Advanced Topics in Digital Communications
Spezielle Methoden der digitalen Datenübertragung

Dr.-Ing. Carsten Bockelmann
Institute for Telecommunications and High-Frequency Techniques
Department of Communications Engineering
Room: SPT C3160, Phone: 0421/218-62386
bockelmann@ant.uni-bremen.de

Lecture
Thursday, 10:00 – 12:00 in N3130

Exercise
Wednesday, 14:00 – 16:00 in N1250

Dates for exercises will be announced during lectures.

Tutor
Tobias Monsees
Room: SPT C3220
Phone 218-62407
tmonsees@ant.uni-bremen.de

www.ant.uni-bremen.de/courses/atdc/

Outline

- Part 1: Linear Algebra
  - Eigenvalues and eigenvectors, pseudo inverse
  - Decompositions (QR, unitary matrices, singular value, Cholesky)
- Part 2: Basics and Preliminaries
  - Motivating systems with **M**ultiple **I**nputs and **M**ultiple **O**utputs (multiple access techniques)
  - General classification and description of MIMO systems (SIMO, MISO, MIMO)
  - Mobile Radio Channel
- Part 3: Information Theory for MIMO Systems
  - Repetition of IT basics, channel capacity for SISO AWGN channel
  - Extension to SISO fading channels
  - Generalization for the MIMO case
- Part 4: Multiple Antenna Systems
  - SIMO: diversity gain, beamforming at receiver
  - MISO: space-time coding, beamforming at transmitter
  - MIMO: BLAST with detection strategies
  - Influence of channel (correlation)
- Part 5: Relaying Systems
  - Basic relaying structures
  - Relaying protocols and exemplary configurations

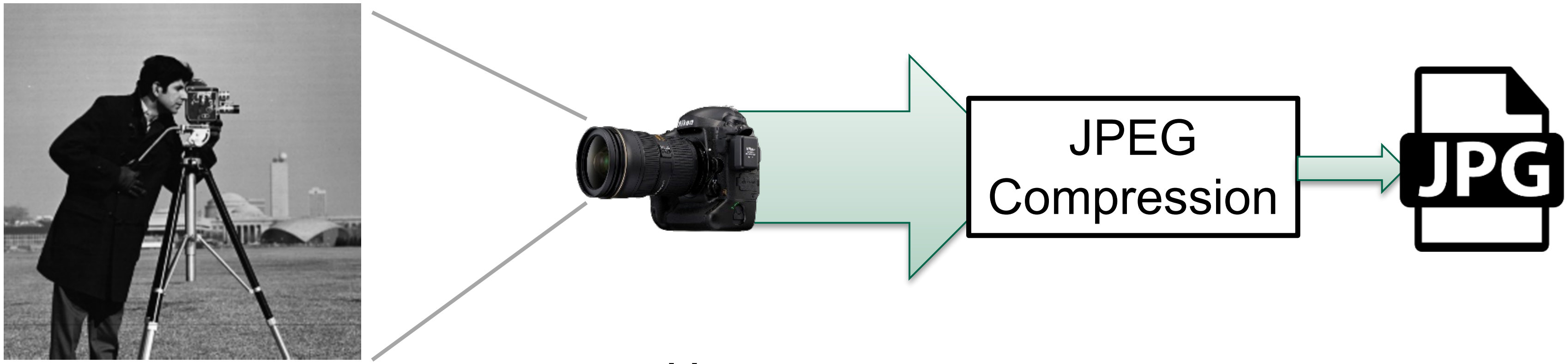
Outline

- Part 6: In Network Processing
  - Basic of distributed processing
  - INP approach
- Part 7: Compressed Sensing
  - Motivating Sampling below Nyquist
  - Reconstruction principles and algorithms
  - Applications

Compressive Sensing

- Motivation
- Basic ideas of Compressed Sensing
  - Undersampling / Underdetermined Systems
  - Sparsity
- Reconstruction principles
  - Basic Optimization task
  - Relaxations and Algorithms
  - Reconstruction Guarantees
- Applications: Sporadic Massive Machine Communications
  - Model and application of CS
  - Differences to standard CS assumptions
  - Adapted and novel CS Multi-User Detection Algorithms

Compressive Sensing Motivation



Scene

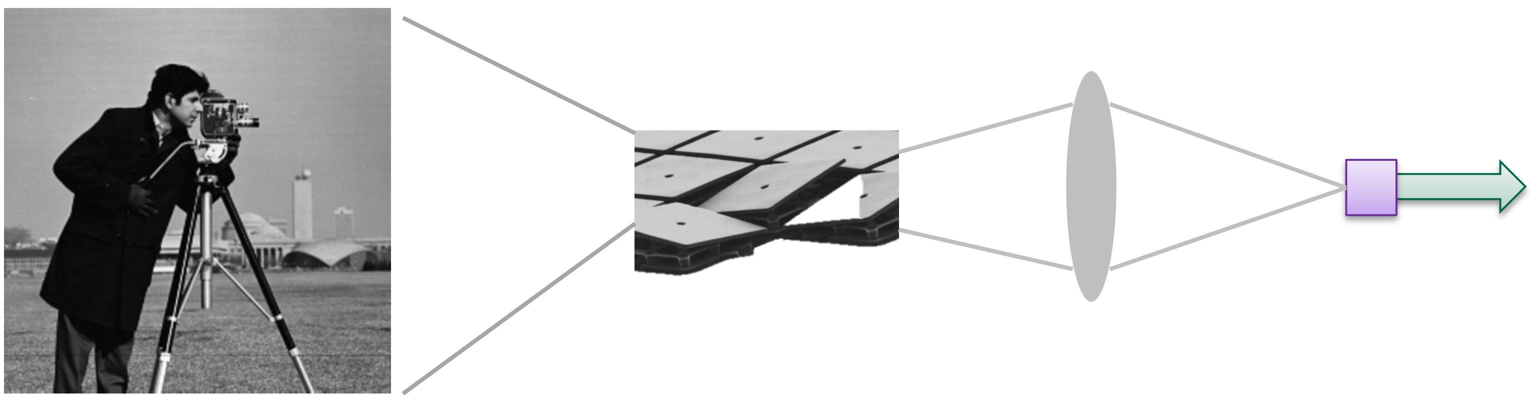
Huge amount of data sampled

Compact file after compression

- Todays signal acquisition systems are often wasteful
  - Huge effort in sampling with high accuracy
  - Removal of redundant information

Sampling, then compress

A First CS Example: The Single Pixel Camera

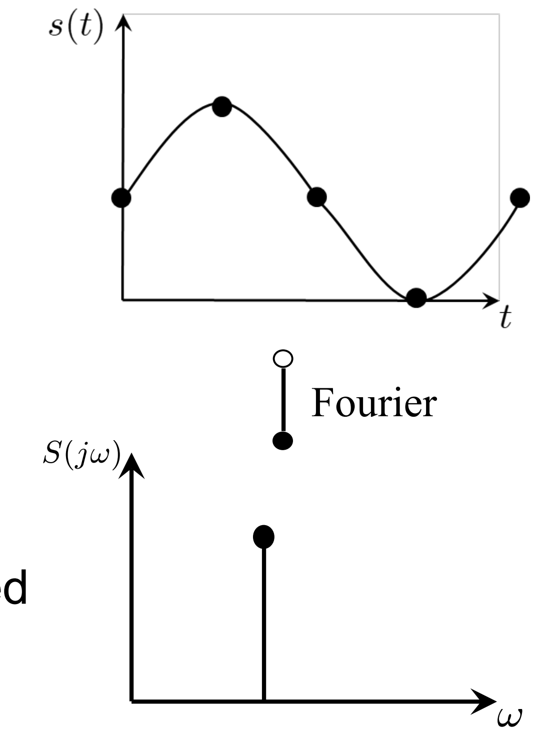


- Idea: Mix scene contents through DMD randomly
  - DMD consists of a high number of tilting micro mirrors
  - Each pixel measurement contains the whole scene

→ Compressed Sampling

Classic Sampling: Shannon/Nyquist

- Assumption:
  - Bandlimited signals with maximum frequency $f_{\max}$
- Shannon/Nyquist sampling:
  - Sampling frequency $f_s \geq 2f_{\max}$
  - Perfect reconstruction by simple low pass interpolation
- **Compressible** signals
  - Lower information content than number of samples
  - Signal properties besides band limitation not considered



Exploit side information of sampled signals!

