Compressed Sensing Bayes-Risk Detection for Frame Based Multi-User Systems

Fabian Monsees, Carsten Bockelmann, Armin Dekorsy

Department of Communications Engineering

University of Bremen, Germany

E-mail: {monsees,bockelmann,dekorsy}@ant.uni-bremen.de

Abstract—**Performing joint activity and data detection has recently gained attention for reducing signaling overhead in multi-user Machine-to-Machine Communication systems. In this context, Compressed Sensing has been identified as a good candidate for joint activity and data detection especially in scenarios where the activity probability is very low. This paper augments activity and data detection for frame based multi-user uplink scenarios where nodes are (in)active for the duration of a frame. We propose a two stage detector which first estimates the set of active nodes followed by a data detector. Our detector outperforms symbol-by-symbol Maximum a posteriori detection.**

I. INTRODUCTION

Machine Type Communication (MTC) is a fast growing field which constantly raises new demands on existing communication systems. Compared to human driven communication, MTC is quite often sporadic and of very low data rate making extensive signaling and complex scheduling inappropriate [1], [2]. Recent research has shown how signaling and scheduling complexity can be reduced by physical layer joint activity and data detection [3]–[5]. Physical layer joint activity and data detection assumes a multi-user uplink network, where nodes sporadically transmit data to a central aggregation node for further processing. Given some knowledge about the node activity, the detector blindly estimates the set of active nodes and the data. The reliability of the activity detection is crucial for the data-detection as erroneous activity detection strongly impacts the overall performance. In [6], [7] the authors have derived the Bayes-Risk detector for symbol-by-symbol joint activity and data detection. This detector allows to control the error rates with respect to activity detection by an additional risk parameter such that certain system dependent error rate requirements are met.

This paper augments symbol-by-symbol Bayes-Risk detection by considering a multi-user uplink scenario where nodes transmit frames to a central aggregation node. For this setup, a two-stage detector is proposed where the first stage estimates the activity with respect to some pre-adjusted Bayes-Risk and the second stage estimates the data of the nodes. As this detection scheme utilizes a whole frame of observations to reliably estimate the activity, it outperforms symbol-by-symbol activity and data detection at the cost of additional complexity.

We start by formulating a general communication model involving a frame based MTC like uplink transmission with a low per-node activity probability in Section II. Section III-A contains the main part of this paper, where we introduce the two step detector for joint activity and data detection. In Subsection III-C the processing of estimated activity pattern is described, which is implemented via a sphere decoder. We show that both stages, i.e., activity detection and data detection, can utilize the same sphere decoder implementation as the two problem formulations are equivalent to each other. The performance of two stage detection is demonstrated in Subsection IV by comparing the error rates with respect to the activity detection for two stage detecting with symbol-by-symbol MAP detection. Moreover we show the impact of the Bayes-Risk factor on the activity detection by investigating the Receiver Operating Characteristic (ROC).

II. SYSTEM MODEL

We start by formulating a general Machine-to-Machine (M2M) uplink communication model where nodes desire to transmit a frame of data to a central aggregation node. The timing is assumed to be synchronized among the nodes. More specifically the system consists of K nodes that transmit a frame of L symbols to a central aggregation point. Motivated by smart-metering and industrial applications [8] each node transmits only occasionally and the probability that the kth node transmits is p_a. If a node is active, it transmits a frame of L consecutive data symbols modulated by a modulation alphabet $\mathcal{A}$, which can be, e.g., Phase shift keying (PSK). For the sake of simplicity modulation symbols in $\mathcal{A}$ are assumed to be distributed uniformly. If a node is inactive, the node does not access the medium and keeps silent for the duration of the whole frame. The detector models inactive nodes as transmitting zeros instead of modulation symbols and uses an augmented alphabet $\mathcal{A}_0 = \mathcal{A} \cup \{0\}$ for detection. This model interprets the set of nodes at any time instance as transmitting symbols from the augmented alphabet $\mathcal{A}_0$. Stacking the symbols of all users at time instance i, results in the multi-user signal vector $\mathbf{x}_i \in \mathcal{A}_0^K$. Concatenating L multi-user vectors of a frame yields the matrix $\mathbf{X} \in \mathcal{A}^{K \times L}$, where the kth row $\mathbf{x}_{\bar{k}}$ contains L consecutive symbols from node k. As a general model, we assume transmission over a wireless channel that can be expressed via the linear input-

