MULTIPLE HYPOTHESIS TESTING FOR COMPRESSIVE WIDEBAND SENSING

Sundeep Prabhakar Chepuri, Geert Leus

Delft University of Technology
Faculty of Electrical Engineering
Delft, The Netherlands

Ruben de Francisco

Holst Centre/imec-nl
Wireless Group
Eindhoven, The Netherlands

ABSTRACT

The classical Compressive Sensing (CS) techniques for wideband spectral sensing consist of a two-stage estimation detection approach. A novel approach is proposed to solve the detection problem directly from observations with incomplete frequency information, for e.g., signals acquired using sub-Nyquist rate sampling. The wideband occupancy detection problem is formulated as a multiple hypothesis testing problem under a non-Bayesian framework, and a Neyman-Pearson-like criterion is proposed. The detector based on exhaustive search performs better than the conventional CS based techniques. However, it is impractical for large block sizes. Hence, we propose a sub-optimal greedy algorithm whose complexity and performance can be traded-off by construction.

1. INTRODUCTION

Efficient spectrum sharing can be achieved in wireless sensor networks through wideband spectrum sensing, by identifying available channels within a large frequency range. In order to sense large bandwidths, in the order of a few hundred MHz, high-rate Analog-to-Digital Converters (ADCs) or complex receiver front-ends are required.

Currently, there is a great interest in reducing the sampling rate for sparse signals and relax the requirements on the ADCs. These are often casted as a Compressive Sensing (CS) problem, where the data is acquired at a rate significantly lower than the Nyquist rate. The sampling rate could be reduced by using the architectures reported in the literature, such as multi-coset sampling [1] or modulated wideband converters [2]. Later, the signal can be recovered with one of the many available sparse recovery algorithms with little or no loss of information. For sensing using classical CS based techniques, either the entire signal or its statistical measures are reconstructed first. Then, in the second stage, detection is performed on this compressive estimate. Here, we avoid this two stage estimation-detection approach. Instead we perform a direct detection on the compressed samples obtained using sub-Nyquist rate sampling. That is the main contribution of this paper. Such detection problems appear in various fields such as event detection in radar, Multi-User (MU) detection in communications, imaging, and spectrum sensing for Cognitive Radios (CRs).

This work was supported in part by STW under FASTCOM project (10551) and in part by NWO-STW under the VICI program(10382).

Here, we consider a multiband occupancy detection problem. In a multiband spectrum, each band could be either busy or free. We formulate this detection problem as a Multiple Hypothesis Testing (MHT) problem under a non-Bayesian philosophy, with each hypothesis describing one possible combination (combinatorial in the number of channels N) of all the channels being busy and/or free. A channel length of N would result in one *null hypothesis* $\mathcal{H}_0$, and $m = 2^N - 1$, *alternative hypotheses* indicated by $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_m$. Solving the MHT problem would give the occupancy of all the N channels at once, and would avoid solving a binary hypothesis problem on every channel.

The MHT problem can be related to model order selection [3, 4], which is generally casted as a parameter estimation problem in a Bayesian framework. However, in this paper we consider a detection problem under a non-Bayesian framework.

We develop a detector for a signal that is acquired using sub-Nyquist rate sampling, i.e., the number of available measurements M is less than the number of Nyquist rate measurements N. This is termed a Compressed Detector (CD). We propose an exhaustive search algorithm as well as a sub-optimal greedy algorithm for CD. The performance of these detectors is analyzed through simulations in terms of the detection probability, the false alarm probability, and the compression rate, which is given by $\frac{M}{N}$.

2. DETECTION MODEL

We start by defining the probability of detection, P_d, and the probability of false alarm, P_{fa}, which are used to analyze the performance of Wideband Sensing (WS) for a given static channel occupancy.

