DETECTION WITH MULTIMODAL DEPENDENT DATA USING LOW DIMENSIONAL RANDOM PROJECTIONS

Thakshila Wimalajeewa and Pramod K. Varshney

Syracuse University, Syracuse, NY USA

ABSTRACT

Performing likelihood ratio based detection with high dimensional multimodal data is a challenging problem since the computation of the joint probability density functions (pdfs) in the presence of inter-modal dependence is difficult. While some computationally expensive approaches have been proposed for dependent multimodal data fusion (e.g., based on copula theory), a commonly used tractable approach is to compute the joint pdf as the product of marginal pdfs ignoring dependence. However, this method leads to poor performance when the data is strongly dependent. In this paper, we consider the problem of detection when dependence among multimodal data is modeled in a compressed domain where compression is obtained using low dimensional random projections. We employ a Gaussian approximation while modeling inter-modal dependence in the compressed domain which is computationally more efficient. We show that, under certain conditions, detection with multimodal dependent data in the compressed domain with a small number of compressed measurements yields enhanced performance compared to detection with high dimensional data via either the product approach or other suboptimal fusion approaches proposed in the literature.

Index terms: Compressive sensing, multimodal data, inter-modal dependence, likelihood ratio based detection, copula theory

1. INTRODUCTION

Fusion of high dimensional heterogenous data for different inference problems is challenging in many applications [1]. While likelihood ratio (LR) based detection (with no unknown parameters) is optimal in the Bayesian setting, its optimality is not guaranteed when the exact joint probability density function (pdf) is not available. It is difficult to compute the joint pdf in the presence of multimodal dependence unless data can be modeled as Gaussian. To model complex dependencies among multivariate data in order to compute the joint pdf, copula theory has been used in [2–8]. While there are several copula density functions available in the literature, finding the best copula function that fits a given set of data is computationally challenging. Further, in order to fuse multimodal data with more than two modalities, finding multivariate copula density functions is another challenge since most of the existing copula functions are derived considering the bivariate case. Thus, the benefits of the use of copula theory for LR based detection with multimodal data comes at a higher computational price. One of the commonly used suboptimal methods for fusion of multimodal data is to neglect inter-modal dependence and compute the likelihood ratio only based on the marginal pdfs of each modality (we call this 'the product approach' in the rest of the paper). However, this approach is expected to lead to poor performance when inter-modal dependence is strong.

[1]This work was supported in part by ARO grant no. W911NF-14-1-0339.

To overcome the computational difficulties in the fusion of high dimensional multimodal data for detection, in this paper, we consider the fusion problem in a compressed domain where compression is achieved via low dimensional random projections as proposed in the compressive sensing (CS) literature [9–12]. The problem of detection with compressive measurements has been addressed by several recent works [13–24]. While some of the work, such as [13, 14, 17, 20, 21, 24] focused on sparse signal detection, some other works [15, 16, 18, 19] considered the problem of detecting signals which are not necessarily sparse. When the signals are not necessarily sparse, it was observed that there is a performance loss when performing LR based detection in the compressed domain compared to that with uncompressed data. However, when the signal-to-noise ratio (SNR) is sufficiently large, this loss is not significant and the compressed detector, i. e., the detector based on compressed data, is capable of providing similar performance as the uncompressed detector. In [23], the authors have extended known signal detection with CS to the multiple sensor case. While intra-signal dependence was considered with Gaussian measurements, the inter-sensor dependence was neglected in [23]. To the best of authors knowledge, the benefits of CS based detection when it is difficult to perform LR based detection with uncompressed data due to inter-modal dependence have not been investigated in the literature.

In this paper, we seek the answer to the following question; is it beneficial, in terms of both performance and computational complexity, to model intermodal dependence to perform LR based detection in the compressed domain via Gaussian approximation over either neglecting dependence (product approach) or model dependence using suboptimal methods (e.g., copula based fusion without knowing exactly the best copula function that models dependence) with uncompressed data? With arbitrary marginal pdfs for each modality with uncompressed data, we show that, under certain conditions, better (or equivalent) detection performance can be achieved in the compressed domain with a small number of compressive measurements compared to performing fusion (i). with the product approach and (ii). when widely available copula functions are used to model dependence of uncompressed data. We briefly discuss how to determine conditions under which performing compressed detection is efficient and effective over suboptimal detection with uncompressed data in the presence of inter-modal dependence.

