Compressed Sensing Bayes Risk Minimization for Under-determined Systems via Sphere Detection

Fabian Monsees, Carsten Bockelmann, Dirk Wübben, Armin Dekorsy

Department of Communications Engineering
University of Bremen, Germany
Email: {monsees, bockelmann, wuebben, dekorsy}@ant.uni-bremen.de

Abstract—The application of Compresses Sensing is a promising physical layer technology for the joint activity and data detection of signals. Detecting the activity pattern correctly has severe impact on the system performance and is therefore of major concern. In contrast to previous work, in this paper we optimize joint activity and data detection in under-determined systems by minimizing the Bayes-Risk for erroneous activity detection. We formulate a new Compressed Sensing Bayes-Risk detector which directly allows to influence error rates at the activity detection dynamically by a parameter that can be controlled at higher layers. We derive the detector for a general linear system and show that our detector outperforms classical Compressed Sensing approaches by investigating an overloaded CDMA system.

I. Introduction

Compressed Sensing (CS) is seen as a promising technology towards novel detectors, optimized for the detection of sparse signals even in under determined systems. Various CS algorithms formulate the problem of solving an under-determined set of equations as a convex optimization problem [1]. Involving a sparsity promoting term such as the l_1 norm guarantees that the algorithm converges and finds the sparsest solution to an under-determined system [2]. Besides that, various greedy approaches exist that solve under-determined problems iteratively. The most prominent ones are the Orthogonal Matching Pursuit (OMP) [3] and the Orthogonal Least Squares (OLS) [4].

In the field of communication technology, e.g, CS was successfully applied to OFDM systems for the reduction of the Peak to Average Power Ratio [5]. Furthermore, in [6] the authors have shown how novel CS techniques can be applied to Machine-to-Machine communication setups in order to detect activity and data jointly, thereby decreasing the required signaling overhead.

Joint detection of activity and data opens a wide field of application for CS algorithms in communications. Detecting sparse multi-user signals in a communication system is mostly motivated by assuming that some users are inactive and, thus, modeled as transmitting zeros, making the signal sparse. However, in such a scenario, the detection task is beyond finding the sparsest solution as pursued in CS. Rather, the detector has to estimate which users are active and which are inactive. The detection of the correct activity pattern, known as vector support in CS, is therefore crucial for successful Compressed-Sensing Multi-User Detection (CS-MUD). Another important point is that signals in communication obey a modulation alphabet which is generally a non-convex set. Accordingly, problems in communication have to be relaxed prior to solving them by common CS solvers.

With this application in mind, we would like to stress that wrongly detecting a user to be inactive is generally worse than doing the opposite, i.e., wrongly detecting a user to be active. In the first case information is lost, whereas higher layer error processing can help to identify erroneous transmissions in the latter case. The so called False Active and False Inactive rate is consequently of crucial importance for CS-MUD.

In this paper, we derive a Bayes-Risk detector for under-determined systems with non-convex finite alphabet constraints, which allows us to control the performance of the activity detection with respect to False Active and False Inactive errors. Moreover, this Bayes-Risk CS-MUD (BR-CS-MUD) can efficiently be implemented using common tree search algorithms such as the Sphere-Detector. As an example, we show the performance of our BR-CS-MUD detector by investigating an overloaded Code Division Multiple Access (CDMA) network, which can be expressed as an under-determined set of equations. The proposed detector utilizes the information about a finite modulation alphabet and finds the optimal solution by finding the closest lattice point. Utilizing this strategy outperforms common convex CS solvers such as the Basis Pursuit De-noising (BPDN) approach.

