and $N = 100$.

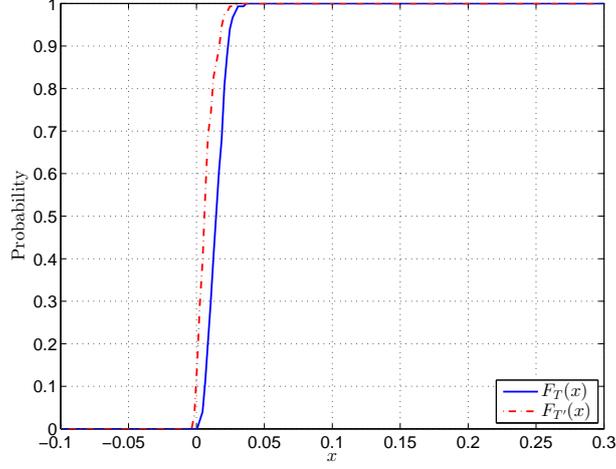


Figure 1: Comparison between empirical CDF of $F_T(x)$ and $F_{T'}(x)$. The simulation is done for $N = 2000$ and other parameters described in Section 5, with $\mathbf{R}_w = \mathbf{R}_w^{(1)}$.

The probability of false alarm is

$$\begin{aligned} p_f &= P(T > \gamma | \mathcal{H}_0) \approx P(T' > \gamma | \mathcal{H}_0) \\ &= P\left(\frac{|\text{tr}(\hat{\mathbf{R}}_w - \mathbf{R}_w)|}{\text{tr}(\hat{\mathbf{R}}_w)} > \gamma\right). \end{aligned}$$

Let $P(\text{tr}(\hat{\mathbf{R}}_w) - \text{tr}(\mathbf{R}_w) > 0) = 1 - P(\text{tr}(\hat{\mathbf{R}}_w) - \text{tr}(\mathbf{R}_w) < 0) = \eta$. After simple algebraic manipulations, we get

$$\eta \left(1 - F\left(\frac{\text{tr}(\mathbf{R}_w)}{1 - \gamma}\right)\right) + (1 - \eta) F\left(\frac{\text{tr}(\mathbf{R}_w)}{1 + \gamma}\right) - p_f = 0, \quad (8)$$

where $F(z) = P(\text{tr}(\hat{\mathbf{R}}_w) \leq z)$ is cumulative density function (CDF) of trace of a Wishart matrix approximated using equation (B.2) in Appendix B, and we have $\eta = 1 - F(\text{tr}(\mathbf{R}_w))$. γ can be computed by solving equation (8) using numerical methods. It can be easily verified that the probability of false alarm is constant over scaling of the noise covariance matrix \mathbf{R}_w .

4.2. Detection probability

When a primary source is present, determining the exact distribution function of T is difficult, even in the case of uncorrelated white noise. However, in practice, we observed that due to non-ideality of the matrix decomposition algorithm and the presence of matrix \mathbf{Q} acting as an additive noise, the matrix \mathbf{Q} partly leaks into $\tilde{\mathbf{R}}_x$. In other words, $\tilde{\mathbf{R}}_x | \mathcal{H}_1$ is highly correlated to $\mathbf{R}_x + \alpha \mathbf{Q}$, where $0 < \alpha \leq 1$ is a constant representing the portion of leakage of \mathbf{Q} into $\tilde{\mathbf{R}}_x$. A rule of thumb approximation for this imperfection of the decomposition step is to set $\alpha = 1$; that is, we assume that \mathbf{Q} is mainly absorbed by $\tilde{\mathbf{R}}_x$. The exact value of α generally depends on the aspects of the problem and the algorithm used in the decomposition process.