Definition 1. Probability of detection and probability of false alarm**.** *Consider a wideband spectrum of B Hz segmented into N channels, such that each channel has a bandwidth of $\frac{B}{N}$ Hz. The channels are indexed from 1 to N. These channels can be either busy or free, depending on whether there is or*

978-1-4673-0971-4/12/$31.00 ©2012 IEEE

there is not a signal transmission, respectively. The indices of such K busy channels are collected in a set

$$\mathbf{b} = \{b_1, b_2, \dots, b_K\}, \tag{1}$$

with $|\mathbf{b}| = K$. The complement of the set $\mathbf{b}$ is denoted by

$$\mathbf{b}^c = \{b_1^c, b_2^c, \dots, b_{N-K}^c\}, \tag{2}$$

with $|\mathbf{b}^c| = N - K$. The probability of detection is then defined as

$$P_d = \frac{1}{K}\sum_{j=1}^{K} \mathcal{X}(b_j \in \tilde{\mathbf{b}} | \tilde{\mathbf{b}} \textit{ detected})\, Pr(\tilde{\mathbf{b}} \textit{ detected}) \tag{3}$$

where, $\mathcal{X}(.)$ denotes the indicator function and $\tilde{\mathbf{b}}$ is the detected busy channel set.

The probability of false alarm is defined as

$$P_{fa} = \frac{1}{N-K}\sum_{j=1}^{N-K} \mathcal{X}(b_j^c \in \tilde{\mathbf{b}} | \tilde{\mathbf{b}} \textit{ detected})\, Pr(\tilde{\mathbf{b}} \textit{ detected}). \tag{4}$$

The MHT detector in the Neyman-Pearson sense is the most likely hypothesis, that would optimize the probability of detection, P_d, (as defined in (3)) with a constraint on the probability of false alarm, P_{fa} (as defined in (4)). However, this detector is complicated to derive. Hence, we propose a simplified detector for the MHT problem using a Neyman-Pearson-like criterion, in the sense that we do not assign any prior probabilities to the hypotheses.

Proposition 1. MHT under a Neyman-Pearson-like criterion. *Let $\mathbf{y}$ denote the observations from a certain process. Let the* null hypothesis *be denoted by, $\mathcal{H}_0$, and the* alternative hypotheses *by, $\mathcal{H}_i$, with $i = 1, \dots, m = 2^N - 1$. The proposed MHT detector under a Neyman-Pearson-like criterion is*

$$i^* = \arg\max_i \left(\Lambda(i) = \ln\frac{Pr(\mathbf{y}|\mathcal{H}_i)}{Pr(\mathbf{y}|\mathcal{H}_0)}\right),\ i = 1, 2, \dots, m. \tag{5}$$

$$\Lambda(i^*) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{i^*}}{\gtrless}} \gamma_{th}. \tag{6}$$

The optimization problem (5) will result in the most likely hypothesis. A selection between the *null hypothesis* $\mathcal{H}_0$ and the most likely *alternative hypothesis* $\mathcal{H}_{i^*}$ is done using the threshold γ_{th}, as in (6).

For the proposed MHT detector we can further re-write P_d and P_{fa} as

$$\begin{aligned} P_d &= \frac{1}{K}\sum_{j=1}^{K}\sum_{i^*=0}^{m} \mathcal{X}(b_j \in \tilde{\mathbf{b}}^{(i^*)} | i^* \text{ is solution})\, P_{H_{i^*}}, \\ P_{fa} &= \frac{1}{N-K}\sum_{j=1}^{N-K}\sum_{i^*=0}^{m} \mathcal{X}(b_j^c \in \tilde{\mathbf{b}}^{(i^*)} | i^* \text{ is solution}) P_{H_{i^*}}. \end{aligned} \tag{7}$$

where, $\tilde{\mathbf{b}}^{(i^*)}$ is one of the possible detected busy channel set and $P_{H_{i^*}}$ is the probability that $\mathcal{H}_{i^*}$ is selected. We choose γ_{th} based on simulations, to maintain a certain P_{fa}.