II. Compressed Sensing Bayes-Risk Detection

A. Sporadic Communication Model

In the following we shortly outline the sporadic communication model which forms the bases for the formulation of the Bayes-Risk Compressed-Sensing detector. In this model we aim to recover an unknown source vector $\mathbf{x}$ whose elements obey an augmented finite modulation alphabet $\mathcal{A}_0 = \mathcal{A} \cup \{0\}$. Here $\mathcal{A}$ stands for the modulation alphabet that is used for data transmission and can be, e.g., Binary Phase Shift Keying. The zero is used to model inactivity in the source vector which we assume to be of dimension $\mathbf{x} \in \mathcal{A}_0^K$. In a communication context, $\mathbf{x}$ can contain the symbols from different nodes in a multi-user scenario [6] or symbols from different antennas in a multiple antenna context. We further assume that the elements x_k of $\mathbf{x}$ are i.i.d and each element is non-zero with probability

p_a. Here p_a can be interpreted as an activity probability. With low p_a, many elements of $\mathbf{x}$ contain zeros, which makes the vector sparse. We summarize the channels between the nodes and the aggregation node with the matrix $\mathbf{T} \in \mathbb{R}^{M \times K}$. $\mathbf{y} \in \mathbb{R}^M$ represents the noise corrupted observation at the aggregation node. The canonical input-output relation of the system is given by

$$\mathbf{y} = \mathbf{Tx} + \mathbf{w}. \tag{1}$$

If $M < K$ the system refers to an under-determined system where the number of variables exceeds the number of observations. The noise vector $\mathbf{w} \sim \mathcal{N}\left(\sigma_n^2, 0\right)$ obeys an uncorrelated zero mean Gaussian distribution with variance σ_n^2. Henceforth, the task is to recover $\mathbf{x} \in \mathcal{A}_0^K$ from M noise corrupted measurements. To ease notation, we consider the communication system to be a real valued system which is not a general restriction. The scheme we propose can also be applied to complex systems using an equivalent real valued system description. With (1) we have a common noisy CS problem with a-priori information about the source vector.

B. Generalized Likelihood Ratio Test

With the previously introduced sporadic communication model, the goal of the detector is to reconstruct the symbols from the augmented modulation alphabet $\mathcal{A}_0$. In [7] the authors derived the MAP detector for an overdetermined system given the knowledge about the a priori distribution of the source vector $\mathbf{x}$. Our approach to recover $\mathbf{x}$ is twofold. First, we derive a detector that copes with under-determined systems given the a-priori knowledge of the source vector. Second, we formulate the detector such that the error rates for the activity detection can be tuned depending on the system by formulating the detection task in a Bayesian framework.

We start by considering the activity detection. For each element $\hat{x}_k$ the detector estimates whether the element was active, i.e., $\hat{x}_k \in \mathcal{A}$ or inactive, i.e., $\hat{x}_k = 0$. The possible outcomes for the activity detection task are summarized as follows. Table I shows the so called Confusion matrix of the

	$x_k \in \mathcal{A}$	$x_k = 0$
$\hat{x}_k \in \mathcal{A}$	True Active	False Active
$\hat{x}_k = 0$	False Inactive	True Inactive

TABLE I
CONFUSION MATRIX FOR ACTIVITY DETECTION

possible mappings between the true classes (element is active or inactive) and the hypothesized classes (element is active or inactive) [8]. The two error events are of special interest. If an element is wrongly estimated to be inactive, the data corresponding to this element is simply lost and cannot be recovered. However, in the the opposite case, when an element is wrongly estimated to be active, higher layer error detection methods can sort out wrong data packets. Thus, we state that these two possible error events have different impact on the overall system performance. With the knowledge that the two error classes have different impact, our approach is to derive a detector that can be controlled by an additional parameter

that balances the activity detection.
To include this balancing in the detector, we consider a binary hypothesis testing problem arising when determining whether an element is active or inactive.

$$H1 : x_k \in \mathcal{A} \rightarrow \text{Element is active}$$
$$H2 : x_k = 0 \rightarrow \text{Element is inactive}$$

With Tab. I we assign the following costs for wrong activity detection.
- C_{Fi} : Cost for estimating False Inactive
- C_{Fa} : Cost for estimating False Active