Following the above reasoning, we analyze the following random variable

$$T'' = \frac{\text{tr}(\mathbf{R}_x + \mathbf{Q})}{\text{tr}(\hat{\mathbf{R}}_y)}$$

as an approximation of $T|\mathcal{H}_1$. This gives

$$\begin{aligned} p_d &= P(T > \gamma|\mathcal{H}_1) \simeq P(T'' > \gamma|\mathcal{H}_1) \\ &= P\left(\frac{\text{tr}(\mathbf{R}_x + \hat{\mathbf{R}}_w - \mathbf{R}_w)}{\text{tr}(\hat{\mathbf{R}}_v)} > \gamma\right) \\ &= P\left(\text{tr}(\hat{\mathbf{R}}_w) > \frac{\text{tr}(\mathbf{R}_w)}{1 - \gamma} - \text{tr}(\mathbf{R}_x)\right). \end{aligned}$$

Thus,

$$p_d \simeq 1 - F\left(\frac{\text{tr}(\mathbf{R}_w)}{1 - \gamma} - \text{tr}(\mathbf{R}_x)\right), \quad (9)$$

where $F(z)$ is the CDF function of trace of a Wishart matrix approximated using equation (B.2) in Appendix B.

In addition to the above approximation, we can find the probability of detection p_d^* , for an ideal case that the decomposition is done perfectly. Substituting $\hat{\mathbf{R}}_x$ by \mathbf{R}_x , we can write

$$p_d^* = P\left(\frac{\text{tr}(\mathbf{R}_x)}{\text{tr}(\mathbf{R}_x + \hat{\mathbf{R}}_w)} > \gamma\right).$$

Thus,

$$p_d^* = F\left(\frac{1 - \gamma}{\gamma} \text{tr}(\mathbf{R}_x)\right). \quad (10)$$

Although this is not the case that happens in practice, by improving the low-rank/sparse decomposition algorithm and increasing its robustness to noise, the actual p_d can approach to the ideal p_d^* .

5. Experimental Results

The proposed spectrum sensing algorithm is simulated with $M = 6, K = 1$. We consider a uniform linear array of antennas with interelement spacing equal to one third wavelength of the transmitted signal. The transmitted signal is independently and identically distributed realizations of a complex zero-mean Gaussian distribution. We perform the simulations with the following noise covariance matrices:

$$\mathbf{R}_w^{(1)} = \begin{bmatrix} 0.7 & 0.2 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.2 & 1.9 & 0.5 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.5 & 1.2 & -0.4j & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.4j & 1.4 & 0.2 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.2 & 0.5 & 0.4 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.4 & 1.3 \end{bmatrix}$$

$$\mathbf{R}_w^{(2)} = \begin{bmatrix} 0.8 & 0.0 & 0.5 & 0.0 & 0.0 & -0.6j \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.5j \\ 0.5 & 0.0 & 1.3 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.8 & 0.4j & 0.0 \\ 0.0 & 0.0 & 0.0 & -0.4j & 1.1 & 0.0 \\ 0.6j & -0.5j & 0.0 & 0.0 & 0.0 & 1.2 \end{bmatrix}.$$

The first choice corresponds to the case that neighboring antennas have spatially colored noises. By the second choice, it is shown that the proposed algorithm can still work efficiently with arbitrary but sparse \mathbf{R}_w 's. If the support of \mathbf{R}_w is known, an assumption in these simulations, we can use matrix completion methods to recover \mathbf{R}_w which, indeed, leads to more accurate estimates. However, we can also exploit low-rank/sparse decomposition algorithms to recover \mathbf{R}_w when its support is unknown. We use the CVX library [36] to decompose $\hat{\mathbf{R}}_x$ by program (5). We use a fixed regularization parameter $\mu = 1.1\|\tilde{\mathbf{s}}\|_\infty \sqrt{M^2 - |\Omega|}$, where $\tilde{\mathbf{s}}$ is a constant vector defined in [32] and is computed by a simple numerical simulation. We compare our algorithm with some well-known detection algorithms