3. SIGNAL MODEL

Let the time domain signal representing N frequency channels be denoted by the $N \times 1$ vector $\mathbf{x} \in \mathbb{C}^N$ and the noise by the $N \times 1$ vector $\mathbf{v} \in \mathbb{C}^N$. The signal $\mathbf{x}$ can be written in terms of its frequency response $\mathbf{x}_f \in \mathbb{C}^N$ as $\mathbf{x} = \mathbf{F}^H\mathbf{x}_f$, where $\mathbf{F} \in \mathbb{C}^{N\times N}$ is the normalized Discrete Fourier Transform (DFT) matrix. Similarly, the noise $\mathbf{v}$ can be written in terms of its frequency response $\mathbf{v}_f \in \mathbb{C}^N$ as $\mathbf{v} = \mathbf{F}^H\mathbf{v}_f$. The passband signal at the receiver can then be written as

$$\mathbf{y} = \mathbf{\Psi}(\mathbf{x}_f + \mathbf{v}_f), \tag{8}$$

where, $\mathbf{\Psi} = \mathbf{F}^H$. We acquire the received signal through a linear measurement process modeled by the *sensing matrix*, $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$, with $M < N$. For sub-Nyquist rate sampling, $\mathbf{\Phi}$ is a fat matrix, resulting in an under-determined system. The acquired signal is denoted by the $M \times 1$ vector

$$\tilde{\mathbf{y}} = \mathbf{\Phi}\mathbf{y} = \tilde{\mathbf{\Psi}}(\mathbf{x}_f + \mathbf{v}_f), \tag{9}$$

where, $\tilde{\mathbf{\Psi}} = \mathbf{\Phi}\mathbf{\Psi}$.

Let the combination of N frequency bands being free (indicated by "0") and/or the frequency bands being occupied (indicated by "1") for the ith hypothesis be denoted by the $N \times 1$ vector $\mathbf{c}_{x|\mathcal{H}_i}$. Such that

$$\begin{aligned} &i = 0 : \mathbf{c}_{x|\mathcal{H}_0} = \left[0, 0, \dots, 0, 0\right]^T, \\ &i = 1 : \mathbf{c}_{x|\mathcal{H}_1} = \left[0, 0, \dots, 0, 1\right]^T, \\ &i = 2 : \mathbf{c}_{x|\mathcal{H}_2} = \left[0, 0, \dots, 1, 0\right]^T, \\ &\vdots \\ &i = 2^N - 1 : \mathbf{c}_{x|\mathcal{H}_{2^N-1}} = \left[1, 1, \dots, 1, 1\right]^T. \end{aligned} \tag{10}$$

The number of non-zero entries of the vector $\mathbf{c}_{x|\mathcal{H}_i}$ is denoted by its ℓ_0-norm, i.e., $\|\mathbf{c}_{x|\mathcal{H}_i}\|_0$. The variance of any active channel is modeled as σ_x^2, and the variance of the noise in each channel as σ_v^2. Assuming that the channels are uncorrelated, we can then define the following covariance matrices $\mathbf{\Sigma}_{x|\mathcal{H}_i} = \mathbb{E}[\mathbf{x}_f\mathbf{x}_f^H|\mathcal{H}_i] = \sigma_x^2\text{diag}(\mathbf{c}_{x|\mathcal{H}_i}) = \sigma_x^2\mathbf{C}_{x|\mathcal{H}_i}$ and $\mathbf{\Sigma}_v = \mathbb{E}[\mathbf{v}_f\mathbf{v}_f^H] = \sigma_v^2\mathbf{I}_N$.

Both the signal and noise are modeled as independently and identically distributed (i.i.d.) Gaussian random variables. Hence, for the ith hypothesis, the signal can be written as $\mathbf{x}^{(i)} \sim \mathcal{CN}(0, \mathbf{\Psi}\mathbf{\Sigma}_{x|\mathcal{H}_i}\mathbf{\Psi}^H)$ and the noise as $\mathbf{v} \sim \mathcal{CN}(0, \sigma_v^2\mathbf{I}_N)$.