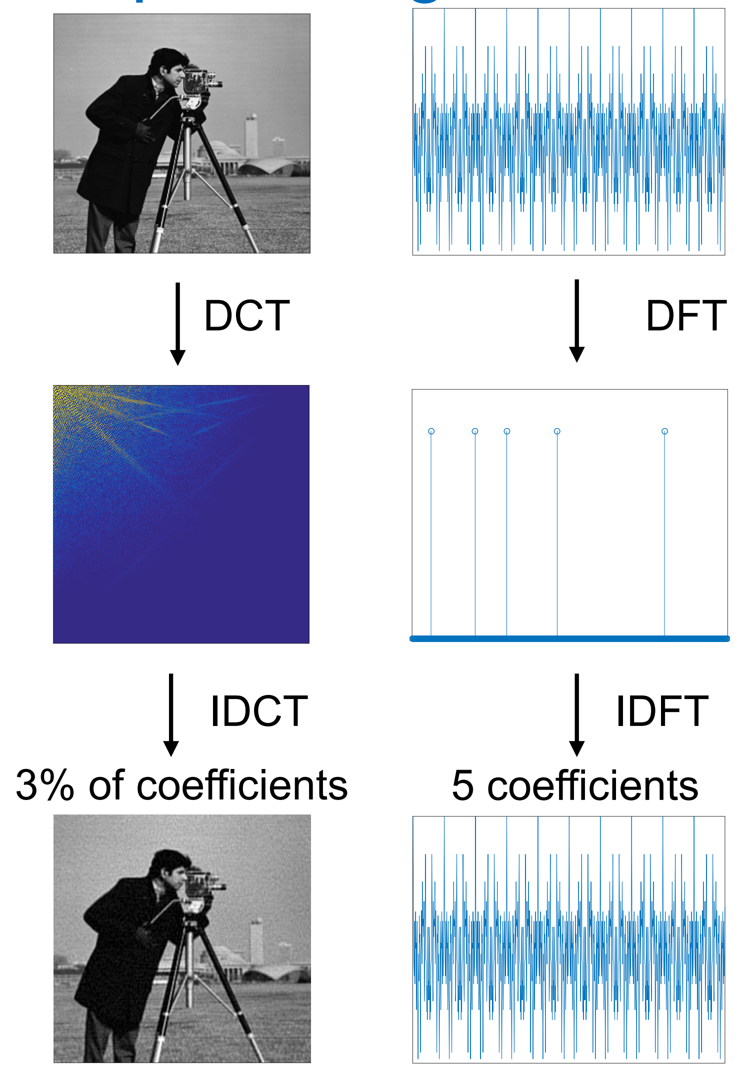
Base Assumption: Compressible / Sparse Signals

- Assume a **compressible** signal $\mathbf{z}$ of length N

- **Compressible**:
  few coefficients in other domain are sufficient
  - Discrete Cosine Transform (DCT)
  - Discrete Fourier Transform (DFT)

- K-sparse representation: $\mathbf{z} = \mathbf{\Psi}\mathbf{x}$
  - $\mathbf{\Psi}$ basis in which $\mathbf{z}$ is compressible / sparse
  - Only K biggest coefficients are relevant

    or
  - Signal only contains exactly K components



The “compressive” in CS

- How to reduce the number of measurements below Nyquist?
  - Image example: sorting and “nulling” of low power coefficients

$$\boldsymbol{z} \approx \boldsymbol{\Psi\Omega}\mathbf{x}$$

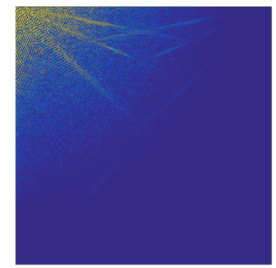
- Nyquist sampling of $\mathbf{z}$ first
- „Nulling“ matrix $\boldsymbol{\Omega} \in \mathbb{R}^{N\times N}$ after transformation

→ no reduction and content dependent

- General subsampling in Compressed Sensing
  - Should be independent of specific signals
  - Should be valid for all K-sparse vectors $\mathbf{x}$ independent of $\boldsymbol{\Psi}$
  - Linear mapping $\boldsymbol{\Phi}: \mathbb{R}^N \rightarrow \mathbb{R}^M$ from dense $\mathbf{z}$ to measurement $\mathbf{y}$

$$\mathbf{y}_N = \boldsymbol{\Phi\Psi}\mathbf{x}$$

DCT



IDCT

3% of coefficients

The Compressive Sensing Problem in a Nutshell

- Problem: Recover $\mathbf{x} \in \mathbb{R}^N$ from the $M < N$ measurements in vector

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \text{(underdetermined linear system)}$$

  - where $\mathbf{n} \in \mathbb{R}^M$ is additive noise and
  - $\mathbf{A} \in \mathbb{R}^{M \times N}$ is given by the sparsity basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ and measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$

$$\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}$$

- Notation: dense signal $\mathbf{z} = \mathbf{\Psi}\mathbf{x}$, noise free measurement $\mathbf{y}_N = \mathbf{A}\mathbf{x}$

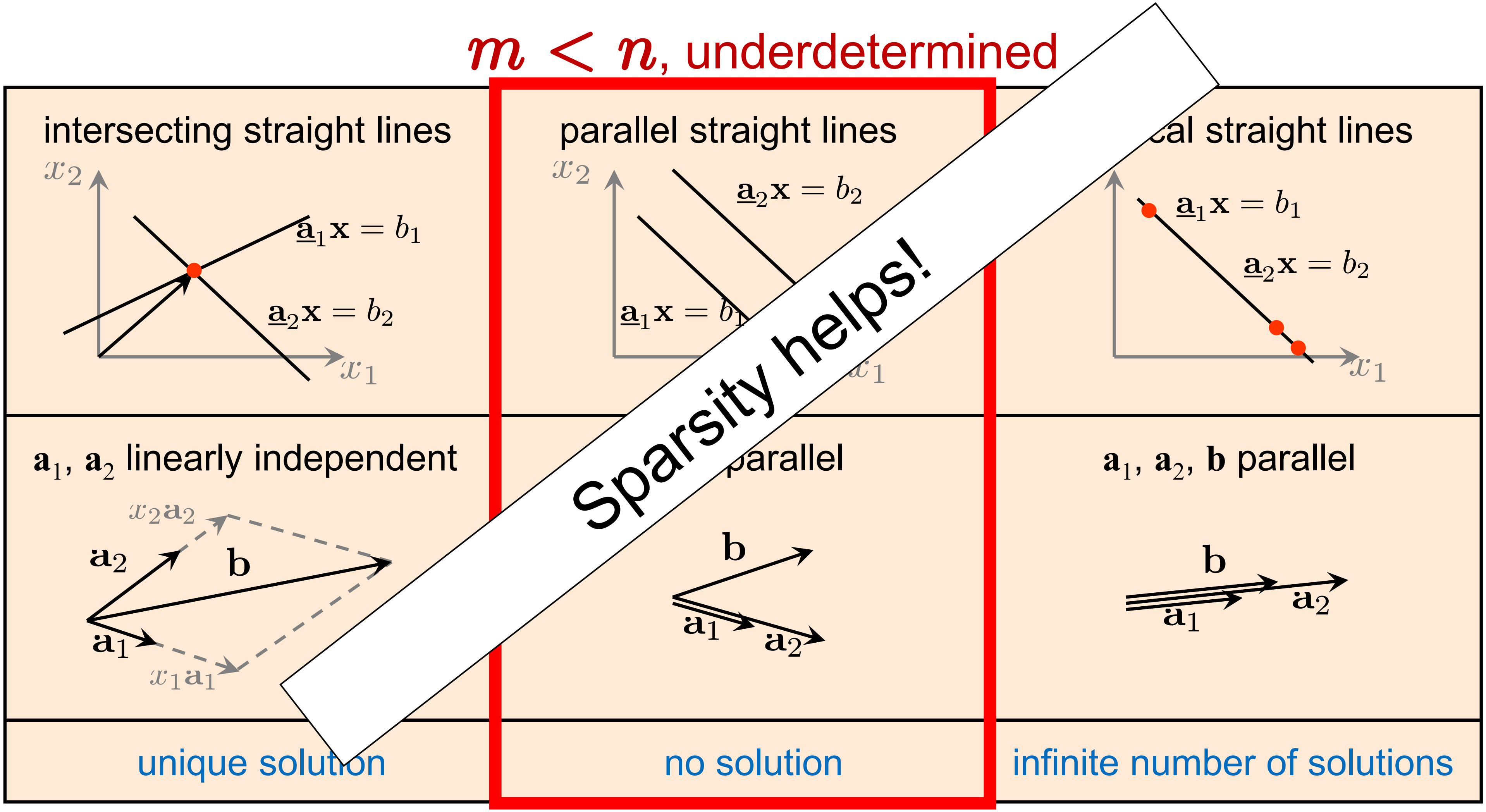
- Open Questions:
  - How to choose $\mathbf{\Phi}$ given a sparse representation in basis $\mathbf{\Psi}$?
  - How to reconstruct $\mathbf{z}$ given (noisy) measurements $\mathbf{y}$?
    subsampling violates sampling rate requirement
    → low-pass reconstruction impossible