2. DETECTION WITH UNCOMPRESSED DATA

Let there be L sensor nodes in a network deployed to solve a detection problem. The measurement vector at each node is denoted by $\mathbf{x}_j \in \mathbb{R}^N$ for $j = 1, \cdots, L$. Under hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, $\mathbf{x}_j$ has the following pdfs:

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}_j \sim f_1(\mathbf{x}_j) \\ \mathcal{H}_0 &: \mathbf{x}_j \sim f_0(\mathbf{x}_j), j = 1, \cdots, L \end{aligned} \quad (1)$$

respectively, where $f_i(\mathbf{x}_j)$ denotes the joint probability density function (pdf) of $\mathbf{x}_j$ under $\mathcal{H}_i$ for $i = 0, 1$ and $j = 1, \cdots, L$. We assume that the elements of $\mathbf{x}_j$ are independent of each other, however, the vectors $\mathbf{x}_j's$ are dependent for $j = 1, \cdots, L$. This is a suitable model when the time samples collected at a given sensor are independent and there is spatial dependence among sensors in a distributed network. To perform LR based detection, it is required to compute the joint pdf of $\{\mathbf{x}_1, \cdots, \mathbf{x}_L\}$, which in general is difficult unless each $\mathbf{x}_j$ has a joint Gaussian pdf.

2.1. Copula based approach

In a parametric framework, copulas are used to construct a valid joint distribution describing an arbitrary, possibly nonlinear dependence structure [25]. According to copula theory, the pdfs of $\mathbf{x} = \{\mathbf{x}_1, \cdots, \mathbf{x}_L\}$ under $\mathcal{H}_i$ can be written as [25],

$$f_i(\mathbf{x}) = \prod_{n=1}^{N} \prod_{l=1}^{L} f_i(x_{nl}) c_{in}(u^i_{n1}, \cdots, u^i_{nL})$$

for $i = 0, 1$ where $c_{in}(\cdot)$ denotes the copula density function, $u^i_{nl} = F(x_{nl}|\mathcal{H}_i)$ with $F(x|\mathcal{H}_i)$ denoting the marginal cdf of x under $\mathcal{H}_i$, and x_{nl} is the n-th element of $\mathbf{x}_l$. Then, the log LR (LLR) can be written in the following form:

$$\begin{aligned} T_{LLR}(\mathbf{x}) &= \log \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} = \sum_{l=1}^{L} \sum_{n=1}^{N} \log \frac{f_1(\mathbf{x}_l[n])}{f_0(\mathbf{x}_l[n])} \\ &+ \sum_{n=1}^{N} \log \frac{c_{1n}(u^1_{1n}, \cdots, u^1_{Ln}|\phi_{1n})}{c_{0n}(u^0_{1n}, \cdots, u^0_{Ln}|\phi_{0n})} \end{aligned} \quad (2)$$

where ϕ_{1n} and ϕ_{0n} are copula parameters under $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively, for $n = 1, \cdots, N$. In this case, in general, N different copulas where each one is L-variate are selected to model dependence.

One of the fundamental challenges in copula theory is to find the copula density function that will best fit the given data set. Further, most of the copula density functions proposed in the literature consider the bivariate case. In order to model dependence of multimodal data with more than two modalities, several approaches such as the use of vines have been proposed in the literature [7], which are in general computationally complex. Thus, in order to better utilize copula theory for multimodal data fusion, these challenges need to be overcome. In the following, we consider a computationally efficient approach for multimodal data fusion in which dependence among data is modeled in a low dimensional transformed domain. We discuss the advantages/disadvantages of the proposed approach over the copula based approach.

