

RESEARCH ARTICLE

Compressive Sensing based Multi-User Detection for Machine-to-Machine Communication

C. Bockelmann*, H. F. Schepker, A. Dekorsy

Department of Communications Engineering, University of Bremen, Germany

ABSTRACT

With the expected growth of Machine-to-Machine (M2M) communication, new requirements for future communication systems have to be considered. More specifically, the sporadic nature of M2M communication, low data rates, small packets and a large number of nodes necessitate low overhead communication schemes that do not require extended control signaling for resource allocation and management. Assuming a star-topology with a central aggregation node that processes all sensor information one possibility to reduce control signaling is the estimation of sensor node activity.

In this paper, we discuss the application of greedy algorithms from the field of Compressive Sensing in a channel coded Code Division Multiple Access context to facilitate a *joint* detection of sensor node activity and transmitted data. To this end a short introduction to Compressive Sensing theory and algorithms will be given. The main focus, however, will be on implications of this new approach. Especially, we consider the activity detection, which strongly determines the performance of the overall system. We show that the performance on a system level is dominated by the missed detection rate in comparison to the false alarm rate. Furthermore, we will discuss the incorporation of activity-aware channel coding into this setup to extend the physical layer detection capabilities to *code-aided* joint detection of data and activity. Copyright © 2013 John Wiley & Sons, Ltd.

*Correspondence

C. Bockelmann, University of Bremen, Department of Communications Engineering, Otto-Hahn-Allee 1, D-28359 Bremen, Germany.
Email: bockelmann@ant.uni-bremen.de

1. INTRODUCTION

1.1. M2M communication

Machine-to-Machine (M2M) communication is expected to grow tremendously in the next years [1], thereby posing new challenges for both existing communication systems, designed with human communication in mind, and existing M2M solutions alike. On the one hand, standards like LTE have to be extended to cope with the requirements of M2M communication, which often differ strongly from requirements for speech communication and human data access. For example, management overhead, such as mobility management or sophisticated resource allocation, is not well suited for low-rate communication of sensor nodes that are only transmitting sporadically. In the current frame structure of LTE-Advanced, at most a few tens of M2M devices can be supported with control channel elements (CCEs) allocated within a sub-frame [2]. This is far below the target number of M2M devices envisioned for the future. On the other hand, a growing number of nodes will be challenging for existing M2M standards like the IEEE standards 802.15.4 and 802.15.1. Both are

based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) and spread spectrum techniques. A thorough overview of current activities in different standardization bodies like ETSI and 3GPP can be found in [1]. Furthermore, detailed requirements on future M2M systems and an assessment of current technologies can be found in the deliverable D2.1 of the European project EXALTED [3].

In order to efficiently cope with a high number of sensor nodes, which are expected to only sporadically transmit small packets with low data rates, new approaches on the physical layer are required. Low control signal overhead and increased power efficiency are a major concern, for long battery life of low-cost sensors. Especially the latter is of great importance for future sensor networks, and is researched in several areas from energy efficient circuits to routing protocols [4]. Here, we focus on the physical layer aspects. One approach for decreasing the overhead is to avoid control signaling regarding the activity of sensor nodes before transmission, e.g., mobility tracking and duty cycle controlling. Another approach is to facilitate an automatic detection of active nodes. Accordingly, a reliable joint data *and* node activity detection has to be achieved,

which substitutes functionality usually implemented on higher layers. A fundamentally new approach to this problem statement is Compressive Sensing Multi-User Detection (CS-MUD).

Considering a *star topology*, where sensors communicate with a central aggregation node, high processing power allows for Multi-User detection. Due to the assumption of sporadic transmission, only a small number of sensors are active at any given time. Therefore, the Multi-User signal composed of signals from all sensors in the network is sparse. Compressive Sensing (CS) theory describes the estimation of sparse signals, or equivalently the solution of linear equation systems [5, 6] for sparse signals. With respect to communications, two properties of CS are especially interesting in M2M scenarios: *both data and activity can be reconstructed at the same time and reconstruction is possible for under-determined equation systems*. The first point is very interesting to implement the proposed activity detection on the physical layer, whereas the importance of the second point becomes clear in the context of the Code Division Multiple Access (CDMA) system that will be assumed as an example for M2M communication in this paper. Any CDMA Multi-User chip-rate model with sporadic activity can be approached by means of CS algorithms [7]. Regarding the second point, however, overloaded CDMA systems, i.e., the number of nodes in the network is larger than the spreading factor, lead to under-determined equation systems whose solution requires sparsity, i.e., sporadic activity. Overall, CDMA is very attractive for M2M scenarios, due to its adaptive and flexible support of different data rates as well as Qualities of Service, flexible and scalable number of supported devices, and low rate channel coding.

1.2. Compressive Sensing

The application of CS theory is a growing research field for many research areas, e.g., distributed signal processing [8] and radar [9]. In communications it is a very recent approach with much potential for further research. In literature, the focus has been on the application of CS theory to communication problems, where some amount of sparsity can be assumed, e.g., in channel estimation [10, 11], for ultra-wideband communication [12] and sensor networks [13, 14, 15, 7]. This section provides a brief overview of CS assumptions, reconstruction approaches and algorithms. For a thorough overview of CS see, e.g., [16].

Generally speaking, CS is concerned with efficient sampling assuming the original data is compressible. Formally, this can be described in the following way: Let \mathbf{z} be a compressible signal with respect to the basis Φ , such that $\mathbf{z} = \Phi\mathbf{x}$ and \mathbf{x} is sparse, i.e., only few entries are non-zero. Note that for a particular family of signals the basis Φ may not be unique, e.g., natural pictures are compressible with respect to many bases, such as Fourier or wavelets, see [17]. However, the basis Φ is always determined by

the physical constraints of the signal. In the following, we refer to Φ as the *sparsity basis*.

The data acquisition process of \mathbf{z} is modeled by determining the correlation of \mathbf{z} to a system Ψ . Further, we incorporate the possibility of inexact measurements into the data acquisition model. Hence, the available information of the physical signal \mathbf{z} is given by

$$\mathbf{y} = \Psi\mathbf{z} + \mathbf{n}, \quad (1)$$

where \mathbf{n} is a noise vector with known stochastic properties. Therefore, the length of \mathbf{y} is determined by the dimensions of Ψ .

The usual task of Compressed Sensing is to find a good *measurement matrix* Ψ for a given sparsity basis Φ . Here, the objective is not only to guarantee good recovery properties for the sparse signal, but to ensure those properties with as few measurements as possible. This allows to use less measurements than Nyquist sampling leading to under-determined equation systems with less equations than variables. The minimum amount of measurements required for reliable reconstruction depends on the sparsity of \mathbf{x} and the dimensions of Φ [18].

Reconstruction for CS is based on finding the sparse description \mathbf{x} of the physical signal \mathbf{z} , i.e., a sparse solution $\tilde{\mathbf{x}}$ that fulfills (1). To obtain such a sparse solution, multiple approaches are discussed in literature, which are principally based on the formulation of well suited optimization problems. The basic CS reconstruction problem is

$$\min_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \Psi\Phi\mathbf{x}, \quad (2)$$

where $\|\cdot\|_0$ denotes the l_0 -pseudo norm that counts the number of non-zero entries in the vector. Problem (2) aims to find a suitable solution $\tilde{\mathbf{x}}$ with as few non-zero entries as possible. Note, that (2) is known to be non-convex and NP-hard to solve, which motivates relaxation approaches. A common relaxation of (2) in the presence of noise is the so-called Basis Pursuit De-Noising (BPDN) [19], where the non-convex l_0 -pseudo norm is relaxed to the convex l_1 -norm. Another approach, stemming from the regression point of view, is the so-called Least Absolute Shrinkage and Selection Operator (LASSO) [20], where the l_1 -norm is used as a penalty term to encourage sparse solutions for the Least Squares (LS) estimation problem. Furthermore, regularization of the LASSO problem by an additional weighted l_2 -norm to enhance convergence for badly conditioned matrices \mathbf{A} is then termed Elastic-Net [21]. For the relaxations of (2) (BPDN/LASSO, Elastic-Net), there are many different approaches and algorithms, which aim to efficiently determine the solution of the general CS reconstruction problem. A thorough overview of the different algorithmic approaches, however, is beyond the scope of this paper.