CS History

- Compressive Sensing Theory originated in 2005/2006 [Candès+Tao], [Donoho]
- Reconstruction properties were quantified with the “Restricted Isometry Property” [Candès+Tao 2006]
- Reconstruction algorithms based on L1/L2-optimization has been widely studied [Candès+Tao 2005] and based on Matching Pursuit approaches has been adapted to CS [Tropp 2007] and further developed [Needell 2008, Dai 2009]
- Special Issue in Signal Processing Magazine, March 2008
- IEEE Transaction on Information Theory: Series on Compressive Sensing
- Application of CS in wireless communications:
  - Channel estimation [Berger 2010]
  - Coding Theory [Dai 2009, Aggarwal 2009]
  - CDMA Transmission [Zhu+Giannakis 2010]

Repetition: Linear Equation Systems

$m < n$, underdetermined



S-sparse

- K-sparse is used as a measurement for how sparse a vector is
- Definition: a vector $\mathbf{x}$ is K-sparse, if

$$\|\mathbf{x}\|_0 \leq K$$

where $\|\mathbf{x}\|_0 = |\text{supp}(\mathbf{x})| = |\{j : x_j \neq 0\}|$ ⟵ l_0-“norm“

Example: 3-sparse

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad \begin{pmatrix} x_1 \\ 0 \\ 0 \\ x_2 \\ 0 \\ \vdots \end{pmatrix} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ 0 \\ \vdots \end{pmatrix}$$

3-sparse — 3-sparse — not 3-sparse

Recovery by l_0/l_2 Optimization

- Underdetermined equations systems can be solved if $\mathbf{x}$ is S-sparse
  - Minimizing the constrained l_0-"norm" yields the sparsest feasible solution

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{A} \cdot \mathbf{x}\|_2^2 < \epsilon$$

Least-Square-Norm

  - Solves Least-Squares problem (Zero-Forcing) with the sparsest solution

    $\rightarrow$ If the position of non-zero entries is known, the reduced problem is solvable!

- Problem:
  - the l_0-"norm" is not convex
  - problem is NP-hard

- Approaches: Approximate l_0-"norm", suboptimal Greedy algorithms

Recovery of Sparse Signals

- Algorithms to solve the underdetermined linear systems for S-sparse solutions can be categorized in three classes
- Convex relaxation:
  Solve a convex program whose minimizer approximates the target signal
  → e.g. interior-point methods, projected gradient methods
  + succeed with very few measurements
  - computational intensive
- Greedy pursuits:
  Find sequentially the support for each active element of $\mathbf{x}$ in measurement $\mathbf{y}$ → e.g. Orthogonal Matching Pursuit (OMP)
  + low complexity
  - less sampling efficient, highly sensitive to correlation properties of matrix $\mathbf{A}$
- Bayesian Methods:
  Belief Propagation and approximations with sparsity inducing priors
  + very sampling efficient
  - high to moderate complexity and very dependent on prior assumptions

Least Absolute Shrinkage and Selection Operator (LASSO)

- Convex Relaxation: approximate the l_0-"norm" is by l_1-norm termed LASSO

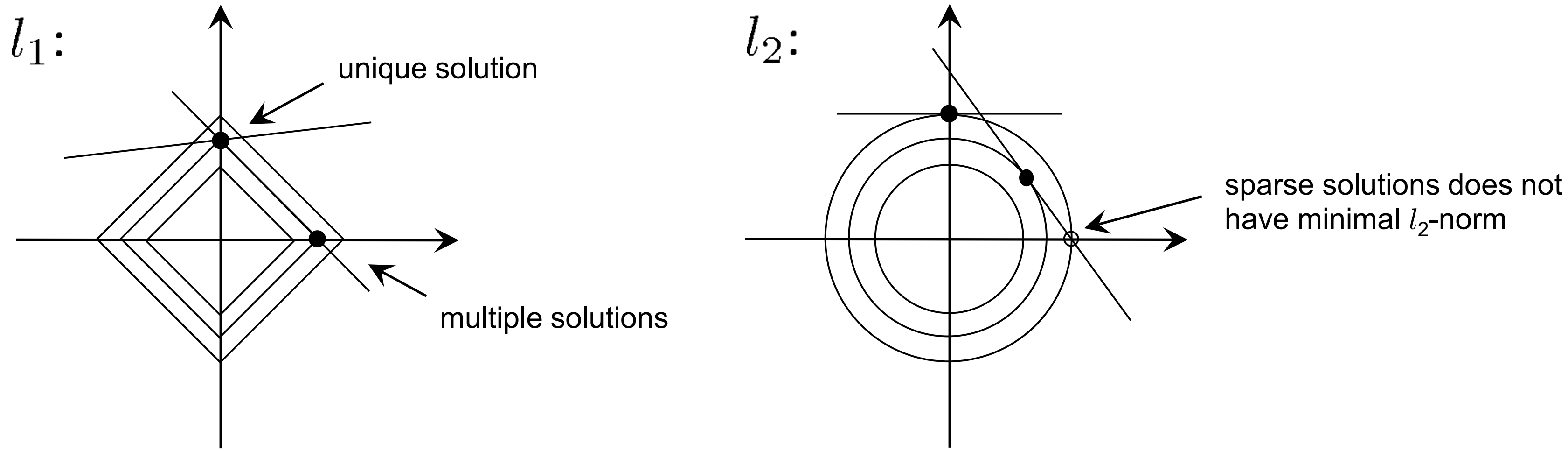
$$\hat{\mathbf{x}}^{\mathrm{LASSO}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A} \cdot \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

- Known from regression analysis [Tibshirani 1996]
  - LS-solution regularized by the l_1-norm
  - Convex problem, can be cast as a quadratic program (for a given λ)
  - Bayes estimation: λ is Laplace-prior

- Task: optimum value of λ is in general not known
  $\Rightarrow$ has to be estimated or determined iteratively

Least Absolute Shrinkage and Selection Operator (LASSO)

- The minimum l_1-norm, defined as $\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$, is convex and favors sparse solutions (noiseless case $\mathbf{y}=\mathbf{Ax}$)