Costs for correct decisions are implicitly set to zero. We cast the problem as the minimization of the Bayes-Risk defined as

$$\mathcal{R} = C_{\text{Fa}} \text{Pr}\left(x_k = 0\right) \int_{Z_{\mathcal{A}}} p\left(y_m | x_k = 0\right) dy_m +$$
$$C_{\text{Fi}} \text{Pr}\left(x_k \in \mathcal{A}\right) \int_{Z_0} p\left(y_m | x_k \in \mathcal{A}\right) dy_m. \tag{2}$$

The Bayes-Risk given in (2) expresses the risk of a wrong activity detection. $p\left(y_m | x_k\right)$ denotes the density for an observation y_m under hypothesis x_k. $\text{Pr}\left(\cdot\right)$ is an event probability. Z_0 and $Z_{\mathcal{A}}$ are the regions in the observation space where the detector estimates the element to be inactive $\hat{x}_k = 0$ or active $\hat{x}_k \in \mathcal{A}$. For the sake of clarity we omit lengthy derivations and refer the reader to [9]. With the results from [10] we cast (2) as a Generalized Likelihood Ratio Test (GLRT) which has the following form

$$\frac{p\left(y_m | x_k = 0\right) \text{Pr}\left(x_k = 0\right) C_{\text{Fa}}}{\max_{x \in \mathcal{A}} p\left(y_m | x_k\right) \text{Pr}\left(x_k\right) C_{\text{Fi}}} \quad \overset{H2}{\underset{H1}{\gtrless}} \quad 1 \tag{3}$$

The GLRT in (3) evaluates the weighted a posteriori probabilities for both hypothesis and assigns the observation to the hypothesis with higher probability. Moreover, hypothesis $H1$ covers the set $\mathcal{A}$ and is therefore a so-called composite hypothesis. As (3) shows we compare the best model in $H1$ with $H2$. It can be seen that (3) can be rewritten as

$$\hat{x}_k = \arg \max_{x_k \in \mathcal{A}_0} p\left(y_m | x_k\right) \text{Pr}\left(x_k\right) C\left(x_k\right), \tag{4}$$

with

$$C\left(x_k\right) = C_{\text{Fi}}^{\mathbf{1}_{\mathcal{A}}(x_k)} C_{\text{Fa}}^{1 - \mathbf{1}_{\mathcal{A}}(x_k)}. \tag{5}$$

Here, $\mathbf{1}_{\mathcal{A}}\left(\cdot\right)$ is the indicator function which takes the value 1 if the argument is contained in the set $\mathcal{A}$. The solution to the optimization problem given in (4) is the optimal scalar estimate for x_k given a scalar observation y_m. Assuming i.i.d. observations and vector hypothesis $\mathbf{x}$, it can be shown that under application of the max-log approximation, (4) can be extended to

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \mathcal{A}_0} \log \prod_{m=1}^{M} p\left(y_m | \mathbf{x}\right) \prod_{k=1}^{K} \text{Pr}\left(x_k\right) C\left(x_k\right). \tag{6}$$

In the following we consider the system given in (1). The a posteriori probability for the element y_m with vector hypothesis $\mathbf{x}$ can now be expressed as

$$p\left(y_m | \mathbf{x}\right) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left[-\frac{1}{2\sigma_n^2}|y_m - \mathbf{T}_m \mathbf{x}|^2\right], \tag{7}$$

and

$$\prod_{m=1}^{M} p\left(y_m | \mathbf{x}\right) = \frac{1}{(2\pi)^{M/2}\sigma_n^M} \exp\left[-\frac{1}{2\sigma_n^2}\|\mathbf{y} - \mathbf{Tx}\|_2^2\right]. \quad (8)$$