output relation

$$\mathbf{Y} = \mathbf{TX} + \mathbf{N}. \tag{1}$$

Here, $\mathbf{Y} \in \mathbb{R}^{M \times L}$ is the matrix of observations, $\mathbf{T} \in \mathbb{R}^{M \times K}$ summarizes the channels from the nodes to the aggregation point and the medium access, e.g., Code division Multiple Access. The matrix $\mathbf{N} \in \mathbf{R}^{M \times L}$ contains white Gaussian noise samples with zero mean and variance σ_n^2. If $M < N$, the system (1) is under-determined like most common Compressed Sensing (CS) problems. In the following we mainly focus on the case where $M < N$ holds.

The previously introduced activity model allows to distinguish between two disjoint events modeling the node (in)activity on a frame basis, i.e., for each row vector denoted as $\mathbf{x}_{\bar{k}}$ we can have

$$\mathbf{x}_{\bar{k}} \in \mathcal{A}^L \to E1_k \quad \text{(Node is active)} \tag{2a}$$
$$\mathbf{x}_{\bar{k}} = \mathbf{0}^L \to E2_k \quad \text{(Node is inactive)} \tag{2b}$$

with event probabilities $\Pr(E1_k) = p_a$ and $\Pr(E2_k) = 1 - p_a$. Here $E1_k$ denotes the event that the kth node is active and $E2_k$ denotes that the kth node is inactive, respectively. Given a frame of observations, the goal is to estimate $E1_k$ or $E2_k \ \forall K$ correctly. In the following we analyze this activity detection and Table I shows the possible outcomes that occur while hypothesizing $E1_k$ and $E2_k$. Note that the two error events False Active and False Inactive also known as *False Alarm* and *Missed Detection* are of major concern as they differently affect the system performance. Having a high False

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

TABLE I
POSSIBLE OUTCOMES FOR ACTIVITY DETECTION

Inactive rate leads to a loss of data since the detector tends to decide more in favor of inactivity, on the other hand a high False Active rate complicates processing at higher layers as obviously wrong symbols are forwarded to higher layers. To this end, we proposed to augment joint activity and data detection by taking the risk for erroneous activity detection into regard resulting in the so called Bayes-Risk MAP detector introduced in [6], [7]. This detector takes cost for an erroneous activity detection into account which is done via the two factors

- C_{Fi} : Cost for False Inactive errors
- C_{Fa} : Cost for False Active errors.

As a major result of [6], the ratio of costs $\Omega = \frac{C_{\text{Fa}}}{C_{\text{Fi}}}$ steers the detector between being more liberal and decide more in favor of activity or to be more conservative and decide more in favor of inactivity. This paper aims at extending the symbol-by-symbol Bayes-Risk MAP detector to a frame based Bayes-Risk detector.

III. FRAME BASED ACTIVITY ESTIMATION

A. Overview of Frame Wise Processing

Given a frame of observations, this section focuses on estimating $E1_k$ and $E2_k$ for all nodes correctly. Previously,

MAP detection has been employed on a symbol-by-symbol basis to jointly estimate activity and data [4], [6]. MAP detection could be extended for frame based systems which is possible in principle but due to the enormous complexity infeasible with state of the art processing capabilities. Instead we make use of the fact that the events $E1_k$ and $E2_k$ are always true for L consecutive symbols and we have the events for the individual symbols

$$E1_k = E1_{k,1} = E1_{k,2} \cdots = E1_{k,i} \cdots = E1_{k,L} \tag{3a}$$
$$E2_k = E2_{k,1} = E2_{k,2} \cdots = E2_{k,i} \cdots = E2_{k,L}. \tag{3b}$$