3. DETECTION WITH COMPRESSED DATA

When the signals $\mathbf{x}_j$'s are high dimensional, it is desired that fusion be performed in a compressed domain. The use of low dimensional random projections for solving inference problems has been addressed in the recent literature [13–23]. Let $\mathbf{A}_j$ be specified by a set of unique sampling vectors $\{\mathbf{a}_{j,m}\}_{m=1}^{M}$ with $M < N$ for $j = 1, \cdots, L$. Then, the low dimensional samples can be expressed as,

$$\mathbf{y}_j = \mathbf{A}_j \mathbf{x}_j \quad (3)$$

for $j = 1, \cdots, L$ where $\mathbf{y}_j$ is the $M \times 1$ compressed measurement vector at the j-th node, and the m-th element of the vector $\mathbf{A}_j\mathbf{x}_j$ is given by $(\mathbf{A}_j\mathbf{x}_j)_m = \langle \mathbf{a}_{j,m}, \mathbf{x}_j \rangle$ for $m = 1, \cdots, M$ where $\langle ., . \rangle$ denotes the inner product. In CS theory, the mapping $\mathbf{A}_j$ is often selected to be a random matrix. Solving (3) when $\mathbf{x}_j$'s are Gaussian is considered in [15] with a single sensor and it is extended to the multiple sensor case in [23]. The degradation of performance in the compressed domain compared to that with uncompressed data while performing LR based detection is expressed in terms of the output SNR or the deflection coefficient in [15, 23]. However, when $\mathbf{x}_j$'s are not Gaussian and there is dependence among them, proper performance comparison for detection in the uncompressed and compressed domains is not available in the literature.

3.1. LR based detection with compressed data

In order to perform LR based detection based on (3), the computation of the joint pdf of $\{\mathbf{y}_1, \cdots, \mathbf{y}_L\}$ is necessary. If the marginal pdf of $\mathbf{x}_j$'s are available, the marginal pdfs of each element in $\mathbf{y}_j$'s can be computed as in the following. The m-th element of $\mathbf{y}_j$, y_{mj}, can be written as, $y_{mj} = \sum_{n=1}^{N} \mathbf{A}_j[m, n] x_{nj}$ where $\mathbf{A}_j[m, n]$ is the (m, n)-th element of $\mathbf{A}_j$. Having the marginal pdfs of x_{nj} and using the independence assumption, the joint pdf of $z = y_{mj}$ can be found after computing the characteristic function of z. It is further noted that $\{y_{mj}\}_{m=1}^{M}$ for a given j are not necessarily uncorrelated of each other although $\{x_{nj}\}_{n=1}^{N}$'s are uncorrelated unless certain conditions are satisfied by $\mathbf{A}_j$ and $\mathbf{x}_j$. For example, if the elements of $\mathbf{x}_j$ are zero mean Gaussian with the covariance matrix $\sigma_v^2\mathbf{I}$, and the projection matrix satisfies the condition $\mathbf{A}_j\mathbf{A}_j^T = \mathbf{I}$, then the elements of $\mathbf{y}_j$ are uncorrelated. However, in general this uncorrelatedness may not hold. Once the marginal pdfs of the elements in $\mathbf{y}_j$ for $j = 1, \cdots, L$ are found, copula theory can be used in order to find the joint pdf of the compressive measurement vectors $\mathbf{y}_1, \cdots, \mathbf{y}_L$. Letting $u_j = F_j(y_{qp})$ for $j = M(p-1) + q$ where $p = 1, \cdots, L$, $q = 1, \cdots, M$, the LLR based on copula functions can be expressed as,

$$T_{LLR}(\mathbf{y}) = \sum_{l=1}^{L} \sum_{k=1}^{M} \log \frac{f_1(y_{kl})}{f_0(y_{kl})} + \log \frac{c_1(u_1, \cdots, u_{ML}|\phi_1^*)}{c_0(u_1, \cdots, u_{ML}|\phi_0^*)}. \quad (4)$$

The second term on the right hand side in (4) requires finding copula density functions of ML variables which is computationally very difficult. Since we assume that the elements in each $\mathbf{x}_j$ are independent under any given hypothesis, each element in $\mathbf{y}_j$ can be approximated by a Gaussian random variable (via Lindeberg-Feller central limit theorem assuming the required conditions are satisfied [26,27]) for given $\mathbf{A}_j$ when N is sufficiently large. Then, LR based detection can be performed via Gaussian approximation, which makes the modeling of dependence among multimodal data with compressed measurements easier.