In the area of sensor networks, two directions can be differentiated: the application of known CS reconstruction algorithms for continuous signals and the extensions

to finite, discrete alphabets. The first investigates the applicability of CS detectors in a communications context [7], and the second aims at new formulations for known detectors, like sparsity-aware CS MAP detection [13]. Such CS MAP detectors extend known approaches, like the sphere detector, to detect the transmitted symbols and non-active nodes. Furthermore, block-detection strategies have been devised to exploit the fact that nodes usually transmit multiple symbols, which implies activity over defined time intervals. Such structure helps to detect node activity more accurately [13, 15, 22].

For the scenario discussed in the following, the CS MAP approach is not applicable as it is numerically too demanding. Here, we focus on efficient CS algorithms that supply continuous estimates for further processing. Therefore, the greedy Group Orthogonal Matching Pursuit (GOMP) detector, which exploits block-information, will be detailed in Section 3.

1.3. Paper Focus and Contribution

In this paper, we provide a detailed introduction to CS based joint Multi-User data and activity detection for CDMA uplink transmission transferring the ideas from CS theory to M2M communication. Special attention is on the application and influence of channel codes in such CS detected CDMA systems. Based on the CDMA chip-rate Multi-User description [23], we will explain how CS algorithms can be applied to *jointly* detect the activity of sensor nodes and their data reducing required control signaling overhead. Furthermore, for the long term practical application of CS-based Multi-User detection schemes, a wider system context has to be considered, which focuses on the impact of CS detectors on other system components. Thus, we provide insight about the behavior of the activity detection, which strongly influences the impact of CS detection. To this end, we extend the analysis of CS detectors beyond the Symbol Error Rate (SER), which has been the only quality indicator in previous publications [7, 13]. In contrast, we also illustrate different activity error events. Specifically, we compare the combination of a computationally efficient greedy CS algorithm to a baseline system employing energy detection and LS filtering that implements *separate* activity and data detection. Additionally, on a frame level channel coding influences the performance of both approaches and offers another opportunity for enhanced activity detection. To this end, we propose to augment the channel decoding for enhanced activity detection to achieve *activity-aware decoding*.

1.4. Organization and Notation

The remainder of this paper is organized as follows: Section 2.1 introduces the general M2M scenario that is considered, and which is mathematically summarized in Section 2.2. Section 3 discusses strategies for separate as well as joint activity and data detection. Furthermore, channel coding in combination with the presented CS

Multi-User detectors is the topic of Section 4. Then, numerical results are discussed in detail in Section 5 and finally a conclusion is given.

Throughout this paper, the following notation is employed: G is a set of indices of groups and \bar{G} is the complementary set. Further, $\Gamma(G)$ specifies the indices of those columns of a matrix \mathbf{A} contained in the groups indexed by G . Furthermore, $\mathbf{A}_{\Gamma(G)}$ specifies the sub-matrix that only contains those columns with indices in $\Gamma(G)$, and likewise $\mathbf{x}_{\Gamma(G)}$ contains only those elements of \mathbf{x} with indices in $\Gamma(G)$. Additionally, \mathbf{x}^ℓ , \mathbf{A}^ℓ and G^ℓ specify the respective variable during the ℓ^{th} iteration. Herein, \mathbf{A}^\dagger is the Moore-Penrose pseudoinverse of \mathbf{A} , and \mathbf{A}^H the Hermitian matrix of \mathbf{A} .

2. SYSTEM DESCRIPTION

2.1. Machine-to-Machine Scenario

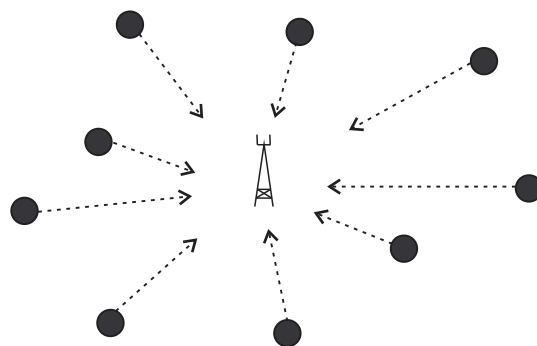


Figure 1. Sensor network - star topology

In this paper, we consider a M2M scenario, where K sensor nodes communicate with a central aggregation node, as shown in Fig. 1, typically denoted as *star topology*. All sensor nodes are devices of low complexity, while the central node allows for advanced signal processing, such as sophisticated detection and message aggregation for forwarding. In principle, the sensor nodes are only active on occasion, i.e., measurements have to be transmitted regularly or event driven, which leads to *sporadic communication*. To model sensor node activity, we adopt a statistical approach, where each sensor node is active for a short time period, with a given activity probability p_a . This activity probability, which we assume to be identical for all sensor nodes, determines the number of active nodes in a statistical sense. In the following, the base assumption is that out of the K sensor nodes in the network the majority are silent, i.e., the activity probability is rather small with $p_a \ll 1$. Given a large number of nodes K , this is a valid assumption for practical applications.

To facilitate sporadic and simultaneous medium access, CDMA with Pseudo random Noise (PN) sequences is used instead of the CSMA/CA scheme employed in IEEE standards. Here, we choose PN sequences due to their well

known correlation properties for uplink communication. CDMA offers a number of attractive properties for M2M communication as noted before. It allows for a very flexible non-orthogonal medium access with some degree of asynchronicity on the chip level between sensor nodes, which is well suited for the low overhead communication that is the goal of efficient M2M communication schemes.

2.2. CDMA Chip-rate System Model

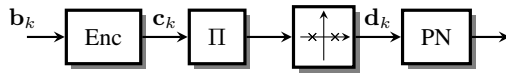


Figure 2. Sensor node transmitter model of sensor k

Based on the previous section, we formulate a mathematical model of the sporadic sensor node communication for further analysis. The basic transmitter setup of the sensor nodes is depicted in Fig. 2. We assume that an active node k_a transmits a data frame of N_b information bits $\mathbf{b}_{k_a} \in \{0, 1\}^{N_b}$. The information is encoded by a channel code of code rate R_c for error protection, yielding a length N_c code word vector $\mathbf{c}_{k_a} \in \{0, 1\}^{N_c}$. After random interleaving, the code bits are mapped to symbols yielding a symbol vector $\mathbf{d}_{k_a} \in \mathcal{A}^{N_d}$ that contains the symbols of a frame, where \mathcal{A} denotes the alphabet. In the following, we will restrict the description to BPSK and, thus, the overall system description is real-valued. An extension to QAM modulation alphabets is straightforward, e.g., by widely linear processing. Note that the activity probability p_a determines activity on a frame level, i.e., a sensor node is active or inactive for multiples of $N_d = N_c$ symbols. Consequently, the symbol vector of an inactive node k_i is modeled as all-zeros, $\mathbf{d}_{k_i} = \mathbf{0}$. Then the so-called *augmented* alphabet $\mathcal{A}_0 = \{\mathcal{A} \cup 0\}$, which is the modulation alphabet \mathcal{A} augmented and extended by the zero symbol that indicates inactivity or “no data”, describes all possible transmit symbols.