Our approach is to estimate node (in)activity $E1_k$ or $E2_k$ for a frame by splitting the estimation problem up into L smaller problems where the solution to the ith problem gives the estimate $E1_{k,i}$ or $E2_{k,i}$. More specifically, we propose a two stage detector based on an activity estimation followed by a data detecting stage. The general structure is shown in Fig. 1. Starting from a frame of observations, some a priori information λ and the current Bayes-Risk Ω, the detector first decomposes the problem into L symbol-by-symbol estimation problems (2) and combines the results to obtain an estimate for $E1_k$ and $E2_k$. This information is the fed as a priori information to a data detector which provides an estimate of the data contained in the frames of active nodes.
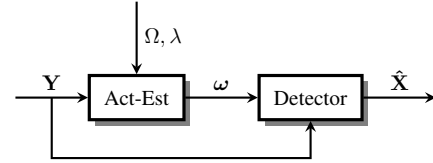

Fig. 1. Signal flow diagram of proposed two stage detector, which first estimates the set of active nodes and then estimates their data.

B. Activity Estimation

We start with the activity estimation by decomposing the frame-wise estimation problem into L individual problems by generalizing the Bayes-Risk MAP detector proposed in [6] for frame based MTC. The symbol-clock Bayes-Risk MAP performs the following Generalized Log-Likelihood Ratio Test (GLRT)

$$\text{cf. [6]} \quad \log \frac{p(\mathbf{y}_i|x_{k,i}=0)\Pr(x_{k,i}=0)C_{\text{Fa}}}{\max_{x \in \mathcal{A}} p(\mathbf{y}_i|x_{k,i})\Pr(x_{k,i})C_{\text{Fi}}} \underset{E1_{k,i}}{\overset{E2_{k,i}}{\gtrless}} 0 \tag{4}$$

in order to obtain a per-symbol estimate for the events $E1_{k,i}$ and $E2_{k,i}$. Note that the denominator of (4) constitutes a composite hypothesis for the kth node being active at time instance i involving all possible modulation symbols from the alphabet $\mathcal{A}$. As, the likelihood $p(\mathbf{y}_i|x_{k,i})$ is a marginal density (with respect to $x_{k,i}$) from the distribution $p(\mathbf{y}_i|\mathbf{x}_i)$ it has to be computed via marginalization. Subsequently, $\mathbb{X}_{x_k=\nu} = \{\mathbf{x} \in \mathcal{A}_0^K : x_k = \nu\}$ denotes the set of K dimensional vectors with elements from $\mathcal{A}_0$ where the kth element is equal to ν. This allows to express (4) explicitly as

$$\underbrace{\log \frac{\sum_{\mathbf{s} \in \mathbb{X}_{x_k=0}} p(\mathbf{y}_i|\mathbf{s})\Pr(\mathbf{s})}{\max_{\nu \in \mathcal{A}} \sum_{\mathbf{s} \in \mathbb{X}_{x_k=\nu}} p(\mathbf{y}_i|\mathbf{s})\Pr(\mathbf{s})}}_{\text{Activity LLR}} \underset{E1_{k,i}}{\overset{E2_{k,i}}{\gtrless}} \underbrace{\log \frac{1}{\Omega}}_{\text{decision threshold}}, \tag{5}$$

where $\mathbf{s}$ is the test-vector taken from the set $\mathbb{X}$ defined previously. Interestingly, the left side of (5) equals the calculation of an activity log-likelihood ratio (LLR) and the right side is the threshold determining when a node is considered to be active or inactive. This threshold is solely determined by the previously defined risk factors as $\Omega = \frac{C_{\text{Fa}}}{C_{\text{Fi}}}$. With (1), and the assumption of white Gaussian noise, the likelihood $p(\mathbf{y}_i|\mathbf{s})$ can be written as

$$p(\mathbf{y}_i|\mathbf{s}) \propto \exp\left(-\frac{1}{2\sigma_n^2}\|\mathbf{y}_i - \mathbf{Ts}\|_2^2\right). \qquad (6)$$

