Iterative Soft Interference Cancellation for Sparse BPSK Signals

Carsten Bockelmann
Department of Communications Engineering,
University of Bremen, 28359 Bremen, Germany
email: bockelmann@ant.uni-bremen.de

***Abstract*—Compressed Sensing based Multi-User Detection (CS-MUD) is a novel MUD approach applied in sporadic Machine Type Communication (MTC) to identify actively transmitting sensors nodes at the same time as the transmitted data. In this context, different reconstruction algorithms from CS as well as detection concepts well established in communications have been adapted, but either are not exploiting finite alphabets or are highly complex. In this paper, we focus on an iterative soft interference cancellation scheme that efficiently exploits the sparsity of the signal, the finite alphabet of the transmit data and the usually employed channel coding to obtain a robust high performance detection scheme.**

I. Introduction

Machine Type Communication (MTC) is a heavily discussed topic in wireless communications due to the paradigm shift from human driven and mostly high data rate communication to machine driven and mostly low data rate communication. This paradigm shift leads to a long list of demanding requirements that need to be fulfilled to achieve the future vision of billions of connected Machine Type Devices (MTDs) in the context of smart X (cities, grid, etc.).

One building block towards a new air interface for MTC is the design of an appropriate physical layer that enables resource *and* signalling efficient communication. Focusing on the uplink, a massive amount of devices challenges the state of the art procedures for communication. In most cellular systems and many local area networks random access methods to establish initial connections or directly transmit data quickly become a bottle neck if too many devices access the channel at the same time. For narrow-band (non-frequency selective) channels uplink scenarios as depicted in Fig. 1 can be described by a multi-user detection problem [1]

$$\mathbf{Y} = \mathbf{A}\mathbf{D} + \boldsymbol{\eta}\,, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{M\times K}$ is the matrix of M receive signals over a block of K symbols. The system matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$ describes the input-output relationship from N sensors to M receive signals and $\mathbf{D} \in \mathcal{S}^{N\times K}$ describes the multi-user matrix of N sensor signals with K symbols from an alphabet $\mathcal{S}$ (e.g. QAM, PSK). Finally, $\boldsymbol{\eta} \in \mathbb{R}^{M\times K}$ denotes the additive noise with iid. $\eta_{ij} \sim \mathcal{N}\left(0, \sigma_\eta^2\right)\ \forall\, i, j$.

While Multi-User Detection (MUD) is a well known solution for problems like (1), the application of ideas from compressive sensing in this context is not. Due to random access only a subset of all users competing for channel access

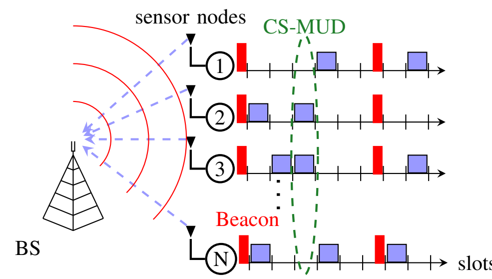

Figure 1. Schematic MTC uplink scenario using a ALOHA like scheme. A base station (BS) starts contention periods by beacon, multiple sensors access the channel concurrently on a slot basis. In each slot CS-MUD will be applied to detect the multi-user signal.

will be active for a given time/slot, which introduces sparsity to the detection problem. Compressed Sensing based Multi-User Detection (CS-MUD) is a novel scheme to efficiently handle this problem by jointly detecting activity and data of sensors in a multi-user detection problem thus alleviating collisions in random access. State of the art solutions either adapt CS algorithms to communication problems, e.g. in terms of finite alphabets [2], [3], or apply MAP detection [4] with sparse priors. However, the latter approach suffers from high complexity which makes it nearly impossible to solve the *massive* access problem.

In this paper, we look at the joint activity and data detection problem from a communications perspective and introduce sparsity in form of a statistical activity model to a well known efficient iterative detection concept in MIMO processing: soft interference cancellation. In the following, we shortly introduce the statistical sparsity model and some general assumptions and summarise the outline of soft interference cancellation. Then, we will highlight the differences due to the assumption of a statistical activity model and derive the soft value calculations to incorporate activity detection and channel decoding in one detection scheme. Finally, the discussion of exemplary results concludes this paper.

In the following bold lower case letters denote column vectors, e.g. $\mathbf{d}_k$ is the k^{th} column of the matrix $\mathbf{D}$, and bold underlined letters denote rows, e.g. $\underline{\mathbf{d}}_n$ is the n^{th} row of $\mathbf{D}$. The superscript t, e.g. $\mathbf{w}^t$, always denotes the t^{th} iteration.