Here $\mathbf{T}_m$ is the mth row vector of the matrix $\mathbf{T}$. In the following we assume that the elements of the source vector x_k from (1) obey a Bernoulli / Uniform distribution. In particular, we have

$$\Pr\left(x_k \in \mathcal{A}\right) = 1 - \Pr\left(x_k = 0\right) = p_a \quad (9)$$

$$p\left(x_k | x_k \in \mathcal{A}\right) = \frac{1}{|\mathcal{A}|} \quad (10)$$

and we write the a priori probability of the source vector $\mathbf{x}$ weighted with the costs from (5) as

$$\prod_{k=1}^{K} \Pr\left(x_k\right) C\left(x_k\right) =$$

$$\prod_{k=1}^{K} \left[C_{\mathrm{Fa}}\left(1 - p_a\right)\right]^{1-\mathbf{1}_{\mathcal{A}}(x_k)} \left(C_{\mathrm{Fi}}\frac{p_a}{|\mathcal{A}|}\right)^{\mathbf{1}_{\mathcal{A}}(x_k)}$$

$$= \left[C_{\mathrm{Fa}}\left(1 - p_a\right)\right]^{K-\sum_k \mathbf{1}_{\mathcal{A}}(x_k)} \left(C_{\mathrm{Fi}}\frac{p_a}{|\mathcal{A}|}\right)^{\sum_k \mathbf{1}_{\mathcal{A}}(x_k)}, \quad (11)$$

with (8) and (11) we can finally rewrite the optimization problem given in (6) to

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{A}_0}\|\mathbf{y} - \mathbf{Tx}\|_2^2 + 2\sigma_n^2 \sum_{k=1}^{K} \mathbf{1}_{\mathcal{A}}\left(x_k\right) \log\left(\frac{C_{\mathrm{Fa}}}{C_{\mathrm{Fi}}}\frac{1 - p_a}{p_a/|\mathcal{A}|}\right). \quad (12)$$

Moreover, the indicator function $\mathbf{1}_{\mathcal{A}}\left(x_k\right)$ can be replaced by the zero pseudo-norm defined as $\|\mathbf{x}\|_0 = \#\{x_k : x_k \neq 0\}$. We further express the ratio of costs as $\Omega = \frac{C_{\mathrm{Fa}}}{C_{\mathrm{Fi}}}$ and we obtain the non-convex optimization problem

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{A}_0}\|\mathbf{y} - \mathbf{Tx}\|_2^2 + 2\sigma_n^2\|\mathbf{x}\|_0\log\left(\Omega\frac{1 - p_a}{p_a/|\mathcal{A}|}\right)$$

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{A}_0}\|\mathbf{y} - \mathbf{Tx}\|_2^2 + \lambda\left(\Omega\right)\|\mathbf{x}\|_0. \quad (13)$$

Where the penalty term $\lambda\left(\Omega\right) = 2\sigma_n^2\log\left(\Omega\frac{1-p_a}{p_a/|\mathcal{A}|}\right)$ reflects the a priori assumption about the source vector $\mathbf{x}$ weighted by the parameter Ω. With $\Omega = 1$ and $M \geq K$ (13) equals to the MAP detector from [7]. The optimization problem (13) can be interpreted as a generalized MAP detector involving the Bayes-Risk for the activity detection. Additionally, Ω scales the a priori assumption and can be interpreted as an additional parameter that determines whether the detector is conservative $\Omega > 1$ or liberal $\Omega < 1$. A conservative detector will decide in favor of inactivity and will produce less False Active errors. In contrast a liberal detector will decide in favor of activity and produce less False Inactive errors than a conservative detector. The problem given in (13) is non-convex but can be implemented via a Sphere-Detector under the condition that $M \geq K$ holds. In the following we augment this detector for under-determined systems, i.e., $M < K$.

C. Constant Modulus Restriction

To solve (13) for $M < K$ efficiently, we restrict to constant modulus alphabets which allows us to replace the l_0 pseudo-norm by the l_p norm, i.e., $\|\mathbf{x}\|_0 = \|x\|_2^2 = \|x\|_p^p$ and we have

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{A}_0}\|\mathbf{y} - \mathbf{Tx}\|_2^2 + \lambda\left(\Omega\right)\|\mathbf{x}\|_p^p. \quad (14)$$

The penalty term $\lambda(\Omega)$ is directly influenced by Ω and can be negative for small values of Ω. We make use of the norm-invariance of the source vector $\mathbf{x}$ and expand (14) to an overdetermined system