3.2. LR based detection via Gaussian approximation

Let $\mathbf{y} = [\mathbf{y}_1^T \cdots \mathbf{y}_L^T]^T$ be a $ML \times 1$ vector. With Gaussian approximation we have $\mathbf{y}|\mathcal{H}_i \sim \mathcal{N}(\boldsymbol{\mu}^i, \mathbf{C}^i)$ where

$$\boldsymbol{\mu}^i = [{\boldsymbol{\mu}_1^i}^T \cdots {\boldsymbol{\mu}_L^i}^T]^T \quad (5)$$

and

$$\mathbf{C}^i = \begin{bmatrix} \mathbf{C}^i_1 & \mathbf{C}^i_{12} & \cdots & \mathbf{C}^i_{1L} \\ \mathbf{C}^i_{21} & \mathbf{C}^i_2 & \cdots & \mathbf{C}^i_{2L} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{C}^i_{L1} & \mathbf{C}^i_{L2} & \cdots & \mathbf{C}^i_L \end{bmatrix} \quad (6)$$

with $\boldsymbol{\mu}_j^i = \mathbb{E}\{\mathbf{y}_j|\mathcal{H}_i\}$, $\mathbf{C}_j^i = \mathbb{E}\{(\mathbf{y}_j - \mathbb{E}\{\mathbf{y}_j\})(\mathbf{y}_j - \mathbb{E}\{\mathbf{y}_j\})^T|\mathcal{H}_i\}$, $\mathbf{C}_{jk}^i = \mathbb{E}\{(\mathbf{y}_j - \mathbb{E}\{\mathbf{y}_j\})(\mathbf{y}_k - \mathbb{E}\{\mathbf{y}_k\})^T|\mathcal{H}_i\}$ with $j \neq k$, $k = 1, \cdots, L$ and $j = 1, \cdots, L$ for $i = 0, 1$. Further, let $\boldsymbol{\beta}_j^i = \mathbb{E}\{\mathbf{x}_j|\mathcal{H}_i\}$, $\mathbf{D}_j^i = \mathbb{E}\{(\mathbf{x}_j - \mathbb{E}\{\mathbf{x}_j\})(\mathbf{x}_j - \mathbb{E}\{\mathbf{x}_j\})^T|\mathcal{H}_i\}$ and $\mathbf{D}_{jk}^i = \mathbb{E}\{(\mathbf{x}_j - \mathbb{E}\{\mathbf{x}_j\})(\mathbf{x}_k - \mathbb{E}\{\mathbf{x}_k\})^T|\mathcal{H}_i\}$ for $j \neq k$. Then we have,

$$\boldsymbol{\mu}_j^i = \mathbf{A}_j\boldsymbol{\beta}_j^i, \mathbf{C}_j^i = \mathbf{A}_j\mathbf{D}_j^i\mathbf{A}_j^T, \text{and } \mathbf{C}_{jk}^i = \mathbf{A}_j\mathbf{D}_{jk}^i\mathbf{A}_k^T \quad (7)$$

for $j, k = 1, \cdots, L$ and $i = 0, 1$. Then, we can write,

$$\boldsymbol{\mu}^i = \mathbf{A}\boldsymbol{\beta}^i \text{ and } \mathbf{C}^i = \mathbf{A}\mathbf{D}^i\mathbf{A}^T \quad (8)$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \cdot & \cdot & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdot & \cdot & \mathbf{0} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \cdot & \cdot & \mathbf{A}_L \end{pmatrix} \quad (9)$$

is a $ML \times NL$ matrix and $\boldsymbol{\beta}^i$ and $\mathbf{D}^i$ are notations analogous to $\boldsymbol{\mu}^i$ and $\mathbf{C}^i$, respectively. Then, the decision statistic of the LLR based detector is simply given by,

$$\Lambda = \mathbf{y}^T(\mathbf{C}^{1^{-1}} - \mathbf{C}^{0^{-1}})\mathbf{y} - 2(\boldsymbol{\mu}^{1^T}\mathbf{C}^{1^{-1}} - \boldsymbol{\mu}^{0^T}\mathbf{C}^{0^{-1}})\mathbf{y}.$$

To illustrate the detection performance with multimodal data in the compressed domain with Gaussian approximation compared to detection with uncompressed data, in the following, we present a numerical example considering $L = 2$. We further consider the elements of $\mathbf{A}_j$ to be iid zero mean Gaussian for $j = 1, 2$.