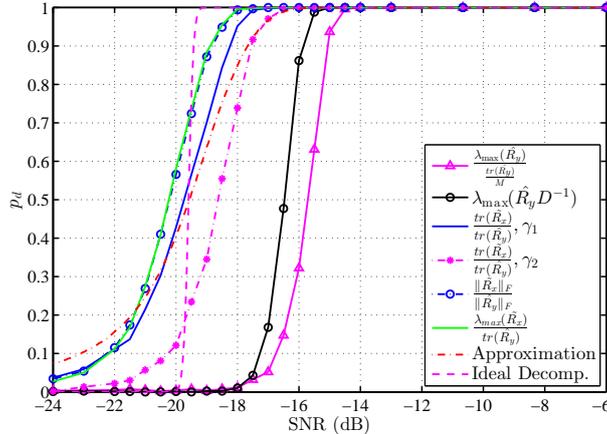


Figure 2: p_d of each detection algorithm versus SNR(dB) for $N = 500$, $\mathbf{R}_w = \mathbf{R}_w^{(1)}$ and other parameters as described in section 5. “ $\text{tr}(\hat{\mathbf{R}}_x)/\text{tr}(\hat{\mathbf{R}}_y), \gamma_1$ ” denotes p_d of our algorithm with the threshold calculated empirically with $p_f = 0.01$, while “ $\text{tr}(\hat{\mathbf{R}}_x)/\text{tr}(\hat{\mathbf{R}}_y), \gamma_2$ ” stands for p_d of our algorithm with the threshold computed using (8) with the same p_f . “Approximation” and “Ideal Decomp.” indicate the probabilities of detection calculated using equations (9) and (10), respectively.

for uncorrelated noise case assuming one transmitter. $\lambda_{\max}(\hat{\mathbf{R}}_y)/(\text{tr}(\hat{\mathbf{R}}_y)/M)$, called Scaled Largest Eigenvalue (SLE) [10], is suitable when $\mathbf{R}_w = \sigma^2 \mathbf{I}$, and $\lambda_{\max}(\hat{\mathbf{R}}_y \mathbf{D}^{-1})$ with $\mathbf{D} = \text{diag}([\hat{\mathbf{R}}_y]_{11}, [\hat{\mathbf{R}}_y]_{22}, \dots, [\hat{\mathbf{R}}_y]_{MM})$ [12] is applicable when $\mathbf{R}_w = \text{diag}(\sigma_1^2, \dots, \sigma_M^2)$.

Based on 1000 Monte-Carlo simulations, for each SNR, which is defined as $10 \log_{10}(\text{tr}(\hat{\mathbf{R}}_x)/\text{tr}(\hat{\mathbf{R}}_w))$, the probability of detection p_d is calculated for each algorithm. The results are shown in Figures 2 to 5 for different \mathbf{R}_w and N values. The thresholds are computed empirically assuming $p_f = 0.01$, except for the plots indicated by “ $\text{tr}(\hat{\mathbf{R}}_x)/\text{tr}(\hat{\mathbf{R}}_y), \gamma_2$ ” in which the threshold is calculated using (8). For larger values of N (as in Figures 3 and 5), our algorithm performs much better since finite sampling effect decreases and matrix decomposition works with higher performance. For example, as depicted in Figure 5, for $N = 2000$, and $\mathbf{R}_w = \mathbf{R}_w^{(2)}$ our method detects the presence of the primary signal with SNR 6 dB lower than the other methods. The approximation for the probability of detection is calculated from (9), and the probability of detection for the ideal case is obtained using (10). As illustrated in Figures 2 to 5, the approximation of p_d for our algorithm is close to the p_d obtained from numerical simulations. Also, the performance decay is negligible when the approximated threshold γ_2 is used instead of the empirical threshold γ_1 .

Finally, using equations (8) and (9) the receiver operational characteristic (ROC) curves can be calculated approximately. Figures 6 and 7 depict the approximated ROC curves for $N = 500$, and $N = 2000$, where $\mathbf{R}_w = \mathbf{R}_w^{(2)}$ in all plots.