Considering (6), we see that (5) evaluates a sum including $\mathcal{O}(|\mathcal{A}^K|)$ summands which leads to a high computational effort, especially if K is large. The sum in (5) over exponential terms can be simplified by utilizing the so called max-log approximation [9] and thus turning the evaluation of the sum into an optimization problem yielding

$$\min_{\mathbf{s}\in\mathbb{X}_{x_k=0}} \left[\frac{1}{2\sigma_n^2}\|\mathbf{y}_i - \mathbf{Ts}\|_2^2 - \log(\Pr(\mathbf{s}))\right] - \qquad (7)$$

$$\min_{\nu\in\mathcal{A}}\min_{\mathbf{s}\in\mathbb{X}_{x_k=\nu}} \left[\frac{1}{2\sigma_n^2}\|\mathbf{y}_i - \mathbf{Ts}\|_2^2 - \log(\Pr(\mathbf{s}))\right] \underset{E1_{k,i}}{\overset{E2_{k,i}}{\gtrless}} \log\frac{1}{\Omega}.$$

Using the activity model (2) and the corresponding event probabilities, the a-priori probability for the vector hypothesis $\mathbf{s}\in\mathcal{A}_0^K$ reads

$$\log(\Pr(\mathbf{s})) = \log\left((1-p_a)^{K-\|\mathbf{s}\|_0}\left(\frac{p_a}{|\mathcal{A}|}\right)^{\|\mathbf{s}\|_0}\right)$$

$$\log(\Pr(\mathbf{s})) \propto -\|\mathbf{s}\|_0 \log\left(\frac{1-p_a}{p_a/|\mathcal{A}|}\right). \qquad (8)$$

Here, $\|\mathbf{s}\|_0$ is the l_0 norm and simply counts the number of non-zero zero elements contained in $\mathbf{s}$ [10]. Without changing the two optimization problems in (7) we subtract $\log\frac{1}{\Omega}$ from both minimizations. Moreover, inserting (8) into (7) and summarizing two $\min$ operations yields,

$$\min_{\mathbf{s}\in\mathbb{X}_{x_k=0}} \left[\frac{1}{2\sigma_n^2}\|\mathbf{y}_i - \mathbf{Ts}\|_2^2 - \log(\Pr(\mathbf{s})) + \log\Omega\right] - \qquad (9)$$

$$\min_{\mathbf{s}\in\mathbb{X}_{x_k\in\mathcal{A}}} \left[\frac{1}{2\sigma_n^2}\|\mathbf{y}_i - \mathbf{Ts}\|_2^2 - \log(\Pr(\mathbf{s})) + \log\Omega\right] \underset{E1_{k,i}}{\overset{E2_{k,i}}{\gtrless}} 0.$$

Both optimization problems in (9) address the same problem with different regions in the domains considered. Through the decision with respect to $E1_{k,i}$ and $E2_{k,i}$, both minimizations can be summarized into one while this implies a decision towards $E1_{k,i}$ or $E2_{k,i}$. Minimizing over $\mathbb{X}_{x_k=0}\cup\mathbb{X}_{x_k\in\mathcal{A}} = \mathcal{A}_0^K$ shows that the previously K different optimization problem now turn be be equal to each other. In turn, we can get the result of the estimated events $E1_{k,i}$ and $E2_{k,i}$ by considering the elements of the optimal argument $\hat{\mathbf{s}}$ that optimizes

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}\in\mathcal{A}_0^K} \|\mathbf{y}_i - \mathbf{Ts}\|_2^2 + \|\mathbf{s}\|_0 \underbrace{2\sigma_n^2 \log\left(\Omega\frac{1-p_a}{p_a/|\mathcal{A}|}\right)}_{\lambda}. \quad (10)$$

Here the kth element $\hat{s}_k$ indicates the estimation of $E1_{k,i}$ or $E2_{k,i}$ with respect to the test defined in (9). Moreover, λ

reflects the a priori knowledge that is available to the detector. The activity estimation is shown to be a direct application of per-node activity Log-Likelihood ratios (5) where the decision threshold separating between activity and inactivity is determined by the quotient of the Bayes-Risk factors and summarized in Ω.