3.3. Example

We consider two cases. In Case I, $\mathbf{x}_1$, and $\mathbf{x}_2$ have the following marginal pdfs under the two hypotheses (as considered in [3]): $x_{i1}|\mathcal{H}_j \sim \mathcal{N}(0, \sigma_j^2)$, and $x_{i2}|\mathcal{H}_j \sim \text{Exp}(\lambda_j)$. It is noted that $x \sim \text{Exp}(\lambda)$ denotes that x has an exponential distribution with $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ and 0 otherwise. Under $\mathcal{H}_1$, x_{i2}'s are generated so that $x_{i2} = x_{i1}^2 + w^2$ where $w \sim \mathcal{N}(0, \sigma_1^2)$. Then we have $x_{i2} \sim \text{Exp}(\lambda_1)$ with $\lambda_1 = \frac{1}{2\sigma_1^2}$. Under $\mathcal{H}_0$, x_{i2}'s are generated independent of x_{i1} for $i = 1, \cdots, N$ with parameter λ_0.

For Case II, we consider that $x_{i1} \sim \text{Exp}(\lambda_j)$ and $x_{i2}|\mathcal{H}_j \sim \text{Beta}(a_i, b_i = 1)$ where $x \sim \text{Beta}(a, b)$ denotes that x has a beta distribution with pdf $f(x) = \frac{1}{\mathcal{B}(a,b)}x^{a-1}(1-x)^{b-1}$ and $\mathcal{B}(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is the beta function. Under $\mathcal{H}_1$, x_{i2}'s are generated so that

$$x_{i2} = \frac{u}{u + x_{i1}}$$

where $u \sim \text{Gamma}(\alpha_1, \beta_1 = 1/\lambda_1)$. Then $x_{i2}|\mathcal{H}_1 \sim \text{Beta}(a_1, b_1 = 1)$ with $a_1 = \alpha_1$. It is noted that $x \sim \text{Gamma}(\alpha, \beta)$ denotes that x has Gamma pdf with $f(x) = \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}$ for $x \geq 0$ and $\alpha, \beta > 0$. Under $\mathcal{H}_0$, x_{i2} is generated independent of x_{i1} with parameters a_0 and $b_0 = 1$.

First, we illustrate how the dependence structure of the data changes from uncompressed domain to the compressed domain. In Fig. 1, we show the scatter plots for both compressed and uncompressed data at the two sensors under $\mathcal{H}_1$. In Fig. 1, the top and bottom subplots are for Case I and Case II, respectively while left and right subplots are for uncompressed and compressed data, respectively. It can be observed that while uncompressed data at the two sensors are strongly dependent of each other, compressed data appears to be weakly dependent with a completely different (Gaussian like) pattern.

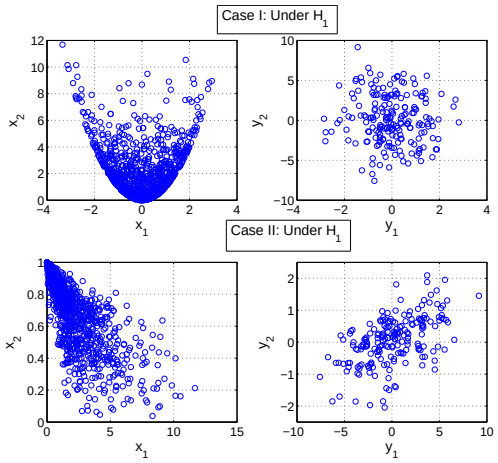
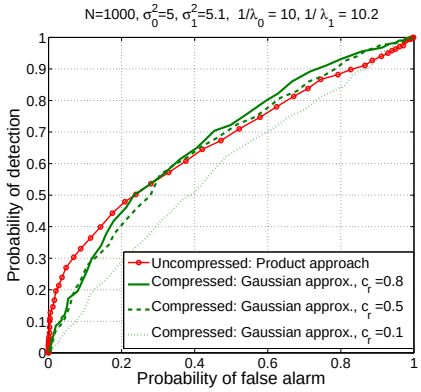
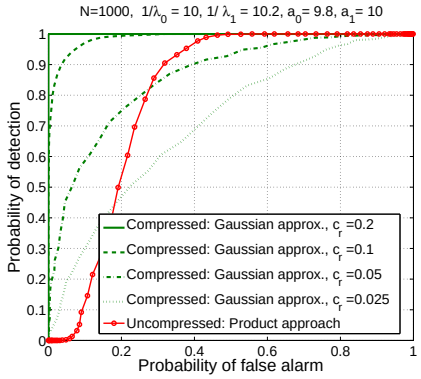