6. Conclusion

In this paper, we considered spectrum sensing for CR nodes equipped with an array of sensors in spatially correlated noise environments with unknown noise covariance. We proposed a detection algorithm based on low-rank/sparse decomposition of the sample covariance matrix of the received signals. We evaluated the performance of the proposed detection algorithm using approximately derived probabilities of detection and false alarm. The simulation results showed that in a correlated noise environment, the proposed detector performs much better than the GLRT detectors for uncorrelated noise case.

Appendix A. Deriving GLRT detector for spatially correlated noise case

We consider the case that $K = 1$; i.e., there is only one transmitter. Let $\mathbf{s}(t) \sim \mathcal{N}(0, p)$ be the transmitted signal from the source and $\mathbf{w}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_w)$ be the noise on the sensors. The received signal is $\mathbf{y}(t) = s(t)\mathbf{a}(\theta) + \mathbf{w}(t)$, where

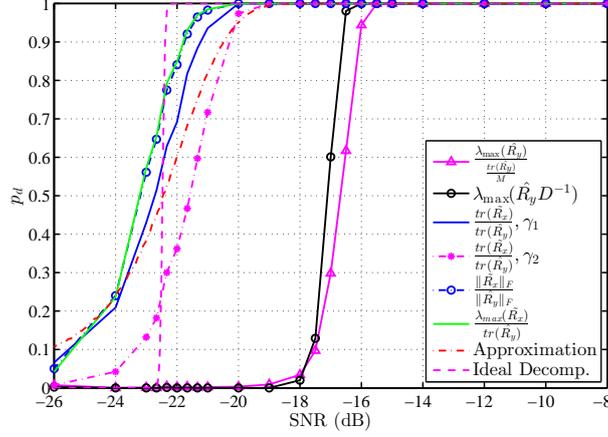


Figure 3: p_d of each detection algorithm versus SNR (dB) under the same conditions as of Figure 2 except for $N = 2000$.

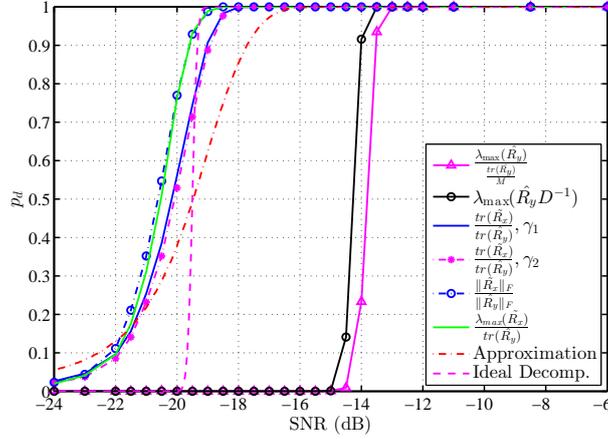


Figure 4: p_d of each detection algorithm versus SNR (dB) under the same conditions as of Figure 2 except for $\mathbf{R}_w = \mathbf{R}_w^{(2)}$, and $N = 500$.

$\mathbf{a}(\theta) = (a_1(\theta), \dots, a_M(\theta))^T$ is the vector of complex gains of sensors in the direction θ and θ is the direction of arrival of the source. Let $\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(N)]$. Under \mathcal{H}_1 , the likelihood function is

$$f(\mathbf{y}|\mathcal{H}_1, \mathbf{R}_w, \theta, p) \propto \frac{1}{|\mathbf{R}_y|^{\frac{1}{2}}} e^{-\frac{1}{2} \mathbf{y}^T \mathbf{R}_y^{-1} \mathbf{y}}$$

$$\Rightarrow f(\mathbf{Y}|\mathcal{H}_1, \mathbf{R}_w, \theta, p) \propto \frac{1}{|\mathbf{R}_y|^{\frac{N}{2}}} e^{-\frac{1}{2} \text{tr}(\mathbf{R}_y^{-1} \mathbf{Y})},$$