Matlab Demo

Greedy Algorithms: Matching Pursuit

- Idea: Iteratively increase the support by the element x_i with the highest correlation to measurement $\mathbf{y}$
- Orthogonal Matching Pursuit (OMP):
  1. Determine index $\lambda_t = \arg \max_{j=1,\ldots d} |\langle \mathbf{r}_{t-1}, \mathbf{a}_j \rangle|$
  2. Augment matrix of chosen element $\mathbf{A}_t = [\mathbf{A}_{t-1} \ \mathbf{a}_{\lambda_t}]$
  3. Solve LS-problem $\mathbf{x}_t = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}_t \mathbf{x}\|_{l_2}$
  4. Calculate new residual $\mathbf{r}_t = \mathbf{y} - \mathbf{A}_t \mathbf{x}_t$
  5. Repeat until K elements were chosen
- Other variants exist, e.g., OLS, StOMP, ROMP, …
  - Pros: Low complexity and fast execution
  - Cons: Sparsity K has to be known / Proper stopping criteria have to be found
    Highly sensitive to correlation properties of matrix $\mathbf{A}$
    Less sampling efficient

Compressed Sensing

- Problem: Recover $\mathbf{x} \in \mathbb{R}^n$ from the $m < n$ measurements in vector

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \text{(underdetermined linear system)}$$

  - where $\mathbf{n}$ is additive noise and $\mathbf{A} \in \mathbb{R}^{m \times n}$ is given by the sparsity
  - basis $\boldsymbol{\Psi}$ (e.g. Fourier basis) and measurement matrix $\boldsymbol{\Phi}$ (e.g. identity)

$$\mathbf{A} = \boldsymbol{\Phi}\boldsymbol{\Psi}$$

- Notation: dense signal $\mathbf{z} = \boldsymbol{\Psi}\mathbf{x}$, noise free measurement $\mathbf{y}_\mathbf{N} = \boldsymbol{\Phi}\boldsymbol{\Psi}\mathbf{x}$
- Question: How to recover $\mathbf{x}$ from the under-determined equation system?
- Assumptions:
  - The signal x is sparse, but it is unknown which entries are non zero

Why does it work and when does it work?

- Coherence:

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \max_{1 \leq k,j \leq N} \frac{|\langle \boldsymbol{\phi}_k, \boldsymbol{\psi}_j \rangle|}{\|\boldsymbol{\phi}_k\|_2 \|\boldsymbol{\psi}_j\|_2}$$

  - measures the maximum correlation between any two elements of $\mathbf{\Phi}$ and $\mathbf{\Psi}$
  - can be checked in practice
  - recovery of $\mathbf{x}$ with ℓ_1-minimization is exact, with probability exceeding $1 - \delta$, if

$$M \geq C \cdot \mu^2(\mathbf{\Phi}, \mathbf{\Psi}) \cdot K \cdot \log(N/\delta)$$

Compressed Sensing only works well for low coherence $\mu(\mathbf{\Phi}, \mathbf{\Psi})$

Example: Fourier Basis $\mathbf{\Psi}$ and random Subsampling by $\mathbf{\Phi}$ (Gaussian, Spike, etc.)

Reconstruction Criteria and Guarantees

- Restricted Isometry Property (RIP): For each positive integer $K = 1,2,\ldots$ define the isometry constant δ_K of a matrix $\mathbf{A}$ as the smallest integer such that

$$(1-\delta_K)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1+\delta_K)\|\mathbf{x}\|_2^2$$

  holds for all K-sparse vector $\mathbf{x} \notin \mathrm{Ker}(\mathbf{A})$

  - Isometry: “does not distort space”
  - If RIP of order $2K$ shall hold for $\delta_{2K} \in (0, 1/2]$ then M measurement are required

$$M \geq C \cdot K \cdot \log\left(\frac{N}{K}\right)$$

  Example: N=10000, K = 20 → $M \approx 40$ required

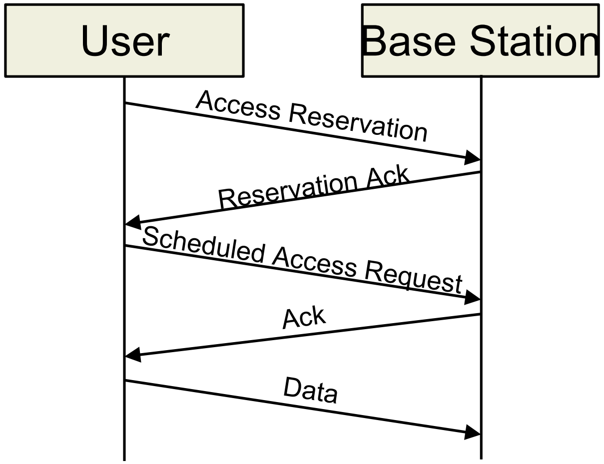
RIP is required for most reconstruction guarantees
→ Random sub-gaussian distributed matrices!

Applications of CS in Communications

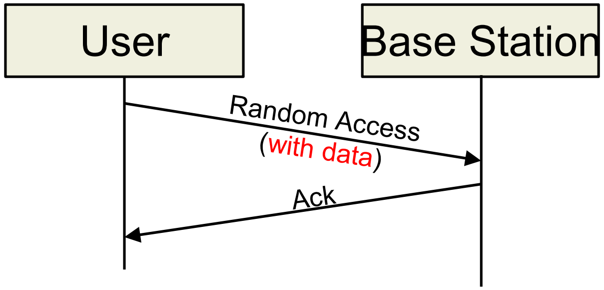
- Sampling of sparse Signals
  - Sampling and transmission of sparse signals, e.g. sensor networks
- Channel Estimation
  - Consider tapped delay line with non-uniform delay distribution
    → sparse wireless channel in terms of delay
  - Most wireless channels are sparse in sample clock
- Spectrum Sensing
  - Cognitive radio idea: find unused spectrum / time in which spectrum is free
  - Either spectrum or the edges of spectrum in time and frequency are sparse
- Sporadic Communication
  - Traffic characteristic of machine type traffic leads to sparse detection problems

Massive Machine Communication

- Today's (cellular) systems (3G/4G)
  - Designed for high data rate / large packets
  - Access reservation and scheduling
  - Control overhead is negligible vs. payload size
- Now: a new massive access problem
  - Massive number of nodes (sensors, etc.)
  - Typically low-data rates / small packets
  - Control overhead for scheduling non-negligible
- Potential solution
  - Reduce control signaling overhead by random access
  - No control overhead, simply send data
  - Major problem: user collisions!