With CDMA as medium access scheme, each sensor node k uses a real-valued PN sequence $\mathbf{s}_k \in \{\pm 1\}^{N_s}$, to spread each of its information symbols $d_{k,i}$ with $i = 1, \dots, N_c$ to N_s chips. Without loss of generality, we assume that the spreading factor N_s , i.e., the number of chips per information symbol, is identical for all sensor nodes. Therefore, the amount of symbols in a fixed frame length will be identical for all nodes.

We assume that sensor nodes switch activity on the same symbol time basis, i.e., they are either active for the next N_c symbols or inactive (frame synchronous), which is illustrated in Fig. 3. In a practical system this has to be ensured, e.g., by a synchronization signal from the aggregation node, which is beyond the scope considered here. However, it is still possible that the chips transmitted at a synchronous symbol time are not received at the same chip time in this uplink scenario, thus yielding an asynchronicity of a few chips at the aggregation node. Such typical CDMA uplink considerations can be easily

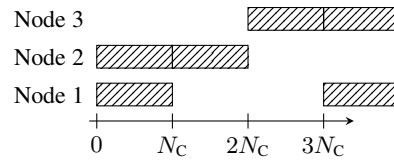


Figure 3. Example for synchronous frame activity from the sensor node perspective

incorporated in the model by delayed versions of the spreading sequences similarly to [24]. In order to simplify the notation, we assume that all frames are received synchronously at the aggregation node. Furthermore, the spreaded symbols of each node are distorted by a sensor node specific frequency selective channel $\mathbf{h}_k \in \mathbb{R}^{L_h}$ of length L_h which is constant for a whole frame the received vector at the aggregation node is

$$\mathbf{y} = \sum_{k=1}^K \mathbf{H}_k \mathbf{S}_k \mathbf{d}_k + \mathbf{n}, \quad (3)$$

where

$$\mathbf{S}_k = \begin{pmatrix} \mathbf{s}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{s}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots \end{pmatrix} \quad (4)$$

denotes the spreading matrix of user k and

$$\mathbf{H}_k = \begin{pmatrix} h_{k,0} & 0 & \dots \\ h_{k,1} & h_{k,0} & \\ \vdots & \vdots & \ddots \\ h_{k,L_h-1} & h_{k,L_h-2} & \\ 0 & h_{k,L_h-1} & \\ 0 & 0 & \ddots \end{pmatrix} \quad (5)$$

is the convolution matrix of the channel of user k and additive white Gaussian noise $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2)$ superimposes the signal at the receiver. The chip-rate model (3) includes Inter-Symbol-Interference (ISI) within one frame, due to the frequency selective channel. For simplicity, we do not capture frame start and end processing in this model. Thus, we omit the last $L_h - 1$ rows of \mathbf{H}_k and set the matrix to be $\mathbf{H}_k \in \mathbb{R}^{N_s N_c \times N_s N_c}$.

Based on [23], equation (3) can be written as a Multi-user chip-rate model as

$$\mathbf{y} = \mathbf{A} \mathbf{x} + \mathbf{n}, \quad (6)$$

where the matrix $\mathbf{A} \in \mathbb{R}^{N_s N_c \times K N_c}$ models the summation in (3), combining spreading and channel influences. The vector $\mathbf{x} \in \mathbb{R}^{K N_c}$ contains the data from all nodes, i.e., it is the stacked vector of all node frames \mathbf{d}_k

$$\mathbf{x} = [\mathbf{d}_1^T, \dots, \mathbf{d}_K^T]^T. \quad (7)$$

3. ACTIVITY AND DATA DETECTION

3.1. Compressive Sensing Detection Model

To clarify the connection of Compressed Sensing and the CDMA model developed in the previous section, recall the CS model (1). In terms of the CDMA model, the sparsity basis Φ is given by the set of PN sequences, which are used to separate the sensor nodes, and the channel influence can be interpreted as the measurement matrix Ψ . However, in contrast to the typical CS problem, the measurement matrix Ψ cannot be chosen arbitrarily, but is given physically, and the sparsity basis Φ is not constrained by physical properties of a signal, but can be chosen freely. Instead of viewing sparsity basis Φ and measurement matrix Ψ as separate parts of the system description, it is also widely accepted to combine the two matrices to an overall matrix \mathbf{A} that characterizes the transition from the sparse domain to the observation [16], which is exactly the CDMA system equation (6). This means that with (6) as the system model by solving (2) we jointly detect data and activity of sensor nodes without requiring activity control signaling.

Unfortunately, the computational complexity of this approach quickly gets prohibitive. As the size of the Multi-user vector \mathbf{x} is given by KN_c a restricted model for detection purposes is required for reasonable code word sizes N_c . Therefore, (6) is divided into sub-problems, each considering L consecutive transmit symbols per node. Thus, each of the N_c/L sub-problems is determined by the system matrix $\mathbf{A}_\nu \in \mathbb{R}^{N_s L \times K L}$, which is the same for all $\nu = 1, \dots, N_c/L$, as both PN sequences and channels are assumed to be constant for an entire frame. In order to simplify the detection model, we neglect the ISI between sub-problems in the detection. Depending on the channel length L_h the choice of the sub-problem size L will determine the performance loss due to remaining ISI with negligible losses for $L/N_c \rightarrow 1$ and more severe losses for $L/N_c \rightarrow 0$. The transmit vector $\mathbf{x}_\nu \in \mathcal{A}_0^{KL}$ then summarizes L out of N_c transmit symbols for each of the K nodes in the ν^{th} sub-problem as

$$\mathbf{x}_\nu = [d_{1,L(\nu-1)+1}, \dots, d_{1,L\nu}, \dots, d_{K,L(\nu-1)+1}, \dots, d_{K,L\nu}]^T. \quad (8)$$

Note, that the properties of \mathbf{A} in a typical CDMA setup are different from the usual assumptions in CS literature. In CS theory, the columns of \mathbf{A} are usually assumed to be normalized to length 1, i.e., columns with identical ℓ_2 -norm, which may have a severe impact on the reconstruction quality. In the previously defined CDMA system, both the channel and the PN sequences will on average have unit norm. However, due to the random behavior of these two influences, the column norms may vary strongly, and thus statistical reconstruction quality will deviate from the ideal CS behavior. Therefore, we introduce an effective channel $\tilde{\mathbf{H}}_\nu$, which is a diagonal

matrix of the column norms of \mathbf{A}_ν such that

$$\mathbf{A}_\nu = \tilde{\mathbf{A}}_\nu \tilde{\mathbf{H}}_\nu. \quad (9)$$

Using this modification, CS detection is applied to a unit norm matrix $\tilde{\mathbf{A}}_\nu$ which has two advantages: i) the performance of the greedy CS algorithm discussed in Section 3 is enhanced [25] and ii) either soft-value calculation or hard quantization after CS detection can exploit $\tilde{\mathbf{H}}_\nu$ as the effective channel.

In summary, the following simple system equation fully characterizes the communication for each of the $\nu = 1, \dots, N_c/L$ equivalent CS detection models

$$\mathbf{y}_\nu = \tilde{\mathbf{A}}_\nu \tilde{\mathbf{H}}_\nu \mathbf{x}_\nu + \mathbf{n}_\nu, \quad (10)$$

where $\mathbf{n}_\nu \in \mathbb{R}^{N_s L}$ is real-valued AWGN noise $\mathcal{N}(0, \sigma_n^2)$ and $\mathbf{y}_\nu \in \mathbb{R}^{N_s L}$ denotes the received vector at the aggregation node.