Remark 1: The role of Ω can be seen very clearly by comparing (5) with (10). Assume $\Omega \gg 1$ then $\log\frac{1}{\Omega} \ll 0$ and the detector decides hypothesis $E2$ as long as the activity LLR is greater than $\log\frac{1}{\Omega}$, which will be the case as long as Ω is sufficiently large. In (10), Ω is linearly connected to the sparsity promoting l_0 norm and controls the cost for estimating non-zero symbols. if Ω is sufficiently large, the all zero vector will be always be the optimal value of the optimization problem. In summary, Ω allows to move the decision threshold for the activity Log-Likelihood between deciding more likely in favor of inactivity $\Omega > 1$ or deciding more likely in favor of activity $\Omega < 1$.

Having (10) solved for a whole frame allows to concatenate the vector estimates column wise to the matrix $\hat{\mathbf{S}} \in \mathcal{A}_0^{K\times L}$ where each row contains estimates for the events $E1_k$ and $E2_k$. If the k,ith element of $\hat{\mathbf{S}}$, $\hat{s}_{k,i}\in\mathcal{A}$, then the detector has hypothesized $E1_{k,i}$ if $\hat{s}_{k,i}=0$, the detector has hypothesized $E2_{k,i}$. Knowing that (in)active nodes repeat their selected activity pattern L times (2) suggests to process the rows of $\hat{\mathbf{S}}$ with a majority criterion for hypothesizing $E1_k$ or $E2_k$ on a per-frame basis. Estimation is done by making use of the indicator function $\mathbf{1}_\mathcal{A}(\cdot)$ which is one if the argument is contained in $\mathcal{A}$ and zero else. We define the majority criterion as

$$\omega_k = \begin{cases} \frac{1}{L}\lambda & \text{if} \quad \sum_{i=1}^{L}\mathbf{1}_\mathcal{A}(\hat{s}_{k,i}) \geq \frac{L}{2} \\ L\lambda & \text{if} \quad \sum_{i=1}^{L}\mathbf{1}_\mathcal{A}(\hat{s}_{k,i}) < \frac{L}{2} \end{cases}, 1 \leq k \leq K. \quad (11)$$

As (11) shows, we setup a vector $\boldsymbol{\omega}$ where the kth element is the scaled version of the a priori assumption λ and the scaling is dependent on the estimated node activity on a frame observation.

C. Data Detection

In the following we use the vector $\boldsymbol{\omega}$, providing information about the activity patterns, for estimating the data. This is done by solving a problem similar to (10) with modified penalty term

$$\hat{\mathbf{x}}_i = \arg\min_{\mathbf{x}\in\mathcal{A}_0^K} \|\mathbf{y}_i - \mathbf{Tx}\|_2^2 + \sum_{k=1}^{K}\mathbf{1}_\mathcal{A}(x_k)\omega_k \qquad (12)$$

As (12) shows, each element contained in $\mathbf{x}_k$ is charged with its individual prior ω_k and sparsity is promoted only for node k if ω_k is large. This data detector directly uses the result of the activity estimation for the data detection in a way that nodes that were identified to be most likely active are preferred to be considered to be active at data detection.

D. Implementation via Modified Sphere Decoder

In the following we first show how (10) can be implemented via sphere decoding, the implementation of (12) is analogous

and shown later. As we address under-determined systems in our analysis, the first step is to regularize the system matrix $\mathbf{T}$ to be fully or even over-determined. To do so, we restrict ourself to constant modulus modulation alphabets $\mathcal{A}$, i.e., $|x_i| = 1 \; \forall x_i \in \mathcal{A}$ such as PSK. With this restriction, it is possible to replace the penalty term based on the l_0 norm by any norm $\|\mathbf{s}\|_0 = \|\mathbf{s}\|_1 = \|\mathbf{s}\|_2^2 \Leftrightarrow |s_k| = 1 \; \forall k$ which reformulates (10) to