Fig. 1: Scatter plots of uncompressed and compressed data under $\mathcal{H}_1$; $N = 1000$, $M = 200$, $L = 2$



(a) Case I: $\sigma_0^2 = 5$, $\sigma_1^2 = 5.1$, $1/\lambda_0 = 10$, $1/\lambda_1 = 10.2$



(b) Case II: $1/\lambda_0 = 10$, $1/\lambda_1 = 10.2$, $a_0 = 9.8$, $a_1 = 10$

Fig. 2: Detection performance with multimodal dependent data in the compressed and uncompressed domains: $N = 1000$

3.4. Product approach with uncompressed data vs. Gaussian approximation with compressed data

In the following, we compare the detection performance with compressed multimodal data and the product approach (where depen-

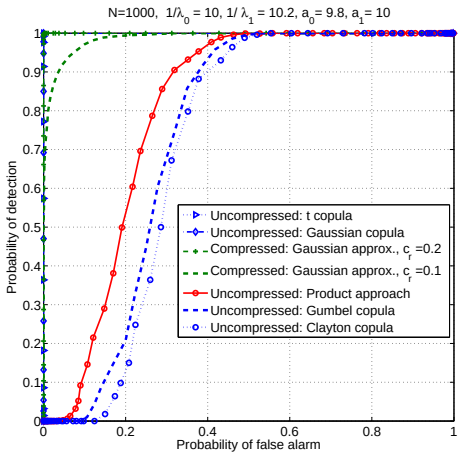


Fig. 3: Detection performance with multimodal dependent data in the compressed and uncompressed domains for case II; $1/\lambda_0 = 10$, $1/\lambda_1 = 10.2$, $a_0 = 9.8$, $a_1 = 10$, $b_0 = b_1 = 1$

dence is ignored) with uncompressed data. Fig. 2 shows the performance in terms of the ROC curves for the two cases considered in Example 3.3. We make several important observations here. In Case I, the detection performance with the Gaussian approximation in the uncompressed domain is only slightly better than that with the product approach in the uncompressed domain when the compression ratio, $c_r = \frac{M}{N}$, is relatively large and the probability of false alarm is high. For small c_r, the product approach with uncompressed data shows better performance than the Gaussian approximation, however, the performance gap is not very significant. In Case II, we observe a significant performance gain when performing detection with compressed data even with relatively small c_r compared to the product approach in the uncompressed domain. It is noted that in Case I, the observations at the two sensors are uncorrelated with uncompressed data (although they are dependent) thus $\mathbf{D}^1$ is diagonal. Thus, not taking dependence into account in the uncompressed domain seems not to result in a large performance loss compared to taking dependence in the compressed domain into account. On the other hand, when considering Case II, it is noted that $\mathbf{D}^1$ is not diagonal, and the uncompressed observations under $\mathcal{H}_1$ are strongly correlated. Thus, ignoring dependence with uncompressed data leads to severe performance loss compared to taking dependence (via Gaussian approximation) into account in the compressed domain even with very small c_r. Further, in that case, it is observed that, there is a threshold for c_r after which the Gaussian approximation in the compressed domain starts to perform better than the product approach with uncompressed data.

3.5. Copula based fusion with uncompressed data vs. Gaussian approximation with compressed data

Next, we compare the detection performance when copulas are used to compute the joint pdf with uncompressed data in Fig. 3. Since finding optimal copula function that models a given data set is computationally complex, we plot the detection performance using widely available bivariate copula functions. To that end, we consider Gaussian, t, Gumbel and Clayton copula functions as described in [3, 8]. Further, we consider Example 3.3 with Case II. We further plot the detection performance with the product approach with uncompressed data. It is observed from Fig. 3 that fusion with Gaussian and t copula functions leads to perfect detection, while fusion with Gumbel and Clayton copula provides poor performance even compared to the product approach. On the other hand, fusion performance with compressed data with $c_r = 0.2$ is capable of providing perfect detection with the parameters considered. Thus, with the considered problem parameters, the use of copula functions with uncompressed data seems to be a waste of resources when perfect detection can be achieved with less computational complexity in the compressed domain via Gaussian approximation. Thus, it is worthwhile to investigate as to when it is beneficial to use copula theory to model dependence with uncompressed data compared to performing fusion by modeling dependence with compressed data in a computationally easier fashion. We briefly address this issue in the following.