3.2. Detection Strategies and Assumptions

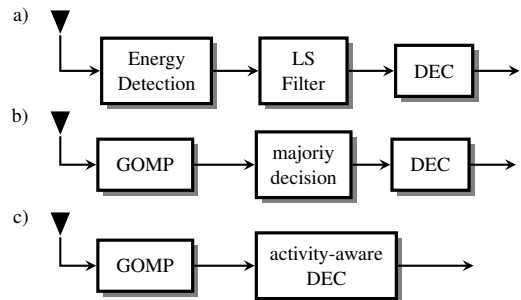


Figure 4. Detection strategies for data and activity detection

Besides the CS based detection, which is the main focus of this paper, other activity detection strategies can be applied. Fig. 4 summarizes the three strategies that will be compared here. Fig. 4a shows *separate* activity and data detection, where energy detection (ED) is employed to identify active nodes and data estimation is then achieved by LS filtering, hard/soft decision and subsequent decoding. Accordingly, ED-LS will be the baseline system to compare with and is briefly described below. Fig. 4b shows the CS-MUD approach using the Group Orthogonal Matching Pursuit (GOMP) detailed below followed by a majority decision to *jointly* detect activity and data. Due to the fact that the GOMP will be applied per sub-problem as described in the previous section node activity is only decided for groups of L symbols requiring an additional frame level decision. The majority decision is the optimal hard activity decision in this context, because frame activity can be interpreted as a repetition code of code rate $1/N_c$ regarding activity. Thus, frame activity knowledge is exploited even though CS-MUD is applied only per sub-problem. Again, hard/soft decision and standard decoding can be applied to recover the data. Finally, Fig. 4c depicts a combined detection strategy of CS-MUD and an extended

decoder that incorporates the decision which nodes are active into decoding based on the knowledge of the code structure. More details about this “activity-aware” decoder will be discussed in Section 4. Regarding complexity, ED-LS is obviously the simplest. The GOMP adds additional complexity due to the iterative approach, where the main complexity stems from repeated LS filtering and finally the most complex scheme employs the activity-aware Viterbi, which has to be used for nearly all nodes to decide activity instead of only the active ones. Subsequently, the energy detector and GOMP will be discussed in more detail.

The base assumption for all detection strategies in this paper is perfect CSI knowledge at the aggregation node, i.e., \mathbf{A} and σ_n^2 are known. In a practical setup, the channel knowledge has to be acquired, e.g., by semi-blind methods exploiting the structure imposed by spreading similar to [26]. Furthermore, for slowly changing channels estimation could be facilitated by a regular estimation phase with channel tracking in active phases of the sensor nodes. A detailed scheme for CSI acquisition, though, is beyond the scope of this paper.

3.3. Energy Detection

The baseline activity detector for CDMA systems employs a matched filter to correlate the columns of \mathbf{A} with the received signal \mathbf{y} and sums this correlation for all N_C symbols of a frame to determine active users. The correlation of the k^{th} user is then given by

$$r_k = \sum_{j \in \Gamma(k)} |\mathbf{A}_j^H \mathbf{y}|, \quad (11)$$

where $\Gamma(k)$ denotes the column indices of user k . Principally, LS and MMSE filtering could also be used here, but are omitted due to complexity reasons.

Usually, active users are identified based on a correlation threshold given by the assumed noise model [27]. However, to provide a fair comparison to the GOMP and to avoid threshold dependencies we assume that the energy detector has perfect knowledge of the instantaneous number of active users K_a . Thus, the K_a nodes with the highest correlations r_k are collected in the set of active nodes G . Subsequently, Least Squares filtering with $\mathbf{A}_{\Gamma(G)}^\dagger$ followed by a hard or soft decision can be applied to estimate the transmitted data of the active users. Note that the computational complexity of the pseudoinverse quickly gets prohibitive for large frame sizes N_C . Therefore, energy detection exploits the full frame structure, but LS filtering will be applied to the sub-problems (9) ignoring some amount of ISI.

3.4. Group Orthogonal Matching Pursuit

The Group Orthogonal Matching Pursuit algorithm [28] is a block-sparsity exploiting member of a family of greedy CS algorithms. The main property of these algorithms is low complexity. However, the drawback of any greedy approach is error propagation, since previous choices

Algorithm 1 Group Orthogonal Matching Pursuit (GOMP)

```

 $G^0 = \emptyset, \ell = 0, \mathbf{r}^0 = \mathbf{y}$ 
repeat
   $\ell = \ell + 1$ 
   $k_{\max} = \arg \max_k \sum_{j \in \Gamma(k)} |\mathbf{A}_j^H \mathbf{r}^{\ell-1}|$  with  $k \in \overline{G^{\ell-1}}$ 
   $G^\ell = G^{\ell-1} \cup k_{\max}$ 
   $\hat{\mathbf{x}}_{\Gamma(G^\ell)}^\ell = \mathbf{A}_{\Gamma(G^\ell)}^\dagger \mathbf{y}$  and  $\hat{\mathbf{x}}_{\Gamma(\overline{G}^\ell)}^\ell = \mathbf{0}$ 
   $\mathbf{r}^\ell = \mathbf{y} - \mathbf{A} \hat{\mathbf{x}}^\ell$ 
until  $\ell = K_a$ 

```

for symbol activity are not re-evaluated. Even though the GOMP will be applied to sub-problems (10), the description in this section omits the index ν for the sake of notational clarity.

The OMP [29, 30], from which the GOMP is derived, is based on two observations: the sparse solution in general contains only a few non-zero values and the measurements in the absence of noise are in a space supported by a sub-matrix of \mathbf{A} determined by the position of these non-zero elements. Therefore, the OMP iteratively builds the support of $\tilde{\mathbf{x}}$, by building the set of columns Γ of matrix \mathbf{A} that define the space for the measurements \mathbf{y} . The GOMP extends this idea by using knowledge about block-sparsity. Due to the fact that L symbols are stemming from a single node, this group of symbols has to be either active or non-active. Therefore, the support has to be build by adding indices of groups to a group index set G instead.

During each iteration, the GOMP adds the group k_{\max} , which has the highest correlation to the previous residual $\mathbf{r}^{\ell-1} = \mathbf{y} - \mathbf{A} \tilde{\mathbf{x}}^{\ell-1}$ to the set $G^{\ell-1}$. This step implements the activity detection based on previous activity and data decisions. After this activity detection, the values of the non-zero elements in $\tilde{\mathbf{x}}$ are determined by Least Squares estimation using the sub-matrix $\mathbf{A}_{\Gamma(G^\ell)}$ of all columns in $\Gamma(G^\ell)$. Afterwards, the current residual \mathbf{r}^ℓ is calculated. As the subsequent iteration depends on the previous residual $\mathbf{r}^{\ell-1}$, the activity detection of the GOMP is influenced by errors in the data estimation. These iterations are continued up to an appropriate stopping criterion, which is discussed in detail in [31]. Without loss of generality, we use the ideal stopping criterion of terminating the algorithm after a number of iterations equal to a known sparsity K_a , i.e., terminating at $\ell = K_a$.

In contrast to the baseline system using energy detection and LS filtering, here the correlation with respect to the combination of spreading sequence and channel is done iteratively and on the level of the sub-problem instead of the whole frame. Furthermore, LS estimation and activity detection are done jointly such that the LS estimation enhances the activity detection step by reducing the residuum. This suppresses highly correlated nodes that are likely chosen by ED, if both exhibit an overall high correlation r_k .