II. Sparsity Model and Assumptions

To model the random access nature of the multi-user detection problem each sensor is assumed to be active with a

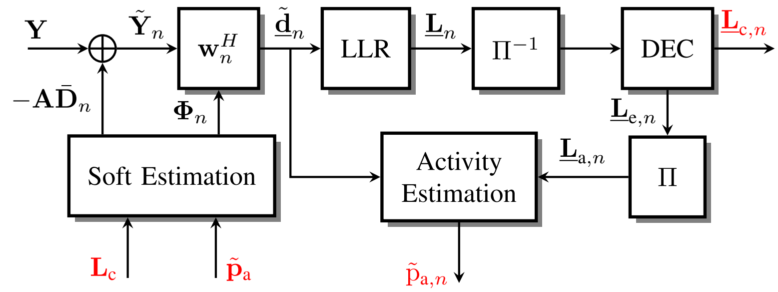

Figure 2. Structure of the iterative SIC for an exemplary sensor n omitting the iteration index.

known and identical probability of p_{a} for the whole frame length K. Using the general uplink model described in (1) the transmit data of sensor n is described by the n^{th} row $\underline{\mathbf{d}}_n$ of $\mathbf{D}$ which leads to a block-sparse matrix. Either blue $\underline{\mathbf{d}}_n \in \mathcal{S}^K$ contains signals from a finite modulation alphabet $\mathcal{S}$ (active) or $\underline{\mathbf{d}}_n = \mathbf{0}_K$ (inactive). Thus, we define the augmented symbol alphabet $\mathcal{S}_0^K = \mathcal{S}^K \cup \{\mathbf{0}_K\}$ to denote all vectors of length K that are either from a finite modulation alphabet $\mathcal{S}^K$ or all-zero, i.e. $\mathbf{0}_K$.

The system matrix $\mathbf{A}$ has to be modelled according to the specific communication setup, e.g for CDMA or OFDMA and appropriate channel assumptions. However, to highlight the performance compared to standard CS approaches, we will chose the elements of the system matrix $\mathbf{A}$ as iid. $\mathbf{A}_{ij} \sim \mathcal{N}(0, 1/\sqrt{M})\ \forall\ i,j$. Choosing $\mathbf{A}$ as Gaussian here is motivated by standard assumptions in CS to fulfil the restricted isometry property (RIP) and thereby ensure proper reconstruction [5], [6]. Performance losses through non-RIP matrices, usually leading to error floors [2], are beyond the scope of this paper.

Furthermore, we restrict the derivations and results presented to BPSK, i.e. $\mathcal{S} = \{-1, 1\}$, and a common activity probability p_{a} for all sensors without loss of generality. A generalisation of the presented derivations to higher order modulations and individual activity probabilities per sensor is straightforward, but would complicate notations unnecessary.

III. Iterative Soft Interference Cancellation

A. General Outline

The iterative soft-SIC is a well known detector in MIMO systems and will only be partly discussed in detail here. For a thorough overview, refer to [7]. In general, the iterative soft-SIC approach follows a structure as depicted in Fig. 2. For each sensor n and in each iteration t soft-estimates of the sensor data $\bar{\underline{\mathbf{d}}}_n^t$ are estimated in an iterative manner. The following procedure is executed iteratively per sensor and iteration, with the novel steps marked:

1) **Interference cancellation** using estimated soft-symbols $\bar{\mathbf{D}}_n^t$.
2) **Interference suppression** by adaptive MMSE filtering with $\mathbf{w}_n^t$ according to the sensor specific average covariance $\mathbf{\Phi}_n^t$ of the interference reduced signal $\mathbf{D} - \bar{\mathbf{D}}_n^t$.
3) **Soft-value processing and decoding** to provide a-posterior LLRs $\mathbf{L}_{\mathrm{c}}$ and extrinsic LLRs $\underline{\mathbf{L}}_{\mathrm{e},n}$.
4) {**new**} **Estimation of the activity** probability $\tilde{\mathrm{p}}_{\mathrm{a},n}^t$ for each sensor n exploiting the extrinsic knowledge gained from decoding by $\underline{\mathbf{L}}_{\mathrm{e},n}$.
5) {**modified**} **Calculation of soft-symbols** $\bar{\underline{\mathbf{d}}}_n^t$ and remaining interference power λ_n^t for interference calculation using $\tilde{\mathrm{p}}_{\mathrm{a},n}^t$ and $\underline{\mathbf{L}}_{\mathrm{c},n}$.