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathcal{A}_0^K} \|\mathbf{y}_i - \mathbf{T}\mathbf{s}\|_2^2 + \|\mathbf{s}\|_2^2 + \underbrace{[\lambda - 1]}_{\Theta} \|\mathbf{s}\|_0 \quad (13a)$$

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathcal{A}_0^K} \left\| \begin{bmatrix} \mathbf{y}_i \\ \mathbf{0}_K \end{bmatrix} - \begin{bmatrix} \mathbf{T} \\ \mathbf{I}_K \end{bmatrix} \mathbf{s} \right\|_2^2 + \Theta \|\mathbf{s}\|_0 \quad (13b)$$

Sphere decoding relies on an accumulating metric in the sequence of estimates symbols, which is clearly fulfilled by the l_2 and the l_0 norm. However, for $\Theta < 0$, (13) has to be reformulated. Without changing the optimization problem, (13) is rewritten such that the penalty term remains positive even if $\Theta < 0$ holds. The key point is that we subtract the magnitude of Θ and add the constant $|\Theta|K$ to the problem. After rearranging, $[K - \|\mathbf{s}\|]$ can be written as a sum of indicator functions $\mathbf{1}_0(\cdot)$, which returns a one if the argument is zero and a zero otherwise. For $\Theta < 0$ we obtain

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

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

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

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathcal{A}_0^K} \|\mathbf{y}_i' - \mathbf{T}'\mathbf{s}\|_2^2 + |\Theta| \sum_{k=1}^{K} \mathbf{1}_0(s_k), \quad (14)$$

where the penalty term is always positive. Comparing (14) and (13) shows that the penalty term in both problems can be summarized as

$$f(\mathbf{s}, \Theta) = \begin{cases} \sum_{k=1}^{K} \Theta \mathbf{1}_{\mathcal{A}}(s_k), & \text{if } \Theta \geq 0, \\ \sum_{k=1}^{K} |\Theta| \mathbf{1}_0(s_k), & \text{if } \Theta < 0, \end{cases} \quad (15)$$

yielding

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathcal{A}_0^K} \|\mathbf{y}_i' - \mathbf{T}'\mathbf{s}\|_2^2 + f(\mathbf{s}, \Theta), \quad (16)$$

where $f(\cdot, \cdot)$ is always positive and reflects the penalty term. Note that (16) can also be used for solving (12) by simply following the steps above and using $f(\mathbf{x}, \boldsymbol{\omega} - \mathbf{1}_K)$, where $\mathbf{1}_K$ is a vector containing only ones.

Applying the skinny QR [11] decomposition on the augmented system description (14) results in $\mathbf{T}' = \mathbf{QR}$ 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. Filtering the received vectors $\tilde{\mathbf{y}}_i = \mathbf{Q}\mathbf{y}_i$ leads to

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathcal{A}_0^K} \|\tilde{\mathbf{y}}_i - \mathbf{R}\mathbf{x}\|_2^2 + f(\mathbf{x}, \Theta). \quad (17)$$

The problem (17) is straightforward to solve by common sphere decoder implementations. For the sake of space reasons we omit the algorithmic description of the utilized Sphere Decoder and refer the reader to [12].

IV. SIMULATIVE RESULTS

A. Setup

In the following, we investigate the performance for an exemplary multi-user uplink system. In particular, we investigate a Multi-User Code Division Multiple Access (CDMA) System, where a set of K nodes accesses the wireless channel via CDMA transmitting data to a central aggregation node [13]. As stated in Section II, we assume a sporadic synchronous frame based medium access of the nodes parametrized by the per node activity probability p_a. In the following we further assume that the nodes are synchronous at chip level, which is not a general restriction as shown in [14]. Moreover perfect Channel State Information is assumed at the aggregation node, which could, e.g., be obtained via a training phase at the beginning of each frame or by random coding inspired techniques as introduced in [15]. Each node spreads its modulation symbols to a CDMA chip sequence of length N. These chips are transmitted over a frequency selective Rayleigh fading channel with length of L_h chips. We model the input-output relation for this system by (1), where $\mathbf{T} \in \mathbb{R}^{N+L_h-1 \times K}$ summarizes spreading and convolution with the underlying frequency selective channel.