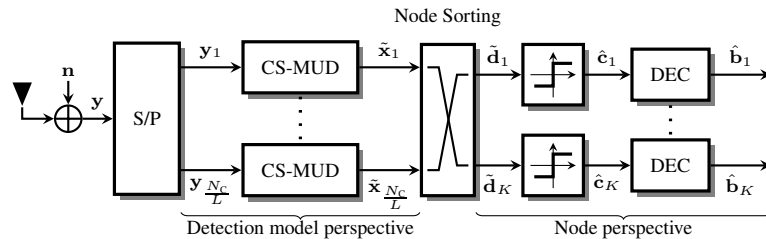


Figure 5. Detection structure at the aggregation node for CS based joint activity and data detection

4. CS DETECTION AND DECODING

4.1. General Aspects

According to the system model and the GOMP algorithm discussed in the previous section, decisions are made on a sub-problem level, neglecting the frame activity of sensor nodes. The received vector \mathbf{y} is separated for multiple CS-MUD detectors per (10), which is illustrated in Fig. 5 and the resulting symbol estimates \tilde{x}_ν have to be sorted to reconstruct the node specific frames. Thus, the estimated frame $\tilde{\mathbf{d}}_k$ of sensor node k is build as the stacked vector of L estimates from $\nu = 1, \dots, N_c/L$ detector runs (cf. Algorithm 1)

$$\tilde{\mathbf{d}}_k = [\tilde{x}_{1,L(k-1)+1}, \dots, \tilde{x}_{1,Lk}, \dots, \tilde{x}_{N_c/L,L(k-1)+1}, \dots, \tilde{x}_{N_c/L,Lk}]. \quad (12)$$

Thereby, each frame contains continuous non-zero entries, where data has been estimated, and zero entries, where the sensor has been estimated as inactive. In other words the effective channel after CS-MUD behaves similar to an erasure channel. As a result, it is not obvious which sensor nodes were active on a frame level and how decoding can be accomplished.

After frame reconstruction, either a quantizer or LLR calculation is required to demodulate the non-zero output of the CS-MUD using the effective channels $\tilde{\mathbf{H}}_\nu$. Originally, the transmit frame of active nodes only contained symbols from the alphabet \mathcal{A}^{N_c} , which poses the question how to handle zero symbols in the estimated frame. The optimal processing should exploit soft-information to preserve reliability information for decoding. Due to the hard decided activity by the GOMP, however, soft-information is not available for zeros and only in the sense of the effective channel coefficients after LS filtering for non-zero entries. Thus, soft-values for non-zeros can be estimated easily similar to soft-processing for linear filters, often employed in communication systems.

From the channel coding point of view each zero is equivalent to an erased bit and in terms of LLR values conveys no decision in either direction. Such, any channel coding scheme that allows to handle erased bits in the decoder should be suitable. In the following, however, we will focus on convolutional codes as a specific channel coding example, as erased code bits pose no difficulty for

a Viterbi decoder. In comparison to optimal symbol-by-symbol decoding through a BCJR decoder, the (Soft-in) Viterbi decoder offers the advantage to determine optimal codewords, which can be exploited directly for activity detection as will be shown in the next section. Still, depending on the percentage of erased bits, decoding errors might be very likely, due to a lack of information. As a rule of thumb, for Viterbi decoding, a decision after $5L_C$ state transitions does not influence the remainder of the code word, where L_C is the constraint length. So, if a whole block of bits in the range of $5L_C$ has been erased, this part of the code word is entirely random.

Independent of the number of symbols in a frame that have been estimated as active, the Viterbi decoder estimates the most likely code word, and thus generates an estimate of the transmitted information bits $\hat{\mathbf{b}}_k$. Consider an extreme example: an inactive node k is detected such that its estimated frame $\tilde{\mathbf{d}}_k$ only contains a single non-zero entry $\tilde{d}_{k,n}$. It seems obvious that this sensor has been inactive, however, the channel decoder cannot estimate inactivity as is. Hence, the two options already discussed as Fig. 4b, majority decision plus CS-MUD (M-CS-MUD) and activity-aware Viterbi decoding plus CS-MUD (V-CS-MUD) seem viable. In the following, we focus on the second approach to exploit the code redundancy for activity detection.

4.2. Activity-aware Viterbi Decoding

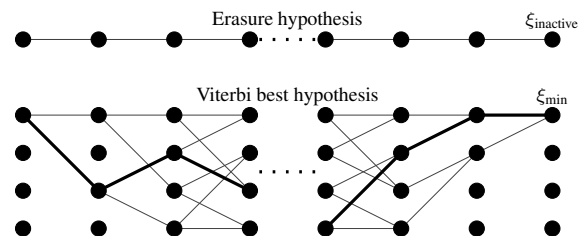


Figure 6. Activity-aware Viterbi extension by erasure path

To incorporate the knowledge that a node could have been inactive instead of sending a code word into the channel decoder, the Viterbi decoder needs to be augmented. Viterbi decoding using the Euclidean distance metric determines the state transition costs using a BPSK

notation of the code bits. This means that an erased symbol, i.e., a zero, simply acts as a “don’t care” in the metric, which is independent of hard or soft decision inputs. If the majority of the bits of a code word are erasures, the accumulated path metric will no longer provide any meaningful decision as most parts of the estimated minimum metric path will be arbitrarily chosen. Therefore, it is necessary to calculate the probability of an all-erasure word in addition to Viterbi decoding, which is illustrated in Fig. 6. The all-erasure path at the top of the figure indicates the necessary augmentation to consider inactivity of a sensor node and has a Euclidean distance metric $\xi_{\text{inactive},k}$, which simply is the l_2 -norm of the quantized code word vector $\hat{\mathbf{c}}_k$ in case of hard decisions

$$\xi_{\text{inactive},k} = \|\hat{\mathbf{c}}_k\|_2. \quad (13)$$

Alternatively, the metric can be calculated in the same sense for soft-values. Comparing the best code word hypothesis determined by Viterbi decoding with this erasure hypothesis is a generalized log-likelihood ratio test comparing the two hypotheses “active” and “inactive” [32]. Using the path metrics the problem can be cast as follows

$$\xi_{\text{inactive},k} - \xi_{\text{min},k} \underset{\text{Inactive}}{\overset{\text{Active}}{\geq}} 0. \quad (14)$$

Thus, if $\xi_{\text{inactive},k}$ is smaller than or equal to the minimum sum path metric $\xi_{\text{min},k}$ after Viterbi decoding, the k^{th} sensor node is estimated as inactive by the decoder. In summary, the channel decoder is used as a decision device for sensor node activity on a frame level.

5. NUMERICAL EVALUATION

5.1. Parameters

In the following, we discuss numerical evaluations of the detectors described in Section 3. More specifically, the focus is on a detailed analysis of the activity detection and the influence of the channel code.

An interesting setup for CS detection is given by *overloaded* CDMA systems, which result in under-determined equation systems for CS detection. Especially in such under-determined systems, CS algorithms offer potential gains over conventional MUD schemes if sensor node activity is unknown [7]. We focus on two scenarios to demonstrate the influence of different problem sizes on the performance. The channel is modeled by $L_h = 6$ i.i.d. Rayleigh distributed taps with an exponential decaying power delay profile for both cases. Due to the constant channel per frame, other channel statistics would mainly influence the amount of ISI and to a smaller degree the correlation, but do not change the overall behavior.