B. Initialisation

For the first sensor ($n = 1$) in the first iteration ($t = 1$) the soft-symbol vectors for all sensors are initialised as $\bar{\underline{\mathbf{d}}}_n^1 = \mathbf{0} \quad \forall n = 1, \ldots, N$. Also, the mean interference powers, which are required for the calculation of filter $\mathbf{w}_n^t$, are initialised as $\lambda_j^1 = p_{\mathrm{a}} \quad \forall j = 1, \ldots, N$ due to the unknown activity of sensors which results in a mean power of p_{a}.

C. Interference Cancellation and Adaptive MMSE Filtering

To suppress interference from all sensors except the n^{th} in the interference cancellation step the n^{th} soft-estimate component is set to zero, i.e.

$$\bar{\mathbf{D}}_n^t = [\bar{\underline{\mathbf{d}}}_1^t; \ldots; \bar{\underline{\mathbf{d}}}_{n-1}^t; \mathbf{0}; \bar{\underline{\mathbf{d}}}_{n+1}^{t-1}; \ldots; \bar{\underline{\mathbf{d}}}_N^{t-1}] \tag{2}$$

The subsequent MMSE filter calculation requires the remaining interference power after cancellation given by $\lambda_n^t = 1/L \sum_{j=1}^{L} \mathrm{cov}\{d_{n,j} - \bar{d}_{n,j}^t\}$.

Then the MMSE filter for sensor n in iteration t after interference cancellation reads

$$\mathbf{w}_n^t = \left(\mathbf{A}\mathbf{\Phi}_n^t\mathbf{A}^H + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{a_n}\,, \tag{3}$$

where $\mathbf{a}_n$ is the n^{th} column of $\mathbf{A}$ and

$$\mathbf{\Phi}_n^t = \mathrm{diag} = [\lambda_1^t, \ldots, \lambda_{n-1}^t, 1, \lambda_{n+1}^{t-1}, \ldots, \lambda_N^{t-1}]\,, \tag{4}$$

which assumes the current sensor to be active.

Finally, the LLR values for sensor n are calculated based on an equivalent AWGN model after filtering and forwarded to the decoder. The decoder output is then used for two purposes: (1) estimating the activity probability and (2) calculating soft-symbol estimates of $\underline{\mathbf{d}}_n$.

D. Soft Estimation

We know that sensors are active for a whole frame, which naturally leads to a frame perspective for the estimation of soft-symbols, i.e. estimation of the vector $\bar{\underline{\mathbf{d}}}_n$, omitting the iteration index t for the sake of notational clarity. Therefore, we aim to estimate

$$\bar{\underline{\mathbf{d}}}_n = \mathrm{E}\left[\underline{\mathbf{d}}_n\right] = \sum_{\mathbf{s}\in\mathcal{S}_0^K} \mathbf{s}\cdot \mathrm{p}\left(\mathbf{s}|\underline{\mathbf{L}}_{\mathrm{c},n}\right)\,, \tag{5}$$

where $\mathrm{p}\big(\mathbf{s}|\underline{\mathbf{L}}_{\mathrm{c},n}\big)$ denotes the probability of $\mathbf{s}$ given the decoder output of sensor n. Note, that $\mathcal{S}_0^K$ denotes the augmented alphabet as defined in section II. Calculating (5) in the given form is quite complex, even for small values of K. However, using the usual assumption of independent symbols due to interleaving, the expression can be simplified to symbol level expressions. First, the all-zero sequence obviously vanishes, which leads to

$$\bar{\underline{\mathbf{d}}}_n = \sum_{\mathbf{s}\in\mathcal{S}^K_{s_i=1}} \mathbf{s}\cdot\prod_{j=1}^{K}\mathrm{p}(s_j|L_{\mathrm{c},n,j}) + \sum_{\mathbf{s}\in\mathcal{S}^K_{s_i=-1}} \mathbf{s}\cdot\prod_{j=1}^{K}\mathrm{p}(s_j|L_{\mathrm{c},n,j})\,, \tag{6}$$

where $\mathcal{S}^K_{s_i=b}$ simply denotes the set of all vectors with $s_i = b$. Therefore, focusing on a single element $\bar{d}_{n,i}$ leads to

$$\bar{d}_{n,i} = [\mathrm{p}(s_i=+1|L_{\mathrm{c},n,i}) - \mathrm{p}(s_i=-1|L_{\mathrm{c},n,i})] \times \sum_{\mathbf{s}_{-i}\in\mathcal{S}^{K-1}}\prod_{j=1,j\neq i}^{K}\mathrm{p}(s_j|L_{\mathrm{c},n,j})\ . \tag{7}$$