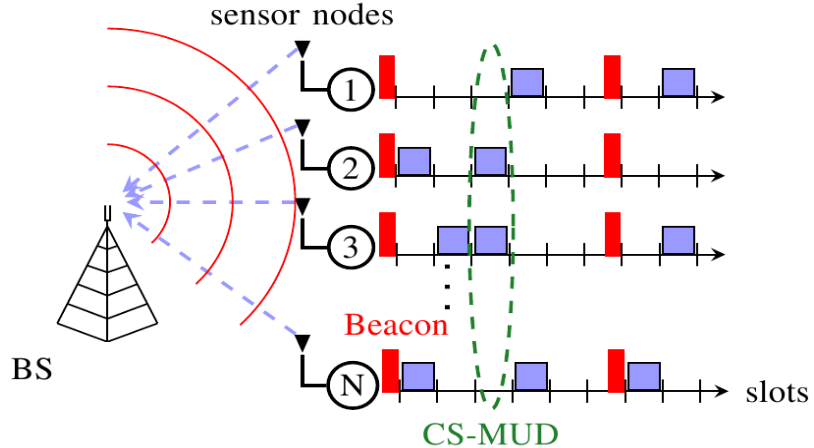


Access Reservation



Random Access

Sporadic Communication Scenario

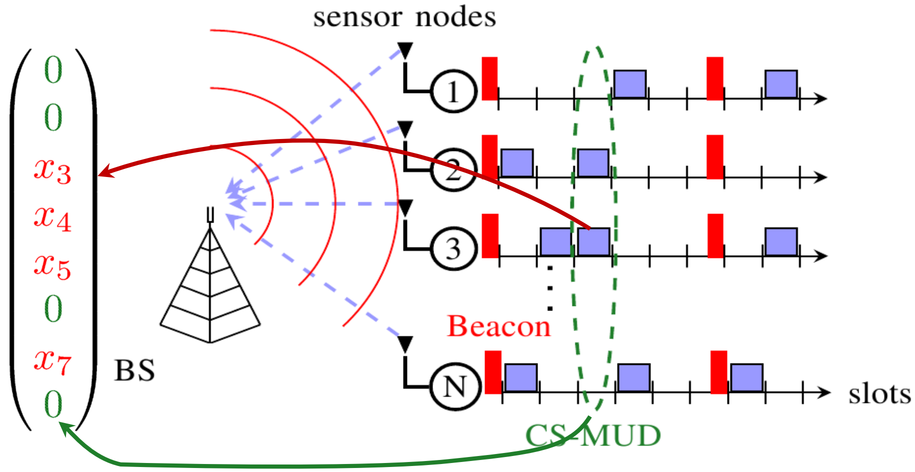


- Random access: M2M uplink communication
  - Base Station (BS) sends a beacon to synchronize sensors and define slots
  - Sensors send to a central Base Station using one or more slots
- Sporadic Communication: users are sporadically active
  - Event driven: e.g. a temperature threshold is met
  - Periodically: e.g. regular energy measurement (smart grid)
- Non-orthogonal medium access scheme
  - User signals interfere during transmission → Collisions!
  - Signal processing to reconstruct user signals → Multi-User Detection (MUD)

Compressed Sensing based Multi-User Detection

- Problem: How to recover sensor data and activity from observations?
- Sporadic communication
  - Inactive nodes “transmit” only zeros
  - Active nodes transmit data symbols

  → **The multi-user vector is sparse**

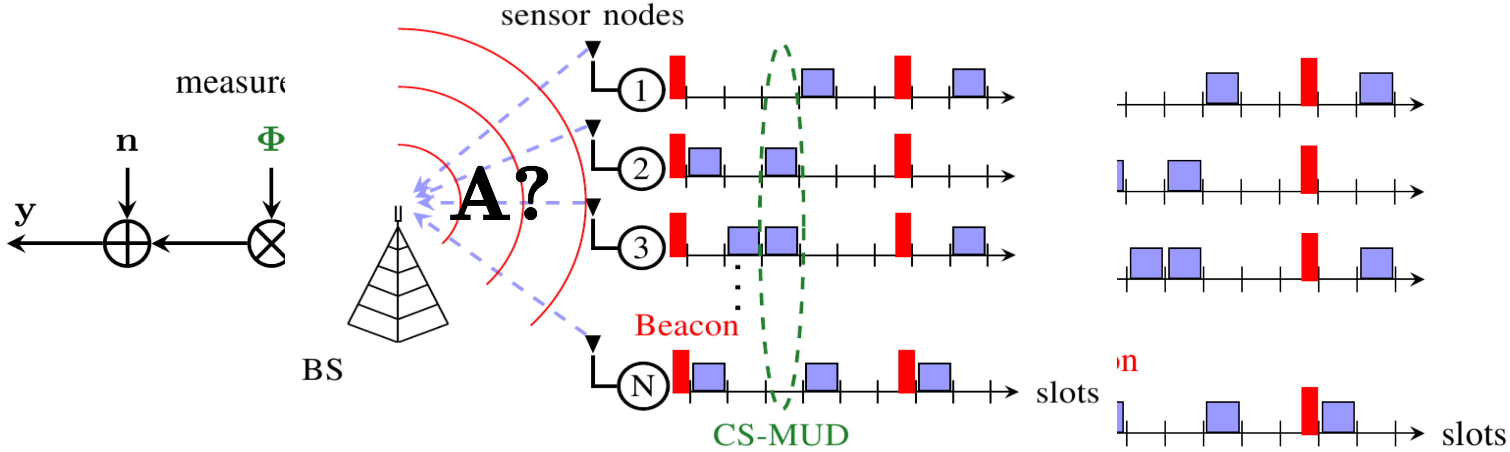


- Idea: Multi-user detection by **Compressed Sensing** exploiting sporadic activity for **joint activity and data detection**

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \mathbf{n}$$

$\mathbf{A} \in \mathbb{C}^{M \times N}$ is a known matrix fully describing the system and $\mathbf{n}$ is additive noise

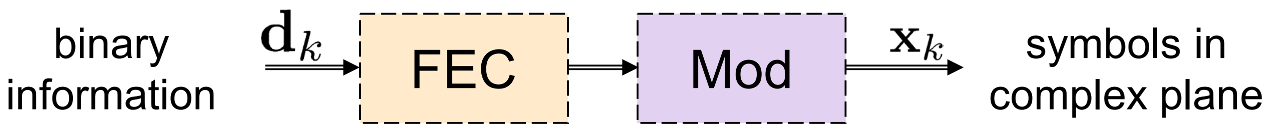
Differences to standard CS problems (1/3)



- $\mathbf{A}$ is actually not fully known!
  - Propagation channel is part of $\mathbf{A}$ and has to be estimated!
  - Unknown radio wave propagation modeled by channel $\mathbf{H}$
  - If a sensor is not active, its channel cannot be estimated
- In terms of common CS notation: $\mathbf{A}=\mathbf{\Phi}\mathbf{H}\mathbf{\Psi}$
  - Usually $\mathbf{\Phi}=\mathbf{I}$ → no compression here
  - Problem is under-determined because of $\mathbf{\Psi}$! → unusual for CS problems
  - $\mathbf{H}$ and x unknown → bilinear CS / CS 3.0

Differences to standard CS problems (2/3)

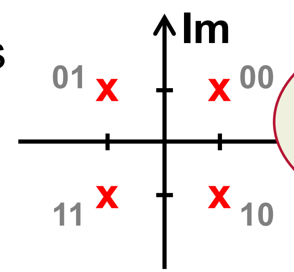
sensor signal



- Forward Error Correction: non-zero elements in $\mathbf{x}$ are part of a codeword
  - Communication systems use error correcting codes for robust communication

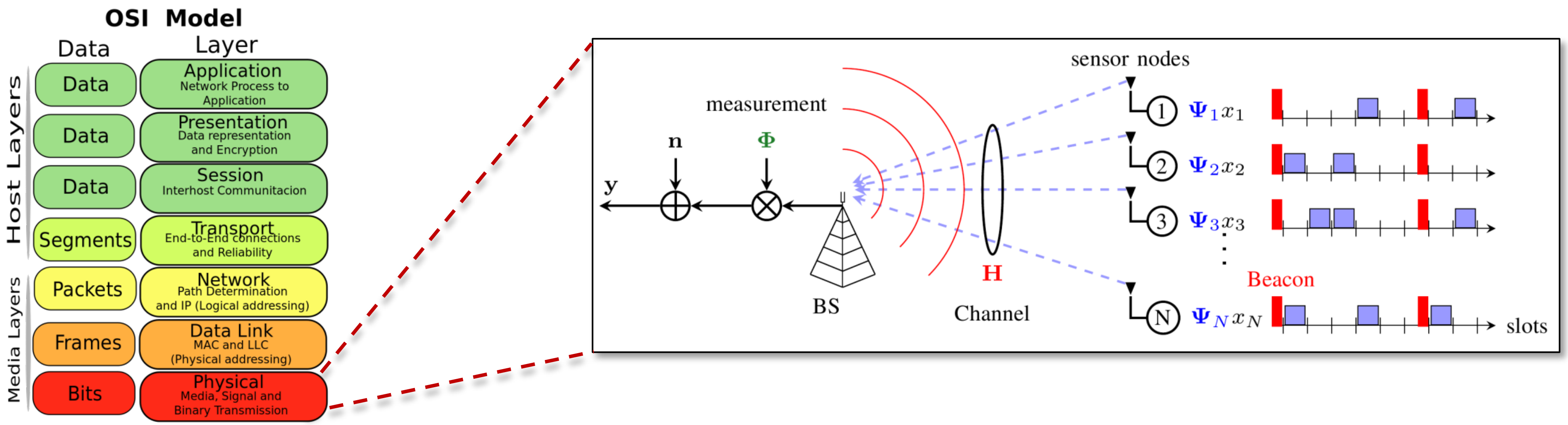
information		information + parity
0100110111	→	010011011110001