The first scenario uses $K = 128$ sensor nodes, where each node transmits using a PN sequence of length $N_s = 32$. Hence, the CDMA system is under-determined by a factor of four. Furthermore, the frame length is $N_c =$

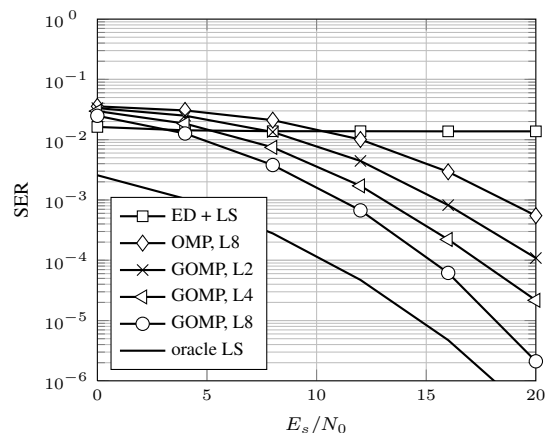


Figure 7. Symbol Error Rate for the augmented alphabet after activity and data detection, before decoding; Parameters: $K = 128$ nodes and activity probability $p_a = 2\%$

104 symbols using a standard terminated non-systematic non-recursive $[5; 7]_8$ convolutional code (code rate $R_C = 1/2$, constraint length $L_C = 3$). Consequently, $N_b = 50$ information bits are transmitted by every active node.

The second scenario with $K = 32$ sensor nodes and a PN sequence length of $N_s = 8$ is used to demonstrate the effect of channel coding and the problem size on performance. Furthermore, two different codes are utilized to demonstrate the achievable gains. On the one hand, the terminated $[5; 7]_8$ convolutional code is used as a “weak” code and the $[133; 145; 175]_8$ convolutional code (code rate $R_C = 1/3$, constraint length $L_C = 7$) as a “strong” code. The frame lengths are fixed to $N_c = 104$ in both cases. Note, that all presented results are restricted to hard decisions after symbol detection. An extension to soft-information is straightforward as discussed in Section 4 and effects all presented results in the same way.

5.2. Results without Channel Coding

Fig. 7 shows the symbol error rate (SER) over the augmented alphabet \mathcal{A}_0 before channel decoding. Thus, the SER contains both errors due to incorrect activity detection, and errors due to incorrect data estimation equally weighted. If not noted otherwise, the sub-problem size $L = 8$ and activity probability $p_a = 0.02$ are used. In addition to the approaches presented in Fig. 4a and c, i.e., ED-LS and GOMP with activity-aware Viterbi decoding, we show simulation results for the OMP. This is a GOMP with index group size 1 in terms of Algorithm 1 and demonstrates the gains by exploiting block-sparsity. Furthermore, the oracle LS is a LS estimation for known activity. Effectively, the SER of the oracle LS is only determined by data errors of the LS estimation and contains no support errors. First, we note that the GOMP of any sub-problem size L has much lower SER than the OMP in this scenario, which is obviously due to

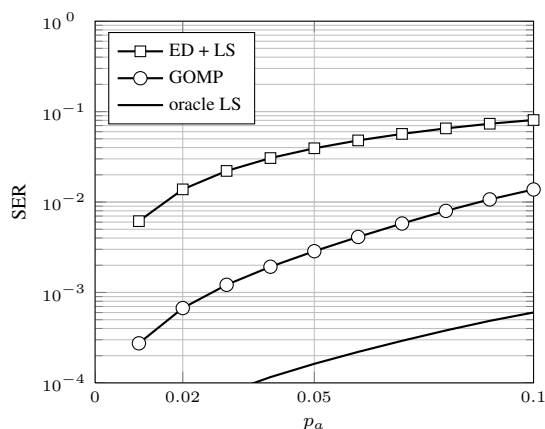


Figure 8. Symbol Error Rate vs. activity probability before decoding; Parameters: $K = 128$ nodes, $E_s/N_0 = 12$ dB and sub-problem size $L = 8$

the exploitation of block-sparsity. The larger the sub-problem size, the larger the gains. Applying the GOMP to the whole problem, i.e., $L = N_c$, can be expected to outperform energy detection over the whole SNR range, but is computationally prohibitive. Secondly, the ED-LS scheme shows superior performance compared to the CS schemes at low SNR, but quickly hits an error floor. The energy detector exploits the full frame activity knowledge, which leads to a good performance for high noise levels due to the averaging. However, systematic activity estimation errors caused by highly correlated columns of \mathbf{A} quickly dominate the performance and lead to an error floor. In overloaded CDMA systems, spreading sequences are highly correlated with a high probability and the OMP/GOMP approach explicitly suppresses nodes, whose columns are highly correlated with a node already detected as active. Lastly, all schemes have a much higher SER than the oracle LS. Therefore, we can conclude that their SER is mainly determined by errors in the activity detection.

To demonstrate the effect of different activity probabilities p_a , Fig. 8 shows the SER at $E_s/N_0 = 12$ dB. As can be seen, the performance of both approaches deteriorates with increased p_a and the difference between GOMP and ED gets smaller. The more nodes are active, the less errors are mainly caused by highly correlated nodes. In terms of CS theory, a higher activity probability corresponds to an overall higher mean sparsity and increased probability of high sparsity values. Sparsity and the degree of underdetermination are tightly connected and determine, if a system is reliably solvable by CS methods. Furthermore, the size of the system, i.e., the dimensions of \mathbf{A} , influences the performance, which is illustrated by Fig. 9.

Here, the second scenario with $K = 32$ nodes exhibits the same degree of sparsity and underdetermination as the first scenario when normalized to the problem dimensions. The overall performance of the CS schemes is visibly degraded in comparison to Fig. 7, whereas the energy

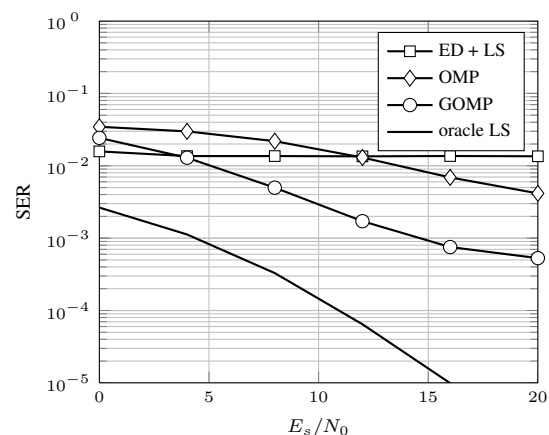


Figure 9. Symbol Error Rate before decoding; Parameters: $K = 32$ nodes, activity probability $p_a = 2\%$ and sub problem size $L = 8$

detector shows the same behavior. This indicates that CS detection performs better for larger system sizes, even though sparsity and underdetermination are comparable. In CS theory, phase transitions [33] are used for the noise free case to describe the parameter sets of underdetermination and sparsity for which perfect recovery is still reliable. Furthermore, both OMP and GOMP lead to an error floor at high SNR, which is due to systematic errors caused by correlations. As data detection errors can be improved by non-linear data estimation and by the subsequent channel decoding, we focus the investigation on activity errors.

5.3. Results with Channel Coding

In the following, we investigate numerical results after decoding. Specifically, activity errors of an entire code word/frame are considered, where we differentiate between frame missed detection (FMDR), i.e., an active node is detected as inactive, and frame false alarm rates (FFAR), i.e., the opposite case.