In order to quantify the performance of detection with both uncompressed and compressed data, we consider Kullback-Leibler (KL) distance to be the performance metric. The KL distance between the pdfs under the two hypotheses in the compressed domain with the Gaussian approximation can be computed as [28]

$$\begin{aligned} & \mathcal{D}_{KL}^{c,G}(f_0||f_1) \\ = \; & \frac{1}{2}\Big\{\mathrm{tr}(\mathbf{A}^{\ddagger}\mathbf{D}^0) + (\boldsymbol{\beta}^1-\boldsymbol{\beta}^0)^T\mathbf{A}^{\ddagger}(\boldsymbol{\beta}^1-\boldsymbol{\beta}^0) \\ & -ML + \log\frac{|\mathbf{A}\mathbf{D}^1\mathbf{A}^T|}{|\mathbf{A}\mathbf{D}^0\mathbf{A}^T|}\Big\} \end{aligned} \tag{10}$$

where $\mathbf{A}^{\ddagger} = \mathbf{A}^T(\mathbf{A}\mathbf{D}^1\mathbf{A}^T)^{-1}\mathbf{A}$ and $\mathrm{tr}(\cdot)$ denotes the trace operator. In the case where $\mathbf{x}_l$'s for $l = 1, \cdots, L$ are assumed to be independent of each other under $\mathcal{H}_0$, we have $f_0(\mathbf{x}) = \prod_{l,n} f_0^m(x_{nl})$ where f_i^m denotes the marginal pdf under $\mathcal{H}_i$. Thus, the KL distance between $f_0(\cdot)$ and $f_1(\cdot)$ with uncompressed data can be written as,

$$\begin{aligned} \mathcal{D}_{KL}^{u}(f_0||f_1) &= \mathcal{D}_{KL}^{u,p}(f_0||f_1^m) \\ &- \underbrace{\mathbb{E}\left\{\sum_{n=1}^{N}\log c_{1n}(u_{1n}^1,\cdots,u_{Ln}^1|\phi_{1n})|\mathcal{H}_0\right\}}_{\Upsilon_{f_0,c}} \end{aligned}$$

where $\mathcal{D}_{KL}^{u,p}(f_0||f_1^m)$ denotes the KL distance under the product approach. When the marginal pdfs are available, $\mathcal{D}_{KL}^{u,p}(f_0||f_1^m)$ can be computed. It is noted that the term $\Upsilon_{f_0,c}$ depends on the particular copula function used to model dependence. Thus, for a given copula function, when $\Upsilon_{f_0,c} > \mathcal{D}_{KL}^{u,p} - \mathcal{D}_{KL}^{c,G}$ performing detection in the compressed domain with given M ($\mathcal{D}_{KL}^{c,G}$ is a function of M) appears to be more effective and efficient than copula based fusion in the uncompressed domain. This issue will be further addressed in detail in future work.

4. CONCLUSION

In this paper, we showed that, under certain conditions, detection with multimodal dependent data with compressive sensing can be better (or equivalent) than detection with the widely considered product approach and copula based fusion with uncompressed data. We briefly discussed the conditions under which modeling dependence for likelihood ratio based detection in the compressed domain is more efficient and effective than modeling dependence with uncompressed data using copula theory which is computationally expensive most of the time. Experiments with real datasets will be considered in future work.

5. REFERENCES

[1] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, 2015.

[2] G. Mercier, G. Moser, and S. Serpico, "Conditional copula for change detection on heterogeneous SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, July 2007, pp. 2394–2397.

[3] S. G. Iyengar, P. K. Varshney, and T. Damarla, "A parametric copula-based framework for hypothesis testing using heterogeneous data," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2308–2318, May 2011.