  - Structure that can be exploited, e.g. by iterative detectors/decoders
- Modulation: non-zero elements in $\mathbf{x}$ are not continuous
  - Communication systems use discrete symbol alphabets in the complex plane
  - Can be exploited, but requires adapted algorithms
  - New quality measure required: symbol error rate (SER)



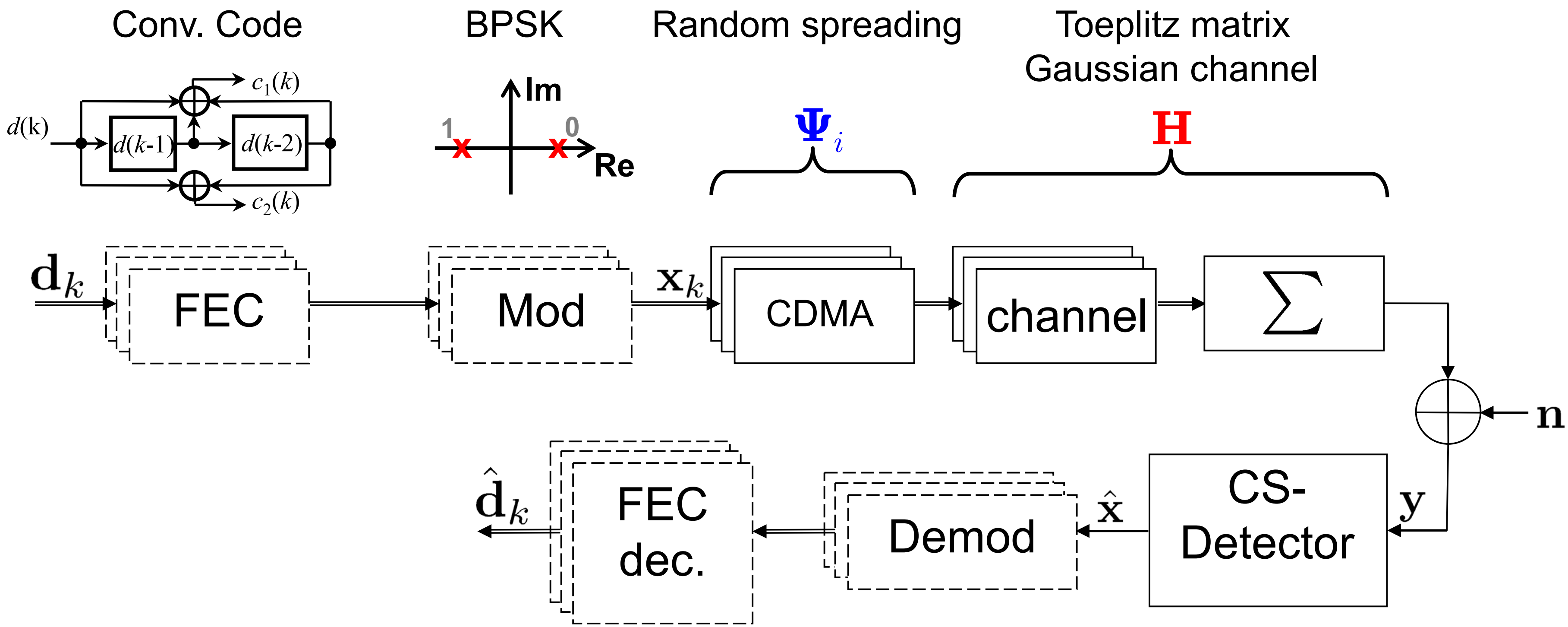
„Digital CS“

Differences to standard CS problems (3/3)



- System context: higher layers are impacted by reconstruction
  - CS theory mainly cares about perfect reconstruction and guarantees
  - But: differentiation of support errors is required
- Impact of False alarms / Missed detections
  - Missed detection: lost sensor information → not recoverable, retransmission
  - False alarm: misinformation → more processing at higher layers
- CS algorithms are not designed to control either support error rate!

Example: CDMA Transmission Model (UMTS)



$\mathbf{d}_k$ data of sensor k

$\mathbf{x}_k$ sensor specific symbols

$\mathbf{y}$ received vector

$\hat{\mathbf{x}}$ estimated (sparse) multi-user vector

$\hat{\mathbf{d}}_k$ estimated data of sensor k

Repetition: OMP

- Idea: Iteratively increase the support by the element x_i with the highest correlation to measurement $\mathbf{y}$
- Orthogonal Matching Pursuit (OMP):
  1. Determine index $\lambda_t = \arg \max_{j=1,\ldots d} |\langle \mathbf{r}_{t-1}, \mathbf{a}_j \rangle|$
  2. Augment matrix of chosen element $\mathbf{A}_t = [\mathbf{A}_{t-1} \ \mathbf{a}_{\lambda_t}]$
  3. Solve LS-problem $\mathbf{x}_t = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}_t \mathbf{x}\|_{l_2}$
  4. Calculate new residual $\mathbf{r}_t = \mathbf{y} - \mathbf{A}_t \mathbf{x}_t$
  5. Repeat until S elements were chosen

Group Orthogonal Matsching Pursuit (GOMP)

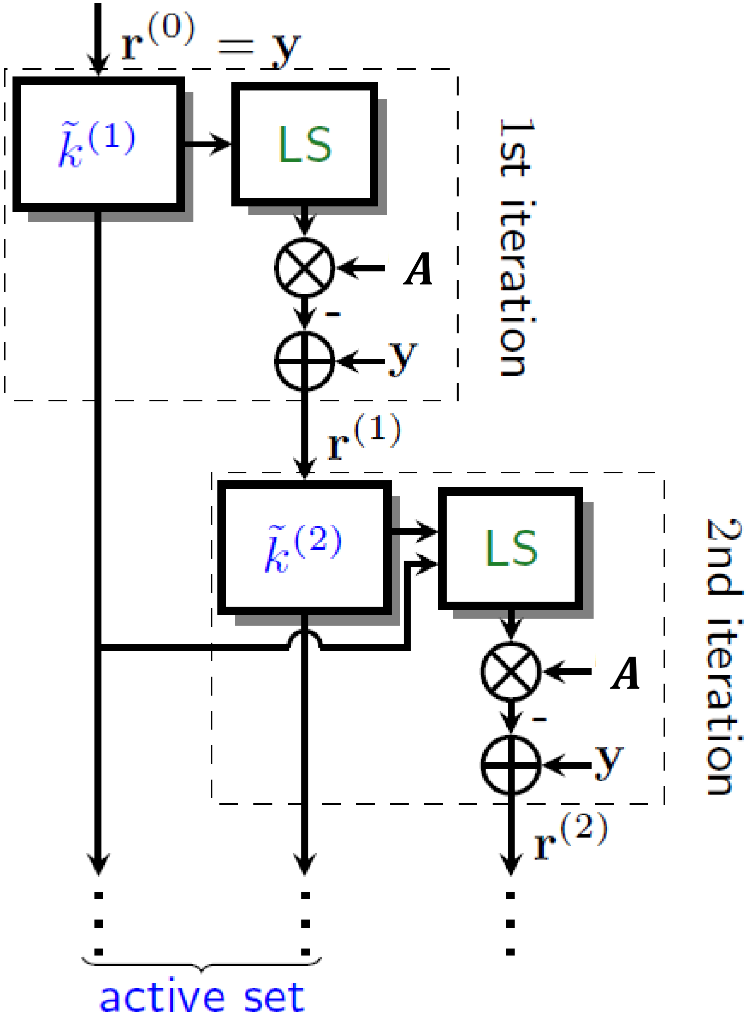
- Initialize $\mathcal{B}^{(0)} = \emptyset$, $\overline{\mathcal{B}}^{(0)} = \{1, \ldots, K\}$, $u = 0$, $\mathbf{r}^{(0)} = \mathbf{y}$
- Repeat
  - $u = u + 1$
  - $\tilde{k} = \arg\max_{k \in \overline{\mathcal{B}}^{(u-1)}} \frac{1}{|\gamma(k)|} \sum_{q \in \gamma(k)} \frac{\left| \boldsymbol{A}_{\{q\}}^T \mathbf{r}^{(u-1)} \right|}{\left\| \boldsymbol{A}_{\{q\}} \right\|_2}$
  - $\mathcal{B}^{(u)} = \mathcal{B}^{(u-1)} \cup \tilde{k}$ and $\overline{\mathcal{B}}^{(u)} = \overline{\mathcal{B}}^{(u-1)} \setminus \tilde{k}$
  - $\tilde{\mathbf{d}}_{\{\gamma(\mathcal{B}^{(u)})\}} = \boldsymbol{A}^{\dagger}_{\{\gamma(\mathcal{B}^{(u)})\}} \mathbf{y}$ and $\tilde{\mathbf{d}}_{\{\gamma(\overline{\mathcal{B}}^{(u)})\}} = \mathbf{0}$
  - $\mathbf{r}^{(u)} = \mathbf{y} - \boldsymbol{A}\tilde{\mathbf{d}}$
- Until stopping criterion is met