In a slight abuse of notation, the shorthand $\mathbf{s}_{-i}\in\mathcal{S}^{K-1}$ is used to denote the vector $\mathbf{s}$ without the i^{th} dimension. Finally, the remaining sum term has to be determined, which can be done using the sum of the probabilities and factoring it the same way as above, which leads to

$$\sum_{\mathbf{s}_{-i}\in\mathcal{S}^{K-1}}\prod_{j=1,j\neq i}^{K}\mathrm{p}(s_j|L_{\mathrm{c},n,j}) = \frac{1-\mathrm{p}\big(\mathbf{s}=\mathbf{0}_K|\underline{\mathbf{L}}_{\mathrm{c},n}\big)}{\mathrm{p}(s_i=+1|L_{\mathrm{c},n,i})+\mathrm{p}(s_i=-1|L_{\mathrm{c},n,i})}\ . \tag{8}$$

Introducing the estimated activity probability $\tilde{\mathrm{p}}_{\mathrm{a},n} = 1-\mathrm{p}\big(\mathbf{s}=\mathbf{0}_L|\underline{\mathbf{L}}_{\mathrm{c},n}\big)$ we finally achieve

$$\bar{d}_{n,i}\ =\ \tilde{\mathrm{p}}_{\mathrm{a},n}\cdot\frac{\mathrm{p}(s_i=+1|L_{\mathrm{c},n,i})-\mathrm{p}(s_i=-1|L_{\mathrm{c},n,i})}{\mathrm{p}(s_i=+1|L_{\mathrm{c},n,i})+\mathrm{p}(s_i=-1|L_{\mathrm{c},n,i})} \tag{9}$$

The calculation of the interference powers λ_n - see (4) - can be approached in a similar way, but is very simple here due to the restriction to BPSK

$$\lambda_n = \tilde{\mathrm{p}}_{\mathrm{a},n} - \frac{1}{K}\|\bar{\underline{\mathbf{d}}}_n\|_2^2\,. \tag{10}$$

To summarise, the probabilities $\mathrm{p}(s_i=\pm1|L_{\mathrm{c},n,i})$ in (9) can be easily obtained by standard LLRs, e.g. unmodified soft decoder output, without probability scaling issues due to the ratio in (9). However, a novelty here is the consideration of the estimated activity probability $\tilde{\mathrm{p}}_{\mathrm{a},n}$ which statistically introduces sparsity to the soft interference cancellation.

E. Activity Estimation

The posterior probability $\mathrm{p}\big(\mathbf{s}=\mathbf{0}|\underline{\mathbf{L}}_{\mathrm{c},n}\big)$ in (8) cannot be determined based on the decoder output LLRs $\underline{\mathbf{L}}_{\mathrm{c},n}$ alone as the decoder does not convey any information about the *inactivity* of node n. Instead we use the estimate after filtering $\tilde{\underline{\mathbf{d}}}_n$ with the decoder output as a-priori knowledge.

Starting with the frame activity LLR for node n

$$L_{\mathrm{act},n} = \log\frac{\mathrm{p}(\mathbf{s}=\mathbf{0}|\tilde{\underline{\mathbf{d}}}_n)}{\mathrm{p}(\mathbf{s}\in\mathcal{S}^K|\tilde{\underline{\mathbf{d}}}_n)} \tag{11}$$

we note that calculation of $L_{\mathrm{act},n}$ requires the enumeration of all vectors in $\mathcal{S}_0^K$ which is infeasible. However, the probability of inactivity for sensor n can be calculated *approximately* employing symbol level activity LLRs. If we wrongly assume symbol activity to be independent, the frame activity LLR $L_{\mathrm{act},n}$ can be approximated as the sum of all $L_{\mathrm{act},n,j}$. Then the activity LLRs of the j^{th} symbol of sensor n are defined as

$$L_{\mathrm{act},n,j} = \log\frac{\mathrm{p}(s=0|\tilde{d}_{n,j})}{\mathrm{p}(s=-1|\tilde{d}_{n,j})+\mathrm{p}(s=+1|\tilde{d}_{n,j})}, \tag{12}$$