where \mathbf{R}_w and $\mathbf{R}_y = p\mathbf{a}(\theta)\mathbf{a}(\theta)^T + \mathbf{R}_w$ are noise and received signal covariance matrices, respectively. Likewise, under \mathcal{H}_0 , we have

$$f(\mathbf{Y}|\mathcal{H}_0, \mathbf{R}_w) \propto \frac{1}{|\mathbf{R}_w|^{\frac{N}{2}}} e^{-\frac{1}{2} \text{tr}(\mathbf{R}_w^{-1} \mathbf{Y})}.$$

The generalized likelihood ratio is then

$$\mathcal{L} = \frac{\max_{\mathbf{R}_w, \theta, p} \frac{1}{|\mathbf{R}_y|^{\frac{N}{2}}} e^{-\frac{1}{2} \text{tr}(\mathbf{R}_y^{-1} \mathbf{Y})}}{\max_{\mathbf{R}_w} \frac{1}{|\mathbf{R}_w|^{\frac{N}{2}}} e^{-\frac{1}{2} \text{tr}(\mathbf{R}_w^{-1} \mathbf{Y})}}.$$

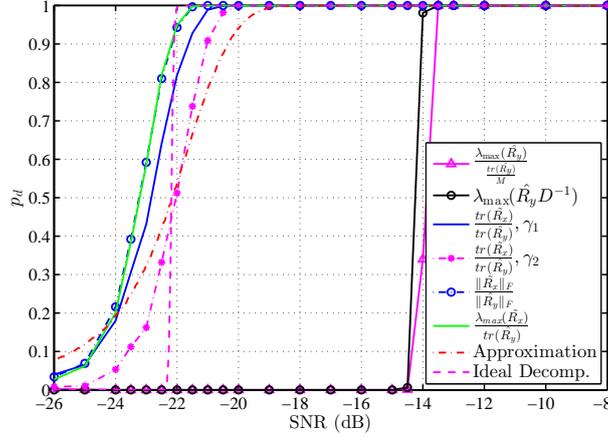


Figure 5: p_d of each detection algorithm versus SNR (dB) under the same conditions as of Figure 4 except for $\mathbf{R}_w = \mathbf{R}_w^{(2)}$, and $N = 2000$.

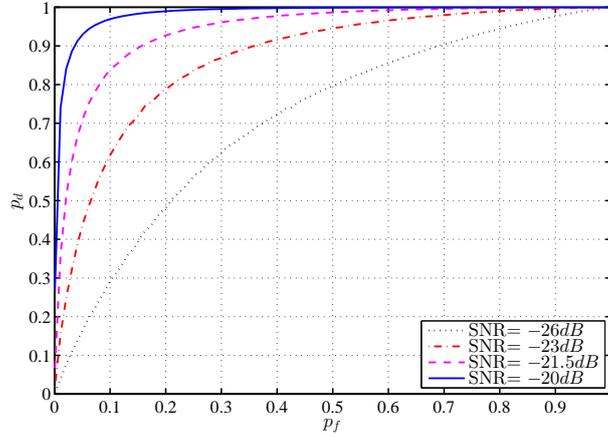


Figure 6: The approximated ROC for different SNR values with $\mathbf{R}_w = \mathbf{R}_w^{(2)}$ and $N = 500$.

It is easy to verify that both optimization problems in numerator and denominator are nonconvex. Furthermore, they become intractable for a general \mathbf{R}_w .