Activity detection

LS estimation

Group Orthogonal Matsching Pursuit (GOMP)

- Greedy algorithm for block-sparse reconstruction
- During iteration u:
  1. **Activity decision:**
     Select node $\tilde{k}$ with highest average correlation to residual $\mathbf{r}^{(u-1)}$
  2. **Data estimation:**
     Least-Square estimation for active set
  3. Compute new residual $\mathbf{r}^{(u)}$
- **Conclusion:**
  - Decision for active set not re-evaluated
  - Data estimation for entire active set



MAP-Detector with sparsity assumption

- In communication: transmit vector $\mathbf{x}$ is defined over discrete, finite alphabet $\mathcal{A}$
  - Greedy algorithms (e.g. OMP) and Convex relaxation methods (e.g. LASSO) are sub-optimal as they search over $\mathbb{C}^n$
- The best possible recovery of sparse signals of discrete, finite alphabets under noisy measurements is the Maximum a-posterori (MAP) detector
  - Proper model of the input distribution of $\mathbf{x}$ is required!
  - Zero / Active and Non-Zero / Active entries have to modeled
  - May be prohibitively complex for large systems!
- Augmented alphabet: Elements of $\mathbf{x}$ are taken from a finite alphabet $\mathcal{A}$ which includes the zero element, e.g. $x_i \in$ {-1,0,1} for a BSPK alphabet

Sparsity Aware MAP-Detector

- Assume: $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$ with $\mathbf{n} \in \mathcal{N}(0, \sigma_n^2)$
- Sparsity aware MAP-detector

$$\hat{\mathbf{x}}^{\mathrm{MAP}} = \arg\max_{\mathbf{x}\in\mathcal{A}^n} \Pr(\mathbf{x}|\mathbf{y}) = \arg\min_{\mathbf{x}\in\mathcal{A}^n} -\ln\Pr(\mathbf{y}|\mathbf{x}) - \ln\Pr(\mathbf{x})$$
$$= \arg\min_{\mathbf{x}\in\mathcal{A}^n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$$

sparsity awareness

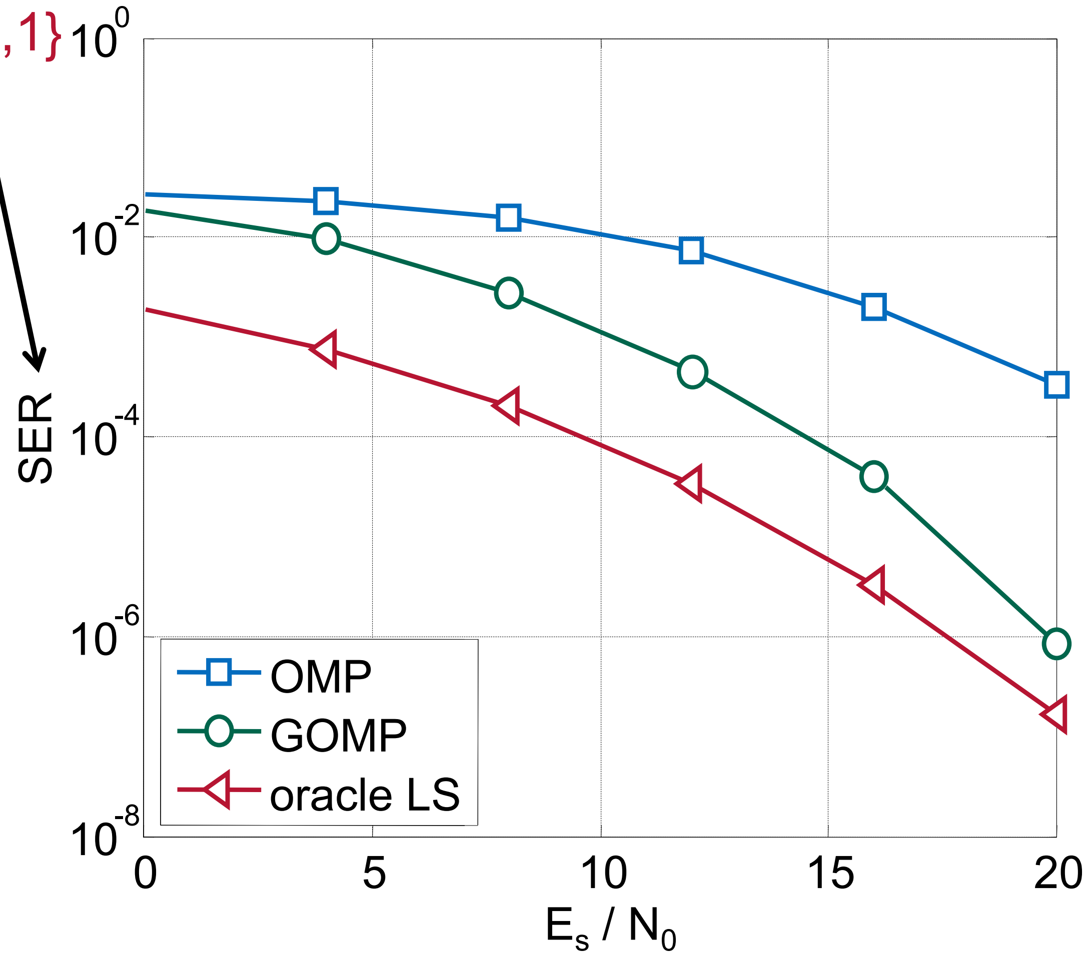
- with $\lambda = 2\sigma_n^2 \ln \frac{1-p_a}{p_a/(|\mathcal{A}|-1)}$, where p_a is the probability that an element x_i is not zero

- LS optimization over finite alphabet with penalty $\lambda\|\mathbf{x}\|_0$
- Penalty parameter λ is related to the a-priori information of elements x_i
- Algorithms: combinatorial search $\rightarrow$ extend Decision Directed Detector / Sphere decoding by l_0-norm