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{A}_0}\|\mathbf{y} - \mathbf{Tx}\|_2^2 + \|\mathbf{x}\|_2^2 + \left[\lambda\left(\Omega\right) - 1\right]\|\mathbf{x}\|_0 \quad (15)$$

$$= \left\|\begin{bmatrix}\mathbf{y}\\\mathbf{0}_K\end{bmatrix} - \begin{bmatrix}\mathbf{T}\\\mathbf{I}_K\end{bmatrix}\mathbf{x}\right\|_2^2 + \underbrace{\left[\lambda\left(\Omega\right) - 1\right]}_{\Theta(\Omega)}\|\mathbf{x}\|_0. \quad (16)$$

If $\Theta(\Omega) \geq 0$, (16) can be directly implemented using a Sphere-Detector. One of the main prerequisites for the Sphere-Detector is an increasing metric in the sequence of detected symbols which is not guaranteed if $\Theta(\Omega) < 0$. However, $\Theta(\Omega)$ controls how non-zero symbols are penalized compared to no costs for zero symbols. For $\Theta(\Omega) < 0$ the cost for non-zero symbols are set to a negative value while zero symbols are still charged with no costs. To ensure that the penalty term remains positive, we re-write the optimization problem such that zero-symbols are penalized with positive costs. To proof this approach, we briefly review that the l_0 pseudo norm can be written as a sequence of indicator functions as $\|\mathbf{x}\|_0 = \sum_{k=1}^{K}\mathbf{1}_{\mathcal{A}}\left(x_k\right)$. If $\Theta(\Omega) < 0$, we can rewrite the optimization problem by subtracting the absolute value of the negative costs while adding the constant $|\Theta(\Omega)|K$. These two step do not change the optimization problem and we obtain for (16)

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{A}_0}\left\|\begin{bmatrix}\mathbf{y}\\\mathbf{0}_K\end{bmatrix} - \begin{bmatrix}\mathbf{T}\\\mathbf{I}_K\end{bmatrix}\mathbf{x}\right\|_2^2 - |\Theta\left(\Omega\right)|\|\mathbf{x}\|_0 + |\Theta(\Omega)|K$$

$$= \arg\min_{\mathbf{x}\in\mathcal{A}_0}\left\|\begin{bmatrix}\mathbf{y}\\\mathbf{0}_K\end{bmatrix} - \begin{bmatrix}\mathbf{T}\\\mathbf{I}_K\end{bmatrix}\mathbf{x}\right\|_2^2 + |\Theta\left(\Omega\right)|\left[K - \|\mathbf{x}\|_0\right]$$

$$= \arg\min_{\mathbf{x}\in\mathcal{A}_0}\left\|\begin{bmatrix}\mathbf{y}\\\mathbf{0}_K\end{bmatrix} - \begin{bmatrix}\mathbf{T}\\\mathbf{I}_K\end{bmatrix}\mathbf{x}\right\|_2^2 + |\Theta\left(\Omega\right)|\sum_{k=1}^{K}\mathbf{1}_0\left(x_k\right)$$

$$= \arg\min_{\mathbf{x}\in\mathcal{A}_0}\|\mathbf{y}' - \mathbf{T}'\mathbf{x}\|_2^2 + |\Theta\left(\Omega\right)|\sum_{k=1}^{K}\mathbf{1}_0\left(x_k\right). \quad (17)$$

Due to the application of $\mathbf{1}_0\left(\cdot\right)$, each zero symbol is penalized by $\Theta(\Omega)$. Moreover, the penalty term in (17) shows that the metric is monotonically increasing in the elements contained in $\mathbf{x}$ which allows direct implementation via a Sphere-Detector. We note that the augmented matrices $\mathbf{y}' \in \mathbb{R}^{M+K}$ and $\mathbf{T}' \in \mathbb{R}^{M+K\times K}$ constitute an overdetermined set of equations.