Appendix B. Distribution of the trace of a Wishart matrix

The exact distribution of $\text{tr}(\hat{\mathbf{R}}_w)$, the trace of a Wishart matrix, in a general case, can be computed using zonal polynomials [37]. However, these polynomials are hard to evaluate specially for large values of M or N . On the other hand, $\text{tr}(\hat{\mathbf{R}}_w)$ equals to a linear combination of chi-square random variables [38]; that is,

$$\text{tr}(\hat{\mathbf{R}}_w) \sim \sum_{i=1}^M \lambda_i \chi_N^2,$$

where λ_i is the i th eigenvalues of \mathbf{R}_w . The distribution of linear combination of chi-square random variables lacks a simple closed form expression [39]. Nevertheless, a simple approximate distribution can be obtained by constructing a random variable with the first three moments equal to the original random variable. Using the results of [40], the probability density function (PDF) of the approximating random variable is given by

$$f(z) = \frac{N}{2\Lambda} \sum_{i=1}^M \frac{\left(\frac{Nz}{2\lambda_i}\right)^{\frac{N\Lambda}{2\lambda_i} - 1} e^{-\frac{Nz}{2\lambda_i}}}{\Gamma\left(\frac{N\Lambda}{2\lambda_i}\right)}, \quad (\text{B.1})$$

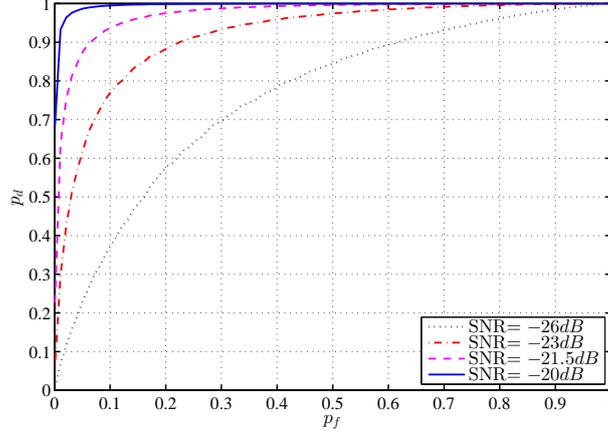


Figure 7: The approximated ROC for different SNR values with $\mathbf{R}_w = \mathbf{R}_w^{(2)}$ and $N = 2000$.

where $\Lambda \triangleq \sum_{i=1}^M \lambda_i$, $z \geq 0$, and $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is the *gamma function*. Thus, the CDF is

$$F(z) = \int_0^z f(z') dz' = \frac{1}{\Lambda} \sum_{i=1}^M \frac{\lambda_i \gamma\left(\frac{N\Lambda}{2\lambda_i}, \frac{Nz}{2\lambda_i}\right)}{\Gamma\left(\frac{N\Lambda}{2\lambda_i}\right)}, \quad (\text{B.2})$$

where $\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt$ is the *lower incomplete gamma function*.

Appendix C. Formulation of (5) as an SDP problem

Let $\mathbf{r} = \text{vec}(\mathcal{P}_\Omega(\mathbf{R}_y - \mathbf{X}))$, where $\text{vec}(\mathbf{A})$ denotes a long vector obtained by stacking the columns of \mathbf{A} . We can rewrite (5) as

$$\tilde{\mathbf{R}}_x = \arg \min_{\mathbf{X}, t} \left\{ \mu \text{tr}(\mathbf{X}) + t \mid \mathbf{X} \geq \mathbf{0}, \|\mathbf{r}\|_2 \leq t \right\},$$

where $\|\mathbf{r}\|_2$ denotes the ℓ_2 norm of \mathbf{r} , and t is an optimization variable. Moreover, using a Schur complement argument, the constraint $\|\mathbf{r}\|_2 \leq t$ is equivalent to

$$\begin{pmatrix} t\mathbf{I}_{M^2} & \mathbf{r} \\ \mathbf{r}^H & t \end{pmatrix} \geq \mathbf{0}.$$

Therefore, (5) can be reformulated as the following SDP program:

$$\tilde{\mathbf{R}}_x = \arg \min_{\mathbf{X}, t} \left\{ \mu \text{tr}(\mathbf{X}) + t \mid \begin{pmatrix} \mathbf{X} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & t\mathbf{I}_{M^2} & \mathbf{r} \\ \mathbf{0} & \mathbf{r}^H & t \end{pmatrix} \geq \mathbf{0} \right\}$$

Acknowledgment

The authors wish to thank Dr. Joakim Jaldén, Dr. Aliakbar Tadaion, and anonymous reviewers for their helpful comments.