B. Performance Analysis

The performance is investigated in a fully loaded $N = K$ and overloaded $N < K$ CDMA setup. For comparison, we also apply symbol-by-symbol Bayes-Risk MAP detection from [6]. The parameters of the simulation setup are summarized in Table II. Fig. 2 compares the overall symbol

Simulation Parameters	
Number of Nodes	$K = 20$
Frame-length	$L = 100$ symbols
Spreading Gain	$1 \leq N \leq 20$
Length of Channel Impulse Resp.	$L_h = 4$ chips
Channel Type	real valued block Rayleigh Fading
Activity Probability	$p_a = 0.2$
Modulation Type	BPSK

TABLE II
SIMULATION PARAMETER

error rate of the Bayes-Risk two step detector with symbol-by-symbol Bayes-Risk detection. The overall symbol error rate summarizes activity and data errors, i.e., $\hat{x} \neq x$. We see that overloading the system from $N = 20$ to $N = 10$ leads to negligible performance losses at the aggregation node. Fig. 2 also shows the extreme case $N = 3$ where the system is highly overloaded, leading to significant losses but, however, the detector is still capable of estimating the activity and the data of the nodes correctly. The performance gain of performing frame based detection is approx. 3dB compared to symbol-by-symbol Bayes-Risk MAP detection for $N = 10$. The performance gains of two step processing compared to symbol-by-symbol
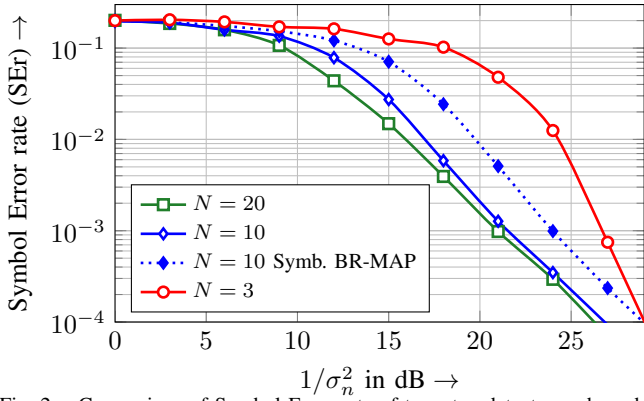
Fig. 2. Comparison of Symbol Error rate of two step detector and symbol-by-symbol Bayes-Risk MAP detector at $\Omega = 1$
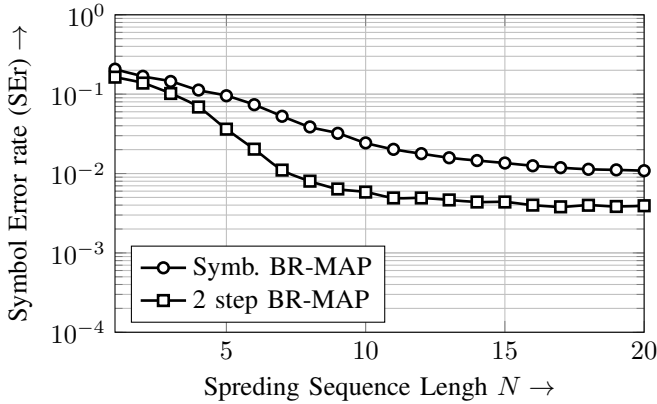

Fig. 3. Comparison of Activity Error rates at SNR = 18 dB for varying spreading sequence length N at $\Omega = 1$