4. OPTIMIZATION PROBLEM

The algorithm to solve the MHT problem will decide on one of the following hypotheses

$$\begin{aligned} \mathcal{H}_0 : &\quad \tilde{\mathbf{y}} = \tilde{\mathbf{\Psi}}\mathbf{v}_f, \\ \mathcal{H}_i : &\quad \tilde{\mathbf{y}} = \tilde{\mathbf{\Psi}}(\mathbf{x}_f^{(i)} + \mathbf{v}_f),\ i = 1, \dots, m = 2^N - 1. \end{aligned} \tag{11}$$

The acquired signal can be written in terms of the compressed signal denoted by $\tilde{\mathbf{x}}^{(i)} = \tilde{\mathbf{\Psi}}\mathbf{x}_f^{(i)}$, and the noise $\tilde{\mathbf{v}} = \tilde{\mathbf{\Psi}}\mathbf{v}_f$. Since, the measurement process is linear, the compressed vectors will still be Gaussian random variables. The covariance matrix of the signal $\tilde{\mathbf{x}}$ for the ith hypothesis is given by

$$\tilde{\mathbf{\Sigma}}_{x|\mathcal{H}_i} = \mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H|\mathcal{H}_i] = \tilde{\mathbf{\Psi}}\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}^H, \tag{12}$$

and for the noise $\tilde{\mathbf{v}}$ it is given by

$$\tilde{\mathbf{\Sigma}}_v = \mathbb{E}[\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H] = \tilde{\mathbf{\Psi}}\mathbf{\Sigma}_v\tilde{\mathbf{\Psi}}^H = \sigma_v^2\mathbf{\Phi}\mathbf{\Phi}^H. \tag{13}$$

Hence, the covariance matrix of the acquired signal can be written as

$$\tilde{\mathbf{\Sigma}}_{y|\mathcal{H}_i} = \mathbb{E}[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^H|\mathcal{H}_i] = \tilde{\mathbf{\Psi}}\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}^H + \sigma_v^2\mathbf{\Phi}\mathbf{\Phi}^H. \tag{14}$$

From (12), (13) and (14) we can write

$$Pr(\tilde{\mathbf{y}}|\mathcal{H}_0) = \frac{|\tilde{\mathbf{\Sigma}}_v|^{-\frac{1}{2}}}{(2\pi)^{(\frac{N}{2})}} \exp\left(-\frac{1}{2}\tilde{\mathbf{y}}^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{y}}\right), \tag{15}$$

$$\begin{aligned} Pr(\tilde{\mathbf{y}}|\mathcal{H}_i) =& \frac{|\tilde{\mathbf{\Sigma}}_{y|\mathcal{H}_i}|^{-\frac{1}{2}}}{(2\pi)^{(\frac{N}{2})}} \\ & \exp\left(-\frac{1}{2}\tilde{\mathbf{y}}^H(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}^H + \tilde{\mathbf{\Sigma}}_v)^{-1}\tilde{\mathbf{y}}\right). \end{aligned} \tag{16}$$

Substituting (15) and (16) in (5), and scaling appropriately we can write the Log Likelihood Ratio (LLR) as

$$\begin{aligned} \Lambda(i) = \ln\left(\frac{|\tilde{\mathbf{\Sigma}}_{y|\mathcal{H}_i}|}{|\tilde{\mathbf{\Sigma}}_v|}\right) + \\ \tilde{\mathbf{y}}^H\left(\tilde{\mathbf{\Sigma}}_v^{-1} - (\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}^H + \tilde{\mathbf{\Sigma}}_v)^{-1}\right)\tilde{\mathbf{y}}. \end{aligned} \tag{17}$$