References

- [1] J. Mitola, M. G., Cognitive radio: making software radios more personal, *IEEE Personal Communications* 6 (4) (1999) 13–18.
- [2] H. Urkowitz, Energy detection of unknown deterministic signals, *Proceedings of the IEEE* 55 (4) (1967) 523–531.
- [3] A. Sonnenschein, P. M. Fishman, Radiometric detection of spread-spectrum signals in noise of uncertain power, *IEEE Transactions on Aerospace and Electronic Systems* 28 (3) (1992) 654–660.

- [4] A. Sahai, D. Cabric, Spectrum sensing: fundamental limits and practical challenges, in: Proc. IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2005.
- [5] W. Gardner, Spectral correlation of modulated signals: Part i-analog modulation, *IEEE Transactions on Communications* 35 (6) (1987) 584–594.
- [6] P. D. Sutton, K. E. Nolan, L. E. Doyle, Cyclostationary signatures in practical cognitive radio applications, *IEEE Journal on Selected Areas in Communications* 26 (1) (2008) 13–24.
- [7] P. Wang, J. Fang, N. Han, H. Li, Multiantenna-assisted spectrum sensing for cognitive radio, *IEEE Transactions on Vehicular Technology* 59 (4) (2010) 1791–1800.
- [8] Y. Zeng, Y. Liang, Maximum-minimum eigenvalue detection for cognitive radio, in: *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2007, IEEE, 2007, pp. 1–5.
- [9] J. Manco-Vázquez, M. Lázaro-Gredilla, D. Ramírez, J. Vía, I. Santamaría, A bayesian approach for adaptive multiantenna sensing in cognitive radio networks, *Signal Processing* 96 (2014) 228–240.
- [10] O. Besson, S. Kraut, L. L. Scharf, Detection of an unknown rank-one component in white noise, *IEEE Transactions on Signal Processing* 54 (7) (2006) 2835–2839.
- [11] R. López-Valcarce, G. Vazquez-Vilar, J. Sala, Multiantenna spectrum sensing for cognitive radio: overcoming noise uncertainty, in: *International Workshop on Cognitive Information Processing (CIP)*, IEEE, 2010, pp. 310–315.
- [12] D. Ramírez, J. Vía, I. Santamaría, Multiantenna spectrum sensing: The case of wideband rank-one primary signals, in: *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, IEEE, 2010, pp. 9–12.
- [13] I. Gupta, A. Ksienski, Effect of mutual coupling on the performance of adaptive arrays, *IEEE Transactions on Antennas and Propagation* 31 (5) (1983) 785–791.
- [14] H. Steyskal, J. S. Herd, Mutual coupling compensation in small array antennas, *IEEE Transactions on Antennas and Propagation* 38 (12) (1990) 1971–1975.
- [15] J. W. Wallace, M. A. Jensen, Mutual coupling in MIMO wireless systems: A rigorous network theory analysis, *IEEE Transactions on Wireless Communications* 3 (4) (2004) 1317–1325.
- [16] T. Svantesson, A. Ranheim, Mutual coupling effects on the capacity of multielement antenna systems, in: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 4, 2001, pp. 2485–2488.
- [17] J. Sala-Alvarez, G. Vazquez-Vilar, R. López-Valcarce, Multiantenna GLR detection of rank-one signals with known power spectrum in white noise with unknown spatial correlation, *IEEE Transactions on Signal Processing* 60 (6) (2012) 3065–3078.
- [18] K. Jitvanichphaibool, Y.-C. Liang, Y. Zeng, Spectrum sensing using multiple antennas for spatially and temporally correlated noise environments, in: *IEEE Symposium on New Frontiers in Dynamic Spectrum*, IEEE, 2010, pp. 1–7.