Fig. 10 shows the FFAR (dashed) and FMDR (solid) for the first scenario with $K = 128$ nodes and $p_a = 2\%$. For simplicity, only the ED-LS and GOMP results for $L = 8$ are shown. It is quite apparent, that the activity detection performance of the energy detector is the determining factor for the overall SER performance. Especially, the missed detection rate dominates in comparison to the false alarm rate. Similarly, the GOMP shows a missed detection rate behavior analogous to the SER results presented in Fig. 7. In contrast, no false alarms up to a FFAR of 10^{-5} could be observed in our evaluation, which allows the conclusion that the FFAR is much lower than the FMDR in this case. These results are especially interesting in terms of the overall CDMA system. On the one hand, higher layer processing does not need to determine “garbage” data estimated for truly inactive users, if false alarms are mostly negligible. On the other hand, most higher layer processing includes some amount of error detection like CRC codes, which implicitly allow to validate the activity

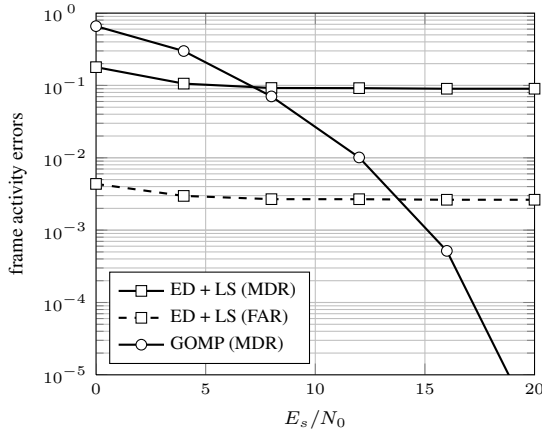


Figure 10. Frame missed detection (solid) and false alarm rates (dashed) after decoding; Parameters: $K = 128$ nodes, activity probability $p_a = 2\%$ and sub-problem size $L = 8$

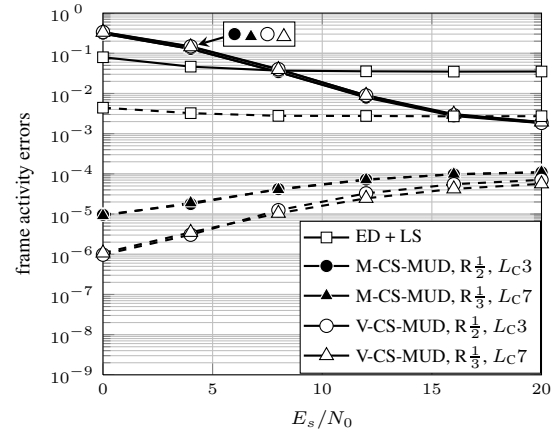


Figure 11. Frame missed detection (solid) and false alarm rates (dashed) after decoding; Parameters: $K = 32$ nodes, activity probability $p_a = 2\%$ and sub-problem size $L = 8$

decision of the physical layer schemes. From this point of view, a heightened false alarm rate would not harm system performance, but an improved missed detection rate could tremendously improve the detection of truly active nodes. Thus, new or modified detection algorithms are required that offer more flexibility to trade-off false alarms and missed detections.

Up to this point, the specific improvements achieved by channel coding have not been discussed. The presented results were focused on Fig. 4a (ED-LS) and Fig. 4c (V-CS-MUD). To illustrate the differences between Fig. 4b (M-CS-MUD) and Fig. 4c (V-CS-MUD), i.e., majority decision before Viterbi decoding and activity-aware Viterbi decoding, Fig. 11 shows the FFAR and FMDR for the second scenario with $K = 32$ nodes, a sub-problem size of $L = 8$ and activity probability $p_a = 0.02$. Results for both approaches are indicated by M-CS-MUD and V-CS-MUD in the figure. Furthermore, two different channel codes have been applied to illustrate the outcome of different coding strengths.

Once again, energy detection exhibits the same performance as in the first scenario. The missed detection rate (solid lines) for the CS-MUD approaches, however, surprisingly indicates no relevant performance differences comparing majority decision and activity-aware decoding, even comparing the two codes. A simple explanation for this behavior lies in the structure of the activity-aware Viterbi detection. The additional erasure path only helps to avoid false alarms, but does not improve missed detections in any way. Consequently, the code strength does not influence this behavior. Furthermore, majority decision is already optimal for hard decisions if the code structure cannot be exploited as in the case of false alarms.

Considering the false alarm rate (dashed lines), activity-aware Viterbi decoding actually shows slight performance gains over majority decision, whose performance obviously does not depend on the coding scheme. The performance differences of activity-aware decoding comparing

the “weak” and the “strong” code, though, are negligible as the main performance factor is the coding gain. The assumed block fading channel offers no diversity.

Interestingly, the false alarm rate increases with the E_s/N_0 instead of decreasing and the performance gain compared to the majority decision gets smaller. There are two opposing influences that cause the observed behavior: in the under-determined scenario it is possible that the columns \mathbf{A}_{k_1} and \mathbf{A}_{k_2} of two nodes k_1 and k_2 , which represent their spreading sequence and channel influences, are highly correlated. Thus, whenever one of these nodes, say k_1 , is active we are likely to estimate the incorrect, highly correlated node k_2 as active instead. As the correlation between nodes is the same for the entire frame, this type of error likely results in a frame error. More importantly, stronger channel coding only slightly helps to cope with it, as the received frame is a valid codeword with high probability due to the systematic nature of this error.

The second influence on the frame false alarm rate is the presence of noise. While the influence of noise leads to more errors on a symbol level, it eases frame error detection. The main reason for this is, that correlation between nodes is constant for an entire frame, while the noise is different for each transmitted chip. Thus, for each transmitted sub-frame, a different node may be incorrectly detected as being active. This increases the likelihood that no node is incorrectly decided as being active for the majority of symbols. We can conclude that the noise helps to disperse the symbol-wise activity errors over nodes, thereby making frame-wise activity errors less likely.

6. CONCLUSION

In this paper, we have presented a scheme for application of CS algorithms to a channel coded CDMA system employed for sensor node communication. To this end,

a CDMA chip-rate model has been formulated and decomposed into smaller sub-problems to lower the computational complexity at the cost of sub optimality. The major advantage of CS-MUD is that activity and data detection can be jointly facilitated on the physical layer, which may be exploited to lower the required signaling overhead for M2M systems. Here, we concentrated on two aspects: i) the detailed behavior of the GOMP detector with respect to activity detection and associated errors to provide meaningful performance indicators and ii) the application of channel coding in combination with such CS-MUD schemes to enhance the overall performance of the system, especially the activity detection.

Regarding the first point we have shown that the two possible activity errors (false alarms, missed detections) have to be differentiated to fully understand the impact of CS detectors on communication systems. Especially, missed detections determine the overall performance, whereas false alarms are less likely and can be further reduced using channel coding and adapted decoders. To achieve the second point, the erasure-like overall behavior of transmitter, channel and CS-MUD has been incorporated in the decoding process by a slightly enhanced Viterbi decoding scheme. The numerical results presented in this paper clearly indicate that the application of channel coding in this context enables an enhanced activity detection. However, this improvement is limited to the case of false alarms, which, in principle, is the less critical error. On the one hand, far less false alarms are occurring, and on the other hand, further processing on higher layers might ensure detection of misinformation. Still, activity errors are tightly coupled and enhancing the false alarm error rate may still prove useful, if the decoding results are fed back to the CS-MUD in an iterative fashion.

Based on our results, a variety of future research topics should be addressed. On the one hand, theoretical analysis regarding the properties of the system matrix \mathbf{A} in a CS context is urgently required to provide insights on the fundamental bounds of different algorithms. On the other hand, the activity detection clearly needs to be enhanced with respect to the missed detections. To this end, possible approaches are iterative schemes, which exploit the channel coding for enhanced false alarm rates or enhanced CS-MUD that primarily focuses on missed detection rates. The main conclusion we draw from Fig. 10 and 11 is that known CS algorithms are not well suited for systems, where the severity of activity estimation errors is not symmetrical. As already mentioned, missed detections are the dominating effect here and may also have an overall higher impact in a systems context, which leads to asymmetrical error importance. The usual goal of CS, however, is to find solutions that are as sparse as possible independent of this aspect, which is only ideal in terms of reconstruction error or the SER over the augmented alphabet. Therefore, for future applications, modified criteria have to be adopted to formulate optimal and suboptimal detection algorithms.