Using the matrix inversion lemma, we obtain

$$\begin{aligned} &(\tilde{\mathbf{\Psi}}\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}^H + \tilde{\mathbf{\Sigma}}_v)^{-1} \\ &= \tilde{\mathbf{\Sigma}}_v^{-1} - \tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}(\mathbf{\Sigma}_{x|\mathcal{H}_i}^{-1} + \tilde{\mathbf{\Psi}}^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}})^{-1}\tilde{\mathbf{\Psi}}^H\tilde{\mathbf{\Sigma}}_v^{-1}. \end{aligned} \tag{18}$$

Further factorizing (18) and simplifying the determinant, we can rewrite the objective function as

$$\begin{aligned} \Lambda(i) = \sigma_x^2\tilde{\mathbf{y}}^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}\mathbf{C}_{x|\mathcal{H}_i}(\mathbf{I}_N + \sigma_x^2\mathbf{C}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}} \\ \mathbf{C}_{x|\mathcal{H}_i})^{-1}\mathbf{C}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{y}} - \|\mathbf{c}_{x|\mathcal{H}_i}\|_0\ln(1+\gamma), \end{aligned} \tag{19}$$

where $\gamma = \frac{\sigma_x^2}{\sigma_v^2}$.

It is mathematically intricate to factor out the matrix $\mathbf{C}_{x|\mathcal{H}_i}$ from the matrix inverse in (19). This makes the optimization an involved non-convex non-linear integer programming problem of high complexity. An exhaustive search would require $O(2^N)$ computations, and is practically not feasible for large N. We term this algorithm based on exhaustive search as MHT-CD. Therefore, we propose a sub-optimal greedy algorithm to solve this optimization problem, based on certain heuristics.

Before presenting the proposed greedy algorithm, we provide some definitions.

Definition 2. Neighborhood. *For $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$, the neighborhood of $\mathbf{c}_{x|\mathcal{H}_i}$ with size S is defined as the set*

$$\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i}) = \{\mathbf{c}_{x|\mathcal{H}_j} \in \{0,1\}^N \mid \|\mathbf{c}_{x|\mathcal{H}_i} - \mathbf{c}_{x|\mathcal{H}_j}\|_1 \le S\}, \tag{20}$$

where, $\|\mathbf{c}_{x|\mathcal{H}_i} - \mathbf{c}_{x|\mathcal{H}_j}\|_1$ denotes the Hamming distance between $\mathbf{c}_{x|\mathcal{H}_i}$ and $\mathbf{c}_{x|\mathcal{H}_j}$.

This means, the vectors $\mathbf{c}_{x|\mathcal{H}_j}$ and $\mathbf{c}_{x|\mathcal{H}_i}$ differ by S bits $\forall\, \mathbf{c}_{x|\mathcal{H}_j} \in \mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$. For any $\mathbf{c}_{x|\mathcal{H}_i}$, the total number of vectors in $\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$ will be $|\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})| = \sum_{r=0}^{S}\binom{N}{r}$.

Definition 3. Local Maximum LLR (LML) point. *Consider $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$ with a neighborhood $\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$. If $\Lambda(p^*) \ge \Lambda(j)$ $\forall \mathbf{c}_{x|\mathcal{H}_j} \in \mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$ and $\forall \mathbf{c}_{x|\mathcal{H}_{p^*}} \in \mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$, then $\mathbf{c}_{x|\mathcal{H}_{p^*}}$ results in the local maximum LLR $\Lambda_S^*(p^*)$ with an LML point p^* within the neighborhood size S.*

4.1. Heuristics

Property 1. *Consider $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$ with a neighborhood $\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$ and LML point $\Lambda_S^*(i)$. Then $\Lambda_1^*(i) \le \Lambda_2^*(i) \cdots \le \Lambda_N^*(i) = \Lambda(i^*)$.*

For $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$ with a neighborhood $\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$, the LLR values between $\mathbf{c}_{x|\mathcal{H}_i}$ and $\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_i})$ are closer to each other for smaller S. As $S \to N$, the difference between the LLR values increases. We make use of this property in the proposed sub-optimal greedy algorithm. Since the algorithm is based on the LML and used to perform CD, we call it LML-CD.