D. Compressed-Sensing Bayes-Risk Sphere-Detector

We here briefly rewrite the optimization problem (17) such that efficient implementation via a Sphere-Detector is possible. We apply the following procedure for solving (17) efficiently.

First, we apply the skinny [11][p. 217] QR decomposition on $\mathbf{T}'$ and obtain $\mathbf{T}' = \mathbf{QR}$ with $\mathbf{Q} \in \mathbb{R}^{M+K \times K}$ being a matrix with orthonormal columns and $\mathbf{R} \in \mathbb{R}^{K \times K}$ being an upper triangular matrix. With this decomposition of $\mathbf{T}'$, we rewrite the optimization problem (17) as

$$= \arg\min_{\mathbf{x} \in \mathcal{A}_0} \|\mathbf{y}' - \mathbf{QRx}\|_2^2 + |\Theta(\Omega)| \sum_{k=1}^{K} \mathbf{1}_0(x_k) \quad (18)$$

$$= \arg\min_{\mathbf{x} \in \mathcal{A}_0} \|\mathbf{Q}^T\mathbf{y}' - \mathbf{Q}^T\mathbf{QRx}\|_2^2 + |\Theta(\Omega)| \sum_{k=1}^{K} \mathbf{1}_0(x_k) \quad (19)$$

$$= \arg\min_{\mathbf{x} \in \mathcal{A}_0} \|\tilde{\mathbf{y}} - \mathbf{Rx}\|_2^2 + |\Theta(\Omega)| \sum_{k=1}^{K} \mathbf{1}_0(x_k) \quad (20)$$

Note that for a $\Theta(\Omega) \geq 0$, the optimization problem can be solved by the direct application of the QR decomposition on (16) [7][12].

III. PERFORMANCE EVALUATION

A. Setup

In the following, we exemplary show the performance of our detector by investigating an overloaded Code Division Multiple Access (CDMA) [13] system. In this model, we assume that in total K sensor nodes are connected to a central aggregation node. The nodes are active only occasionally with a probability factor p_a which is equal for all K nodes in the system. Active nodes transmit Binary Phase Shift Keying (BPSK) modulated symbols and we have $\mathcal{A} = \{\pm 1\}$. The symbols are spread by pseudo-noise sequences to chips by a factor of N which is equal for all nodes in the network. The chips are transmitted over a frequency selective Rayleigh fading channel with impulse response length of L_h which is also assumed to be of equal length for all nodes. Inactive nodes are modeled as transmitting zeros. The joint activity and data detection task is thus the detection of symbols from the augmented alphabet $\mathcal{A}_0 = \mathcal{A} \cup \{0\}$. In the following we write $\mathbf{x} \in \mathcal{A}_0^K$ as the source vector containing the symbols from the K nodes. To be consistent with the previously derived detector, we express the node specific spreading and convolution with the underlaying channel with the matrix $\mathbf{T} \in \mathbb{R}^{N+L_h-1 \times K}$. This allows us to formulate the symbol-rate CDMA system as

$$\mathbf{y} = \mathbf{Tx} + \mathbf{w}. \quad (21)$$

Here $\mathbf{y} \in \mathbb{R}^{N+L_h-1}$ is the received signal at chip-rate and $\mathbf{w}$ is the white uncorrelated noise vector with zero mean and variance σ_n^2. We apply a pre-whitening filter in order to ensure white Gaussian noise at symbol-rate. We neglect the details here, for further information the reader is referred to [14].

B. Symbol Error Rate

We investigate the performance of the detector by considering errors on the augmented alphabet $\mathcal{A}_0$ which we term Gross Symbol Errors (GSE) since activity and data errors are summarized in this class. For assessing the performance at the GSE, we start by investigating it for different degrees of overloading the system. Overloading is done by varying N while keeping K constant. The remaining simulation parameters are summarized in Table II