[4] A. Sundaresan and P. K. Varshney, "Location estimation of a random signal source based on correlated sensor observations," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 787–799, Feb. 2011.

[5] A. Sundaresan, P. K. Varshney, and N. S. V. Rao, "Copula-based fusion of correlated decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, p. 454=471, Jan. 2011.

[6] S. G. Iyengar, P. K. Varshney, and T. Damarla, "Biometric authentication: A copula-based approach," in *Multibiometrics for Human Identification*, B. Bhanu and V. Govindaraju, Eds. Cambridge University Press, 2011, pp. 95–119.

[7] A. Subramanian, A. Sundaresan, and P. K. Varshney, "Fusion for the detection of dependent signals using multivariate copulas," in *14th International Conference on Information Fusion*, Chicago, Illinois, USA, 2011, pp. 740–747.

[8] H. He and P. K. Varshney, "Fusing censored dependent data for distributed detection," *IEEE Trans. Signal Process.*, vol. 63, no. 16, pp. 4385–4395, Aug. 2015.

[9] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489 – 509, Feb. 2006.

[10] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406 – 5425, Dec. 2006.

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

[12] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[13] M. F. Duarte, M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, "Sparse signal detection from incoherent projections," in *Proc. Acoust., Speech, Signal Processing (ICASSP)*, May 2006.

[14] J. Haupt and R. Nowak, "Compressive sampling for signal detection," in *Proc. Acoust., Speech, Signal Processing (ICASSP)*, vol. 3, Honolulu, Hawaii, Apr. 2007, pp. III–1509 – III–1512.

[15] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. Baraniuk, "Signal processing with compressive measurements," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 445 – 460, Apr. 2010.

[16] T. Wimalajeewa, H. Chen, and P. K. Varshney, "Performance analysis of stochastic signal detection with compressive measurements," in 44^{th} *Annual Asilomar Conf. on Signals, Systems and Computers*, Nov. 2010, pp. 913–817.

[17] G. Li, H. Zhang, T. Wimalajeewa, and P. K. Varshney, "On the detection of sparse signals with sensor networks based on Subspace Pursuit," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Atlanta, GA, Dec. 2014, pp. 438–442.

[18] B. Kailkhura, T. Wimalajeewa, L. Shen, and P. K. Varshney, "Distributed compressive detection with perfect secrecy," in *2nd Int. Workshop on Compressive Sensing in Cyber-Physical Systems (CSCPS'14)*, Oct. 2014.

[19] B. Kailkhura, T. Wimalajeewa, and P. K. Varshney, "On physical layer secrecy of collaborative compressive detection," in 48^{th} *Annual Asilomar Conf. on Signals, Systems and Computers*, 2014.

[20] B. S. M. R. Rao, S. Chatterjee, and B. Ottersten, "Detection of sparse random signals using compressive measurements," in *Proc. Acoust., Speech, Signal Processing (ICASSP)*, 2012, pp. 3257–3260.

[21] J. Cao and Z. Lin, "Bayesian signal detection with compressed measurements," *Information Sciences*, pp. 241–253, 2014.

[22] B. Kailkhura, S. Liu, T. Wimalajeewa, and P. K. Varshney, "Measurement matrix design for compressed detection with secrecy guarantees," *IEEE Wireless Commun. Lett.*, 2016, Accepted.

[23] B. Kailkhura, T. Wimalajeewa, and P. K. Varshney, "Collaborative compressive detection with physical layer secrecy constraints," *IEEE Trans. Signal Process.*, 2016, Submitted.

[24] T. Wimalajeewa and P. K. Varshney, "Sparse signal detection with compressive measurements via partial support set estimation'," *IEEE Trans. on Signal and Inf. Process. over Netw.*, vol. 3, no. 1, Mar. 2017.

[25] R. Nelsen, *An Introduction to Copulas*, 2nd ed. New York: Springer, 2006.

[26] H. Cramer, *Mathematical Methods of Statistics*. Princeton, New Jersey: Princeton University Press, 1946.

[27] T. Wimalajeewa, H. Chen, and P. K. Varshney, "Performance limits of compressive sensing-based signal classification," *IEEE Trans. on Signal Process.*, vol. 60, no. 6, pp. 2758–2770, June 2012.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, Inc., NY, 2006.