MAP over N can be observed in Fig. 3. The error rates are shown at a fixed SNR = 18 dB while the spreading sequence length varies from $1 \leq N \leq 20$. Two step based detection outperforms symbol-by-symbol MAP detection for all simulated spreading sequence length N. Additionally, two step detection converges faster to a lower error rate. In the following we investigate the impact of the Risk quotient Ω by investigating the Receiver Operating Characteristic (ROC). With respect to Table I, ROC plots the True Active Rate vs. the False Active Rate for different SNRs. The optimal detector reaches the point $(0|1)$ at high SNR. Of specific interest is the path from low SNR to $(0|1)$ showing if the detector allows a low False Active rate (FAr) and can be considered to be conservative or if the detector is more liberal having a high True Active rate (TAr) at the cost of a also high FAr. We see that while changing Ω, different detector profiles can be achieved and in a practical system Ω has to be controlled by higher layers such that certain error rates are met. Moreover it might be useful to change Ω over the SNR to adaptively choose optimal performance.

V. CONCLUSION

This paper derives a two stage detector that estimates user activity in a multi-user framework based on frame-wise Bayes-Risk MAP detection. This detector takes the risk for erroneous activity detection as a free parameter into account
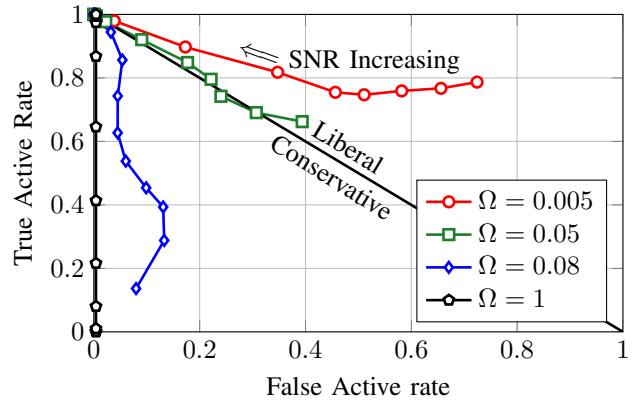

Fig. 4. Receiver Operating Characteristic (ROC) parametrized by Bayes-Risk quotient Ω

and outperforms symbol-by-symbol detection by approx. 3 dB.

REFERENCES

[1] P. Jain, P. Hedman, and H. Zisimopoulos, "Machine type communications in 3gpp systems," *IEEE Communications Magazine*, vol. 50, no. 11, pp. 28–35, November 2012.

[2] V. Misic, J. Misic, and D. Nerandzic, "Extending LTE to support machine-type communications," in *IEEE International Conference on Communications (ICC), 2012*, Ottawa, Canada, June, pp. 6977–6981.

[3] 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.

[4] 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.

[5] 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.

[6] F. Monsees, C. Bockelmann, D. Wübben, and A. Dekorsy, "Compressed sensing bayes risk minimization for under-determined systems via sphere detection," in *77th IEEE Vehicular Technology Conference (VTC2013-Spring)*, Dresden, Germany, Jun 2013.

[7] F. Monsees, C. Bockelmann, and A. Dekorsy, "Joint activity and data detection for machine to machine communication via bayes risk optimization," in *14th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Darmstadt, Germany, Jun 2013.

[8] B. Andrea, M. Dohler, J. Hernndez-Serrano, A. Kountouris, and D. Barthel, "Low-Power Low-Rate goes Long-Range: the case for secure and cooperative Machine-to-Machine communications," in *NETWORK-ING 2011 Workshops*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, vol. 6827, pp. 219–230.

[9] C. Studer, M. Wenk, A. Burg, and H. Bolcskei, "Soft-output sphere decoding: Performance and implementation aspects," in *Fortieth Asilomar Conference on Signals, Systems and Computers, 2006. ACSSC 2006.*, Pacific Groove, USA, Nov 2006, pp. 2071–2076.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[11] G. H. Golub and C. F. V. 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, November 1998.

[14] H. Schepker, C. Bockelmann, and A. Dekorsy, "Coping with CDMA asynchronicity in compressive sensing multi-user detection," in *77th Vehicular Technology Conference (VTC2013-Spring)*, Dresden, Germany, Jun 2013.

[15] M. Asif, W. Mantzel, and J. Romberg, "Channel protection: Random coding meets sparse channels," Oct 2009, pp. 348 –352.