An approach based on the LML has also been proposed in the literature for MU detection [5]. These detectors are for uncompressed signals and are gradient based, where the objective function is sequentially updated by bit-flipping so that its likelihood monotonically increases in each step. The detectors in [5] are for the uncompressed signal, hence it is possible to easily compute the likelihood for each component of the vector. However, this is complicated for the compressed signal as in (19). In this paper, we propose a novel approach for the compressed signals based on the LML, where we select an initial neighborhood region and then build upon the selected region depending on the required performance and complexity.

4.2. Algorithm

The algorithm LML-CD is initialized with a vector $\mathbf{c}(0) \in \{0,1\}^N$ uniformly at random out of 2^N possible vectors. From property 1, we know that the LLR values are closer to each other for smaller S. Hence, we choose N vectors $\mathbf{c}(u) \in \{0,1\}^N$, $u = 1, 2, \ldots, N$, uniformly at random such that these N vectors have all possible Hamming distances from 1 to N, with the initial vector. An observation space from these N vectors and the initial vector is formed. The maximum LLR and hence the hypothesis $\mathcal{H}_{i_1^*}$ within this initial observation space is selected. Next, the observation space is updated with all the possible vectors in the neighborhood $\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_{i_1^*}})$. The hypothesis $\mathcal{H}_{i_2^*}$ resulting in the maximum LLR with in this updated observation space is computed. If the LLR value $\Lambda(i_2^*)$ exceeds the pre-determined threshold γ_{th} then we choose hypothesis $\mathcal{H}_{i_2^*}$ otherwise we choose $\mathcal{H}_0$. The algorithm LML-CD is summarized in Table 1.

For $S = 3$, the algorithm LML-CD requires $\frac{1}{6}(N^3 + 11N + 6)$ computations, with complexity order of $O(N^3)$. The advantage of the LML-CD algorithm is that the complexity and the performance can be traded-off and can be adapted through construction by selecting an appropriate maximum neighborhood size S. In other words, if the neighborhood size S is increased from one, two, etc., up to N, the computational complexity is linear, quadratic, etc., up to exponential in the number of channels and also the performance increases.

Table 1: LML-CD algorithm for WS.

Objective: Select a hypothesis $\mathcal{H}_i$, $i = 0, 1, \ldots, 2^N - 1$.

1. **Initialization**: Start with an initial vector $\mathbf{c}(0) \in \{0,1\}^N$ uniformly at random.
2. **Generate** N **vectors**: $\mathbf{c}(u) \in \{0,1\}^N$, $u = 1, 2, \ldots, N$, each with a Hamming distance of 1 to N from the initial vector $\mathbf{c}(0)$.
3. **Observation space and index set**: $\mathcal{U}_1 = \{\mathbf{c}(u), u = 0, 1, \ldots, N\}$, $|\mathcal{U}_1| = N + 1$, $\mathcal{I}_1 = \{i | \mathbf{c}_{x|\mathcal{H}_i} \in \mathcal{U}_1\}$.
4. **Compute**: $i_1^* = \arg \max_{i \in \mathcal{I}_1} \Lambda(i)$.
5. **Update observation space and index set**: $\mathcal{U}_2 = \{\mathcal{N}_S(\mathbf{c}_{x|\mathcal{H}_{i_1^*}})\}$, $|\mathcal{U}_2| = \sum_{r=1}^{S} \binom{N}{r}$, $\mathcal{I}_2 = \{i | \mathbf{c}_{x|\mathcal{H}_i} \in \mathcal{U}_2\}$.
6. **Compute**: $i_2^* = \arg \max_{i \in \mathcal{I}_2} \Lambda(i)$.
7. **Selection**: if $\Lambda(i_2^*) \geq \gamma_{th}$ then choose $\mathcal{H}_{i_2^*}$ otherwise choose $\mathcal{H}_0$.