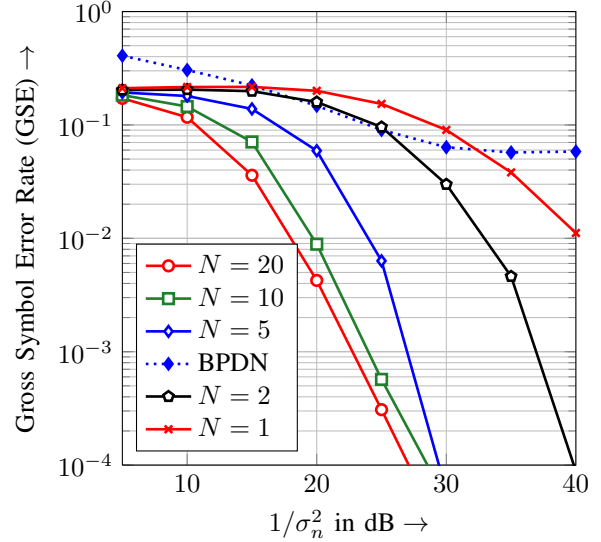| Simulation Parameters	
Number of Nodes	$K = 20$
Spreading Gain	$1 \leq N \leq 20$
Length of Channel Impulse Resp.	$L_h = 4$ chips
Channel Type	real valued Rayleigh Fading
Channel State Information	Perfect
Activity Probability	$p_a = 0.2$
Bayes Factor	$\Omega = 1$
Modulation Type	BPSK

TABLE II
SIMULATION PARAMETER

Fig. 1. Gross Symbol Error Rate parametrized by the length of the spreading sequence N. For $N = 5$, the dotted curve shows the performance of a Basis Pursuit De-noising for comparison.

Fig.1 shows the GSE for varying the spreading gain N. Most significantly, decreasing from $N = 20$ chips down to $N = 10$ chips results only in small losses. Overloading the system further, decreases the performance. However, even with $N = 1$, communication is still possible at high SNR. One should note that for $N = 1$, the system has the dimension 4×20 due to the channel. Additionally, the BR-CS-MUD performs exhaustive search among all possible hypothesis with an efficient search method. This explains the good performance of our detector. Additionally, the performance of the Basis Pursuit De-noising (BPDN) is shown for a system with a spreading gain of $N = 5$ for comparison. Comparing the performance of the BPDN with the BR-CS-MUD shows that the BPDN converges into a error floor even for high SNR. Our detector gains performance from the knowledge that the source vector obeys a constant alphabet with a known a priori distribution. In contrast, the BPDN relaxes the problem to a convex optimization problem prior to solving, which results in significant performance losses.

C. Performance of the Activity Detection

To assess the performance of the activity detection, we employ the so called Receiver Operating Characteristic (ROC)

curve [8]. The ROC evaluates the True Active and the False Active rates of the activity detection which are dependent on the SNR and the choice of Ω. An optimal detector will produce a True Active rate of 100% and a False Active rate of 0%. Fig. 2 shows the ROC for the setup summarized in Table II. The spreading gain is set to a fixed value of $N = 5$ for all nodes and the Bayes factor Ω is varied to influence the activity detection. For a Bayes factor $\Omega = 1$, the curve
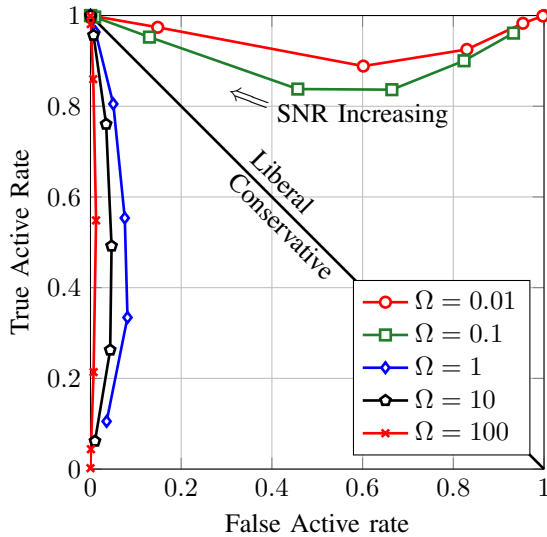