where the denominator can be rewritten using Bayes rule to exploit decoder information through $\mathrm{p}(s=\pm1)$. Note, that calculation of these probabilities through decoder LLRs requires rescaling by $1-\mathrm{p}(s=0)$ to fulfill completeness of probabilities for the ternary alphabet $\mathcal{S}_0$. Thus, we get

$$L_{\mathrm{act},n,j} = \log\frac{\mathrm{p}(s=0)}{1-\mathrm{p}(s=0)} + \underbrace{\log\frac{\mathrm{p}(\tilde{d}_{n,j}|s=0)}{\sum\limits_{b\in\mathcal{S}}\mathrm{p}(\tilde{d}_{n,j}|s=b)\frac{1}{1+\exp(-bL_{\mathrm{c},n,j})}}}_{L_{\mathrm{e,act},n,j}}\ . \tag{13}$$

However, the first part in (13) is not well defined on a symbol level due to the assumed frame activity. Therefore, we choose to take the sum over the extrinsic part $L_{\mathrm{e,act},n,j}$ only and consider a priori information once for the whole frame. Then, the estimate $\tilde{\mathrm{p}}_{\mathrm{a},n}$ reads

$$\tilde{\mathrm{p}}_{\mathrm{a},n} = 1-\mathrm{p}\Big(\mathbf{s}=\mathbf{0}|\tilde{\underline{\mathbf{d}}}_n\Big)\ = 1-\frac{1}{1+\exp(-L_{\mathrm{act},n})} \tag{14}$$

with

$$L_{\mathrm{act},n} \approx \log\frac{1-\mathrm{p_a}}{\mathrm{p_a}} + \sum_{j=1}^{K}L_{\mathrm{e,act},n,j}\,. \tag{15}$$

This reflects that activity is defined on a frame level consistently and exploits the extrinsic knowledge gained from all observations of the frame. Note that the a priori information $\mathrm{p_a}$ will not be updated from iteration to iteration in this scheme but remains fixed.

IV. Numerical Results

Fig. 3 shows results for two parameter sets in terms of the frame error rate after decoding. The SIC approach is depicted in comparison to a Group Orthogonal Matching Pursuit (GOMP) [8] assuming perfect knowledge of *the number* of active sensors required for proper termination of the algorithm. After detection a BCJR decoder is applied. Furthermore, two oracle detectors with perfect activity knowledge (SIC and LS filter) are used as a benchmark for each scheme, respectively.

To highlight the gains achieved by exploitation of the code structure due to soft interference cancellation as well as the novel activity estimation two additional SIC variants are shown. The derived algorithm is simply termed “SIC”. The first variant “SIC w/o DEC” excludes the decoder from the SIC and uses separate decoding *after* SIC. For large number of sensors and low activity probabilities the number of estimated sensors after SIC will be considerably lower than the number

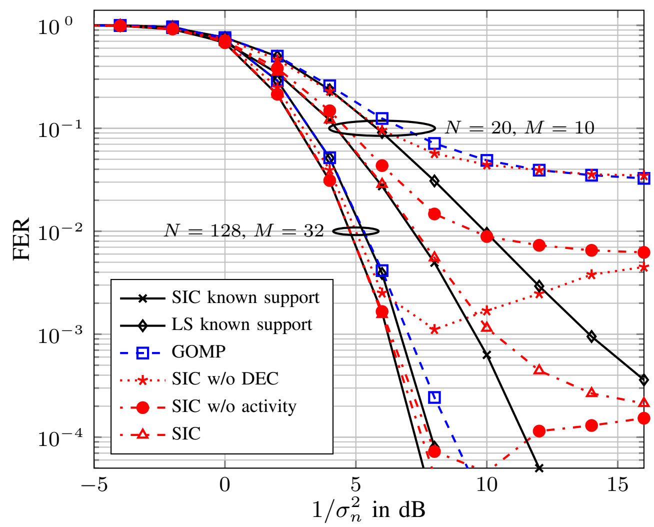

Figure 3. Frame error rate vs. $1/\sigma_n^2$ for two parameter sets: (i) $N = 128$, $M = 32$, $\mathrm{p_a} = 0.02$ and (ii) $N = 20$, $M = 10$, $\mathrm{p_a} = 0.1$, both employing a [7;5] convolutional code with $K = 102$ code bits per frame. SIC uses 2 iterations. 10^6 Monte-Carlo trials.