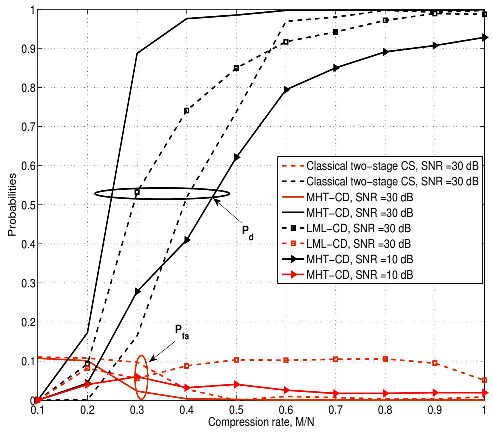
5. SIMULATION RESULTS

The proposed MHT-CD based on an exhaustive search and the proposed sub-optimal LML-CD based on a greedy search are tested in this section. For the simulations, the following parameters are considered: the number of channels is $N = 10$, and two fixed sensing matrices, $\mathbf{\Phi}$ are used, i) *Gaussian*: matrix whose elements are drawn i.i.d. from a random Gaussian distribution of zero mean and variance $\frac{1}{M}$, and ii) *multi-coset*: constructed by selecting M columns uniformly at random from $\mathbf{I}_N$ and multiplied with a normalization factor of $\sqrt{\frac{N}{M}}$. The *Gaussian* and *multi-coset* matrices are used in the simulations, as these are the standard sensing matrices used in the CS framework for recovery with ℓ_1-norm optimization. However, for direct detection, the required properties of the sensing matrices have to be analyzed and this is the subject of future work. The simulations are averaged over 1000 trials. In every trial, the vector $\mathbf{x}_f$ is randomly generated with i.i.d. Gaussian distributed entries of zero mean and variance according to a certain static channel occupancy (σ_x^2 for busy channels and zero for free channels). The vector $\mathbf{v}_f$ is generated with i.i.d. Gaussian distributed entries of zero mean and variance σ_v^2 in each trial. The time domain vectors $\mathbf{x}$ and $\mathbf{v}$ are then obtained using the DFT matrix $\mathbf{\Psi}$. The variances are set according to the required SNR, given by $10 \log_{10}\left(\frac{K\sigma_x^2}{N\sigma_v^2}\right)$ dB. A static channel occupancy is considered, with a busy channel set $\mathbf{b} = \{3, 4\}$, number of free channels $N - K = 8$, and number of busy channels $K = 2$. γ_{th} is chosen to keep the P_{fa} below 10% in order to have a fair comparison between different algorithms.

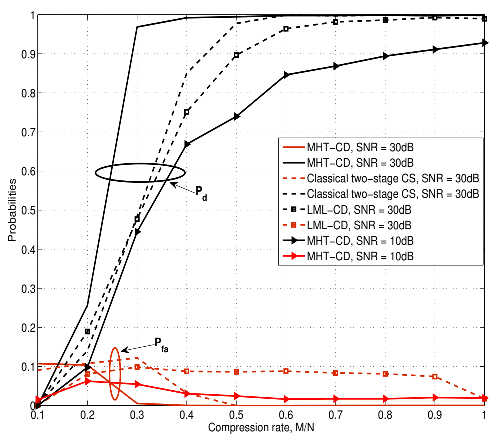
The proposed MHT based compressive wideband sensing is compared with the conventional approach for CS based spectrum sensing of [6]. Then the detection is performed on the frequency response estimate (power or amplitude). We make use of the Regularized Orthogonal Matching Pursuit (ROMP) [7] to solve the convex program in [6], and then perform energy detection to determine the occupancy. We make use of ROMP, as it has the speed of the greedy iterative methods (matching pursuit) and the robustness of ℓ_1-minimization. The complexity order of ROMP is $O(NMK)$. Additionally, in the second stage, the threshold based detection has a complexity order of $O(N)$.