ACKNOWLEDGEMENTS

This work was supported in part by the German Research Foundation (DFG) under grant DE 1858/1-1.

REFERENCES

1. Bartoli A, Dohler M, Hernandez-Serrano J, Kountouris A, Barthel D. Low-Power Low-Rate goes Long-Range: the case for secure and cooperative Machine-to-Machine communications. *NETWORKING 2011 Workshops*, vol. 6827. Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; 219–230.
2. Lien S, Chen K, Lin Y. Toward ubiquitous massive accesses in 3GPP machine-to-machine communications. *Communications Magazine, IEEE* Apr 2011; **49**(4):66–74.
3. EXALTED consortium. D2.1 - description of baseline reference systems, scenarios, technical requirements & evaluation methodology. *Technical Report*, European Collaborative Research FP7 - Project Number: 258512 May 2011.
4. Zhai C, Liu J, Zheng L, Xu H, Chen H. Maximise lifetime of wireless sensor networks via a distributed cooperative routing algorithm. *Transactions on Emerging Telecommunications Technologies* 2012; .
5. Donoho D. Compressed sensing. *IEEE Transactions on Information Theory* 2006; **52**(4):1289–1306.
6. Candes EJ, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* Feb 2006; **52**(2):489–509.
7. Schepker H, Dekorsy A. Sparse multi-user detection for cdma transmission using greedy algorithms. *8th International Symposium on Wireless Communication Systems (ISWCS 11)*, Aachen, Germany, 2011.
8. Xu W, Lin J, Niu K, He Z. A joint recovery algorithm for distributed compressed sensing. *Transactions on Emerging Telecommunications Technologies* 2012; .
9. Baraniuk R, Steeghs P. Compressive radar imaging. *Radar Conference, 2007 IEEE*, 2007; 128–133.
10. Bajwa W, Haupt J, Sayeed A, Nowak R. Compressed channel sensing: A new approach to estimating sparse multipath channels. *Proceedings of the IEEE* June 2010; **98**(6):1058–1076.
11. Berger CR, Wang Z, Huang J, Zhou S. Application of compressive sensing to sparse channel estimation. *IEEE Communications Magazine* November 2010; **48**(11):164–174.
12. Schenk A, Fischer RF. Compressed-Sensing (Decision-Feedback) Differential Detection in Impulse-Radio Ultra-Wideband Systems. *Proceedings of IEEE International Conference on Ultra-Wideband*, Bologna, Italy, 2011; 121–125.
13. Zhu H, Giannakis GB. Exploiting sparse user activity in multiuser detection. *IEEE Transactions on*

- Communications* February 2011; **59**(2):454–465.
14. Angelosante D, Grossi E, Giannakis GB, Lops M. Sparsity-aware estimation of CDMA system parameters. *EURASIP Journal on Advances in Signal Processing* April 2010; **2010**:1–10.
 15. Schepker H, Dekorsy A. Compressive sensing multi-user detection with block-wise orthogonal least squares. *2012 IEEE 75th Vehicular Technology Conference: VTC2012-Spring*, Yokohama, Japan, 2012.
 16. Candès EJ, Wakin MB. An introduction to compressive sampling. *IEEE Signal Processing Magazine* March 2008; **25**(2):21–30.
 17. Strutz T. *Bilddatenkompression: Grundlagen, Codierung, MPEG, JPEG*. second edn., Praxiswissen, Vieweg, 2002.
 18. Candès EJ, Romberg JK. Sparsity and incoherence in compressive sampling. *Inverse Problems* 2007; **23**(3):969–985.
 19. Chen SS, Donoho DL, Sanders MA. Atomic decomposition by basis pursuit. *SIAM Journal of Scientific Computing* 1998; **20**(1):33–61.
 20. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B* 1996; **58**(1):267–288.
 21. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B* 2005; **67**(2):301–320.
 22. Ben-Haim Z, Eldar YC. Near-oracle performance of greedy block-sparse estimation techniques from noisy measurements. *Selected Topics in Signal Processing, IEEE Journal of Sep* 2011; **5**(5):1032–1047.
 23. Verdú S. *Multiuser Detection*. Cambridge Univ. Press: Cambridge, U.K., 1998.
 24. Applebaum L, Bajwa WU, Duarte MF, Calderbank R. Asynchronous code-division random access using convex optimization. *Physical Communication* 2012; **5**(2):129 – 147.
 25. Rangan S, Fletcher A, Goyal V. Asymptotic analysis of MAP estimation via the replica method and applications to compressed sensing. *IEEE Transactions on Information Theory* Mar 2012; **58**(3):1902–1923.
 26. Asif M, Mantzel W, Romberg J. Channel protection: Random coding meets sparse channels. *IEEE Information Theory Workshop, 2009. ITW 2009*, 2009; 348 –352.
 27. Urkowitz H. Energy detection of unknown deterministic signals. *Proceedings of the IEEE* april 1967; **55**(4):523 – 531.
 28. Majumdar A, Ward RK. Fast group sparse classification. *Electrical and Computer Engineering, Canadian Journal of* 2009; **34**(4).
 29. Pati Y, Rezaifar R, Krishnaprasad P. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. *Signals Systems and Computers* November 1993; **1**:40–44.
 30. Tropp JA, Gilbert AC. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory* December 2007; **53**(12):4655–4666.
 31. Tropp JA. Average-case analysis of greedy pursuit. *Proc. SPIE Wavelets XI*, 2005.
 32. Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning*. 2nd edn., Springer, 2011.
 33. Donoho D, Tanner J. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2009; **367**(1906):4273–4293.

AUTHORS' BIOGRAPHIES

Prof. **Armin Dekorsy** holds the chair of the Department of Communications Engineering, University of Bremen, since April 2010. He received his Dipl.-Ing. (FH) (B.Sc.) degree from Fachhochschule Konstanz, Germany; Dipl.-Ing. (M.Sc.) degree from University of Paderborn, Germany; PhD degree from the University of Bremen, Germany, all in communications engineering. From 2000 to 2007 he worked as research engineer at Deutsche Telekom AG and as Distinguished Member of Technical Staff (DMTS) at Bell Labs Europe, Lucent Technologies. In 2007 he joined Qualcomm GmbH as European Research Coordinator conducting Qualcomms' internal and external European research projects like ARTIST4G, BeFemto, and WINNER+. His current research interests include resource management, transceiver design and digital signal processing for wireless communications systems in health care, automation and mobile communications. Prof. Dekorsy is member of ITG expert committee "Information and System Theory", VDE and IEEE communications and signal processing society.

Dr.-Ing. **Carsten Bockelmann** received his Dipl.-Ing. (M.Sc.) degree in electrical engineering in 2006 and his PhD degree 2012 both in electrical engineering and from the University of Bremen, Germany. Since 2012 he is working as a post doctoral researcher at the University of Bremen coordinating research activities regarding the application of compressive sensing/sampling to communication problems. His current research interests include compressive sensing and its application in communications contexts, as well as channel coding and transceiver design.

Dipl.-Ing. **Henning F. Schepker** graduated from the RWTH Aachen University in Aachen, Germany, in 2009. He started working as a PhD candidate at the Department of Communications Engineering at the University of Bremen in Germany in 2010. His research is focused on Compressive Sensing in communication technology.