Fig. 2.   Reciever Operating Characteristic for the activity detection for an SNR range of $5\text{dB} - 40\text{dB}$

starts at the point $(0/0)$ which means that the detector neither correctly detects active nodes nor detects nodes wrongly to be active. Consequently, the detector estimates all nodes to be inactive. For increasing SNR (moving upwards on the curve), this detector increases the True Active rate while still meeting a very low False Active rate. This shows that the detector only detects nodes to be active if the particular probability that the node is active is very high, which can be interpreted as a conservative behavior of the detector. Increasing the Bayes factor to $\Omega = 10$ and $\Omega = 100$ makes the detector even more conservative and the False Active rate is kept very low at increasing SNR. At very high SNR, the detector reaches the $(1/0)$ point and no errors for the activity detection occur.

Decreasing the Bayes factor down to $\Omega = 0.1$ and $\Omega = 0.01$ promotes the detection of symbols instead of zeros. This setting tunes the detector such that the activity detection decides more likely in favor of activity instead of inactivity. At very low SNR, the curves start at the point $(1/1)$ and the detector estimates all nodes to be active, producing a False Active and True Active rate of 100%. The behavior for increasing SNR is analogous to the previous setting, the False Active rate decreases down to 0%. The True Active Rate also decreases slightly and reaches 100% for high SNR. Since the detector has the tendency to decide in favor of activity instead of inactivity, we term this a liberal detector. In a practical setup Ω can be tuned by higher layer applications such that certain False Active of True Active rates are met. Moreover, it might

be useful to scale Ω as a function of the current SNR.

IV. CONCLUSION

In this work, we introduced a novel Compressed Sensing Bayes-Risk Detector that minimizes the risk of an erroneous decision with respect to the activity detection. This detector allows reliable communication even in under-determined communication systems. As a degree of freedom the performance of the activity detection is controlled by an additional parameter. This provides a system depended flexible management of the False Active and False Inactive rate and can be adjusted by higher layers such that certain system specific error rates are met. We showed that our detector utilizes the finite alphabet constraint of the modulation alphabet as side information, which allows reliable communication even in highly under-determines systems. With this side information we our detector outperforms a classical Basis Pursuit De-noising approach.

REFERENCES

[1] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*.   Cambridge, U.K.: Cambridge Univ. Press, May 2012.

[2] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21 –30, Mar. 2008.

[3] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, nov 1993, pp. 40 – 44 vol.1.

[4] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of Control*, vol. 50, pp. 1873–1896, 1989.

[5] R. F. Fischer and F. Waeckerle, "Peak-to-Average Power Ratio Reduction in OFDM via Sparse Signals: Transmitter-Side Tone Reservation vs. Receiver-Side Compressed Sensing," *17th International OFDM Workshop 2012*, pp. 1 –8, Aug. 2012.

[6] H. F. Schepker and A. Dekorsy, "Sparse Multi-User Detection for CDMA transmission using greedy algorithms," in *8th International Symposium on Wireless Communication Systems (ISWCS)*, Aachen, Germany, Nov. 2011, pp. 291 –295.

[7] H. Zhu and G. Giannakis, "Exploiting sparse user activity in multiuser detection," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 454 – 465, Feb. 2011.

[8] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning*, 2nd ed.   Springer, Jul. 2011.

[9] H. L. van Trees, *Detection, estimation and modulation theory*.   Wiley, 1967.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing Detection Theory*, 1st ed.   Prentice Hall, 1993.

[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed.   The Johns Hopkins University Press, 1989.

[12] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2389 – 2402, oct. 2003.

[13] S. Verdú, *Multiuser Detection*.   Cambridge, U.K.: Cambridge Univ. Press, Nov. 1998.

[14] F. Monsees, C. Bockelmann, D. Wübben, and A. Dekorsy, "Sparsity aware multiuser detection for machine to machine communication," in *Second International Workshop on Machine-to-Machine Communication at IEEE Globecom 2012*, Anaheim, USA, Dec 2012.