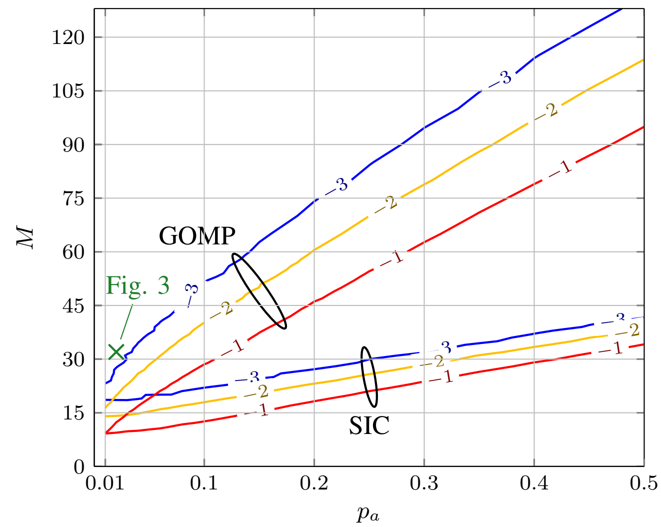

Figure 4. Contour plot of logarithmic FERs (exponent) vs. measurements M and activity probability $\mathrm{p_a}$ at 7dB for $N = 128$ employing a [7;5] convolutional code with $K = 102$ code bits per frame. SIC uses 2 iterations. 10^4 Monte-Carlo trials.

of all sensors thus reducing the decoding complexity. However, as a consequence activity estimation as well as soft-symbol calculation are solely based on the LLRs after MMSE filtering. The second variant "SIC w/o activity" is only using $\underline{\tilde{\mathbf{d}}}_n$ when calculating (15) and completely neglects the decoder output in activity estimation.

Comparing the two parameter sets $N = 128$ and $N = 20$ immediately illustrates the dependence of the FER performance on the size of the detection problem. In the larger system GOMP and SIC seem to perform comparable, whereas in the smaller system the SIC exhibits superior performance due to the decoding gains. Generally, CS approaches show better performance for large problems even at the same relative sparsity whereas SIC already performs well for smaller problems. The major performance changes of the SIC variants are an early error floor ("SIC w/o activity") and loss of decoding gain ("SIC w/o DEC"). Note, that depending on the SNR working point and required performance the variants may be viable complexity reduced alternatives.

More interestingly, the SIC scales vastly better in terms of required measurements M and increases in activity probability $\mathrm{p_a}$ as depicted by the FER contours for $N = 128$ in Fig. 4. Evidently, the comparable performance of GOMP and SIC in Fig. 3 for $N = 128$, indicated by the (green) marker, is a result of the choice $M = 32$ which already enables FER $< 10^{-3}$ for both GOMP and SIC. The main reason for the superiority of the SIC algorithm here is the exploitation of code structure due to soft interference cancellation as well as the consideration of the statistical activity/sparsity model. Channel coding introduces structure to the problem which is usually not available in standard CS contexts like image reconstruction. A straightforward comparison to known reconstruction bounds is therefore not possible.

V. Conclusion

In this paper we have derived a soft interference cancellation algorithm for statistically sparse finite alphabet signals exploiting channel coding to improve the activity and data estimation for sporadic communication setups. The presented results indicate that this approach is vastly superior compared to known CS algorithms like GOMP which are just applied but not tailored to communications assumptions through proper exploitation of the code structure.

VI. Acknowledgement

Part of this work has been performed in the framework of the FP7 project ICT- 317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project.

References

[1] F. Monsees, C. Bockelmann, and A. Dekorsy, "Compressed sensing soft activity processing for sparse multi-user systems," in *9th IEEE Broadband Wireless Access Workshop at IEEE Globecom 2013*, Atlanta, USA, Dec 2013.

[2] C. Bockelmann, H. F. Schepker, and A. Dekorsy, "Compressive sensing based multi-user detection for machine-to-machine communication," *Transactions on Emerging Telecommunications Technologies*, vol. 24, no. 4, pp. 389–400, 2013.

[3] Y. X., Y. Eldar, and A. Goldsmith, "Reduced-dimension multiuser detection," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3858–3874, June 2013.

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

[5] E. J. Candés and T. Tao, "Decoding by linear programming," *IEEE Transaction on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[6] E. J. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[7] M. Witzke, S. Baro, F. Schreckenbach, and J. Hagenauer, "Iterative detection of mimo signals with linear detectors," in *Asilomar Conference on Signals, Systems and Computers*, vol. 1, Nov 2002, pp. 289–293.

[8] A. Majumdar and R. K. Ward, "Fast group sparse classification," *Electrical and Computer Engineering, Canadian Journal of*, vol. 34, no. 4, 2009.