Fig. 1a and Fig. 1b show the performance of the MHT based detector for *Gaussian* and *multi-coset* compression matrices, respectively. The detector MHT-CD based on exhaustive search is illustrated for SNR of 10 dB and 30 dB. MHT-CD performs better than the classical two-stage approach. A P_d of 0.9 is achieved for $\frac{M}{N} \approx 0.3$ with MHT-CD. Using the classical two-stage approach, a P_d of 0.9 is achieved for $\frac{M}{N} \approx 0.6$ and $\frac{M}{N} \approx 0.5$ with *Gaussian* and *multi-coset* compression matrices, respectively. In the case of the LML-CD algorithm, $P_d \approx 0.9$ is obtained for $\frac{M}{N} \approx 0.6$ and $\frac{M}{N} \approx 0.5$ for *Gaussian* and *multi-coset* sensing matrices, respectively. The performance of the cubic complexity LML-CD algorithm is similar to that of the classical two-stage approach.

Using a small N in the simulations reveals an important aspect of CS. In the CS literature, *Gaussian* or any other



(a) $\mathbf{\Phi}$ = *Gaussian* compression matrix.



(b) $\mathbf{\Phi}$ = *multi-coset* compression matrix.

Fig. 1: Performance of MHT detector based on exhaustive search (MHT-CD), classical estimation-detection two stage approach, and the sub-optimal LML-CD algorithm, with $N = 10$, $K = 2$, $S = 3$.

random matrices are suggested as a favorable choice for signal recovery with ℓ_1-norm optimization, but this choice holds mostly for $N \to \infty$ [8]. As can be seen in Fig. 1b, structured matrices like *multi-coset* matrices perform better in case of a smaller N, which appear more often in digital communications.

6. CONCLUSIONS

In this paper, we considered wideband spectral sensing in the multiple hypothesis sense. Each hypothesis corresponds to one of the possible combinations of all channels being busy and/or free. To reduce the complexity and capitalize on the sparsity of the spectrum, the sampling rate is reduced as in the CS framework. Here, we avoid the conventional CS based sensing, which usually involves reconstruction of the compressed signal before detection. Instead a direct detection based on the compressed samples is performed, resulting in a CD. The detector MHT-CD was developed and requires $O(2^N)$ computations. The detector MHT-CD performs better than the classical two-stage approach, but is impractical for large N. Hence, we proposed a sub-optimal greedy algorithm, LML-CD, based on the observed properties of the local maximum LLRs. The LML-CD is of complexity order $O(N^S)$, and has a performance comparable to that of the conventional two-stage approach of the complexity order $O(KMN)$. However, there is still a need for a lower complexity algorithm with a performance near to the MHT-CD and is a subject for future work. It should be noted that the exact knowledge of the sparsity level is not required for direct detection.

7. REFERENCES

[1] 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.

[2] M. Mishali and Y. Eldar, “From theory to practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals,” *IEEE Trans. Signal Process.*, vol. 4, no. 2, pp. 375 –391, april 2010.

[3] P. Stoica and Y. Selen, “Model-order selection: a review of information criterion rules,” *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36 – 47, july 2004.

[4] E. G. Larsson and Y. Selen, “Linear regression with a sparse parameter vector,” *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 451 –460, feb. 2007.

[5] Y. Sun, “A family of likelihood ascent search multiuser detectors: an upper bound of bit error rate and a lower bound of asymptotic multiuser efficiency,” *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1743 –1752, june 2009.

[6] Z. Tian and G. Giannakis, “Compressed sensing for wideband cognitive radios,” in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. ICASSP 2007.*, vol. 4, april 2007, pp. IV–1357 –IV–1360.

[7] D. Needell and R. Vershynin, “Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit,” *IEEE Trans. Signal Process.*, vol. 4, no. 2, pp. 310 –316, april 2010.

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