- [19] S. Vorobyov, A. Gershman, K. Wong, Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays, *IEEE Trans. Signal Process.* 53 (1) (2005) 34–43.
- [20] T. Li, A. Nehorai, Maximum likelihood direction finding in spatially colored noise fields using sparse sensor arrays, *IEEE Trans. on Signal Processing* 59 (3) (2011) 1048–1062.
- [21] E. J. Candès, X. Li, Y. Ma, J. Wright, Robust principal component analysis?, *Journal of the ACM (JACM)* 58 (3) (2011) 11.
- [22] F. Li, R. J. Vaccaro, Performance degradation of DOA estimators due to unknown noise fields, in: *IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, 1991, pp. 1413–1416.
- [23] C. P. Domizioli, B. L. Hughes, K. G. Gard, G. Lazzi, Noise correlation in compact diversity receivers, *IEEE Transactions on Communications* 58 (5) (2010) 1426–1436.
- [24] S. Krusevac, P. Rapajic, R. A. Kennedy, Channel capacity estimation for MIMO systems with correlated noise, in: *IEEE Global Telecommunications Conference*, Vol. 5, 2005, pp. 2812–2816.
- [25] M. L. Eaton, *Multivariate statistics: a vector space approach*, Wiley New York, 1983.
- [26] A. Kislansky, R. Shavit, J. Tabrikian, Direction of arrival estimation in the presence of noise coupling in antenna arrays, *IEEE Transactions on Antennas and Propagation* 55 (7) (2007) 1940–1947.
- [27] O. Zeitouni, J. Ziv, N. Merhav, When is the generalized likelihood ratio test optimal?, *IEEE Transactions on Information Theory* 38 (5) (1992) 1597–1602.
- [28] E. J. Candès, B. Recht, Exact matrix completion via convex optimization, *Foundations of Computational Mathematics* 9 (6) (2009) 717–772.
- [29] B. Recht, M. Fazel, P. A. Parrilo, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization, *SIAM Rev.* 55 (2010) 471–501.
- [30] E. J. Candès, Y. Plan, Matrix completion with noise, *Proceedings of the IEEE* 98 (6) (2010) 925–936.
- [31] R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)* (1996) 267–288.
- [32] A. Belloni, V. Chernozhukov, L. Wang, Square-root lasso: pivotal recovery of sparse signals via conic programming, *Biometrika* 98 (4) (2011) 791–806.
- [33] O. Klopp, et al., Noisy low-rank matrix completion with general sampling distribution, *Bernoulli* 20 (1) (2014) 282–303.
- [34] K.-C. Toh, M. J. Todd, R. H. Tütüncü, Sdpt3a matlab software package for semidefinite programming, version 1.3, *Optimization methods and software* 11 (1-4) (1999) 545–581.
- [35] I. Polik, T. Terlaky, Y. Zinchenko, Sedumi: a package for conic optimization, in: *IMA workshop on Optimization and Control*, Univ. Minnesota, Minneapolis, 2007.
- [36] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.0 beta, <http://cvxr.com/cvx> (2012).
- [37] R. I. Murihead, *Aspects of Multivariate Statistical Theory*, Wiley Interscience, 1981.
- [38] S. Kourouklis, P. G. Moschopoulos, On the distribution of the trace of a noncentral wishart matrix, *Metron* 43 (2) (1985) 85–92.
- [39] A. Castaño-Martínez, F. López-Blázquez, Distribution of a sum of weighted central chi-square variables, *Communications in Statistics Theory and Methods* 34 (3) (2005) 515–524.
- [40] S. Gabler, C. Wolff, A quick and easy approximation to the distribution of a sum of weighted chi-square variables, *Statistical papers* 28 (1)

(1987) 317-325.