Overloaded CDMA Transmission

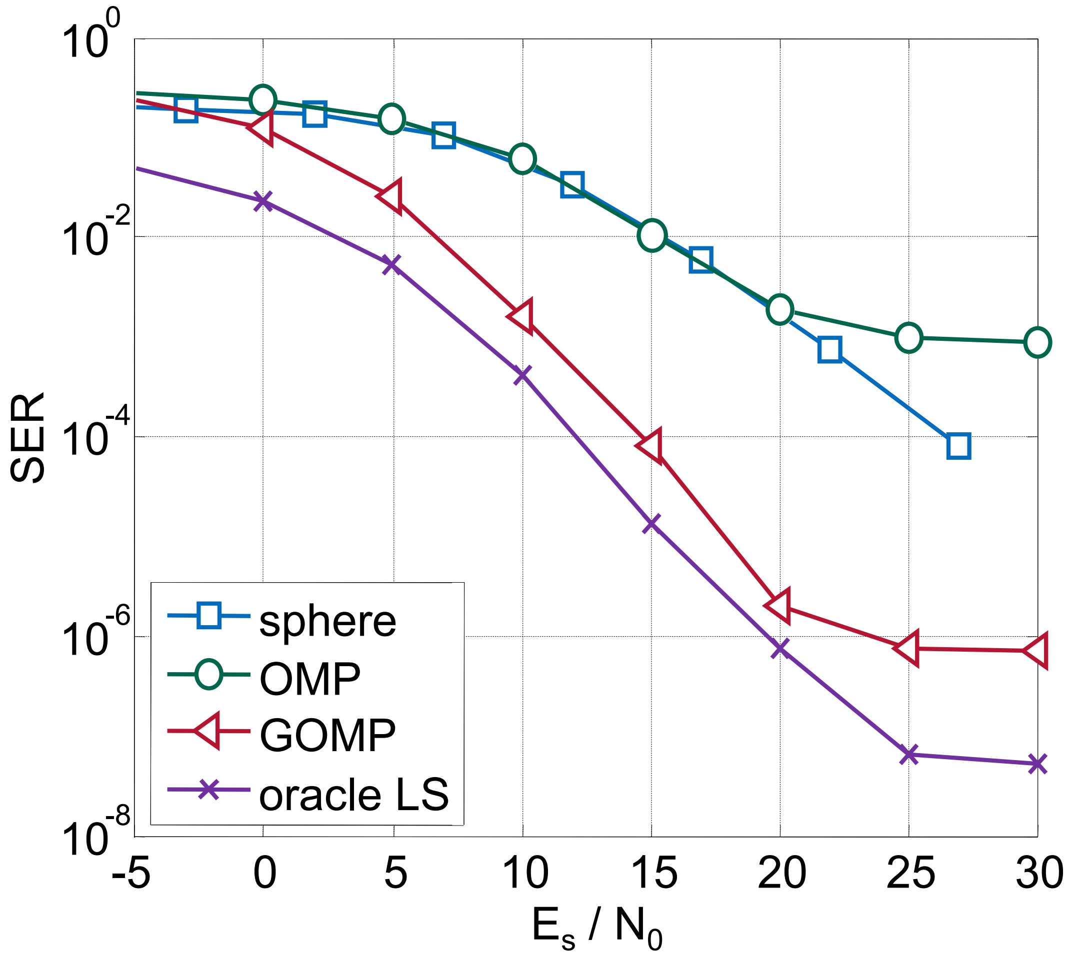
- Parameters
  - K=128 Users
  - L=8 Symbols per User
  - N=32 Spreading Sequence
  - F=256 Frame Length in Chips
- Observations
  - Underdetermined by factor 4, i.e., 4 times more users than length of the spreading sequence
  - OMP ignores and GOMP exploits block sparsity (L=8)
  - Exploitation of block sparsity improves performance significantly

Symbol Errors w.r.t {-1,0,1}



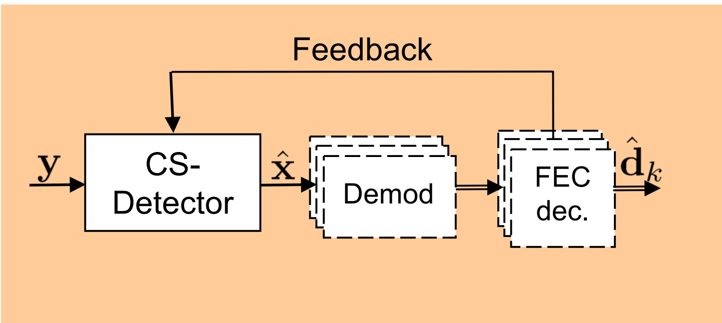
Underloaded CDMA Transmission

- Parameters
  - K=10 Users
  - L=10 Symbols per User
  - N=16 Spreading Sequence
  - F=160 Frame Length in Chips
- Observations
  - Overdetermined System $K<N$
  - OMP/GOMP show error floor
  - Sphere Detection just as good as OMP?

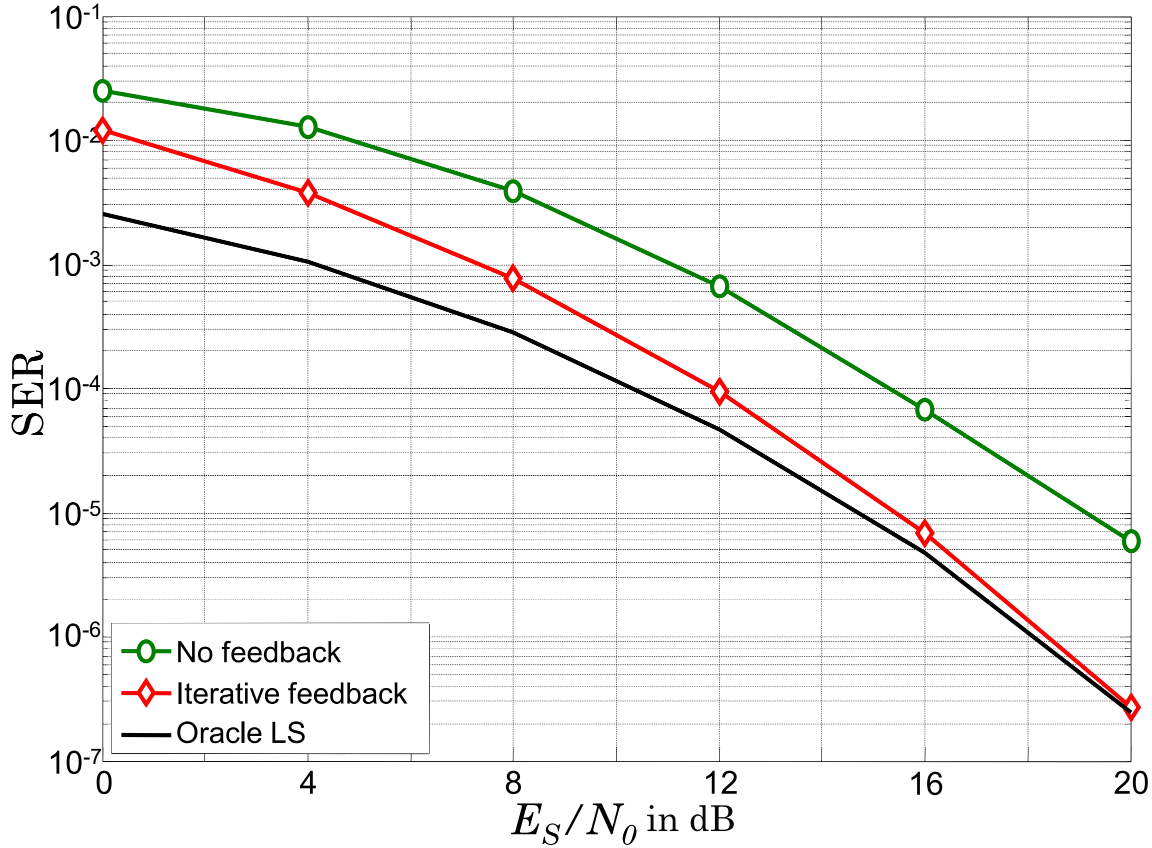


CS-MUD exploiting FEC with Iterative Feedback

- Idea: Use information of FEC to improve detection



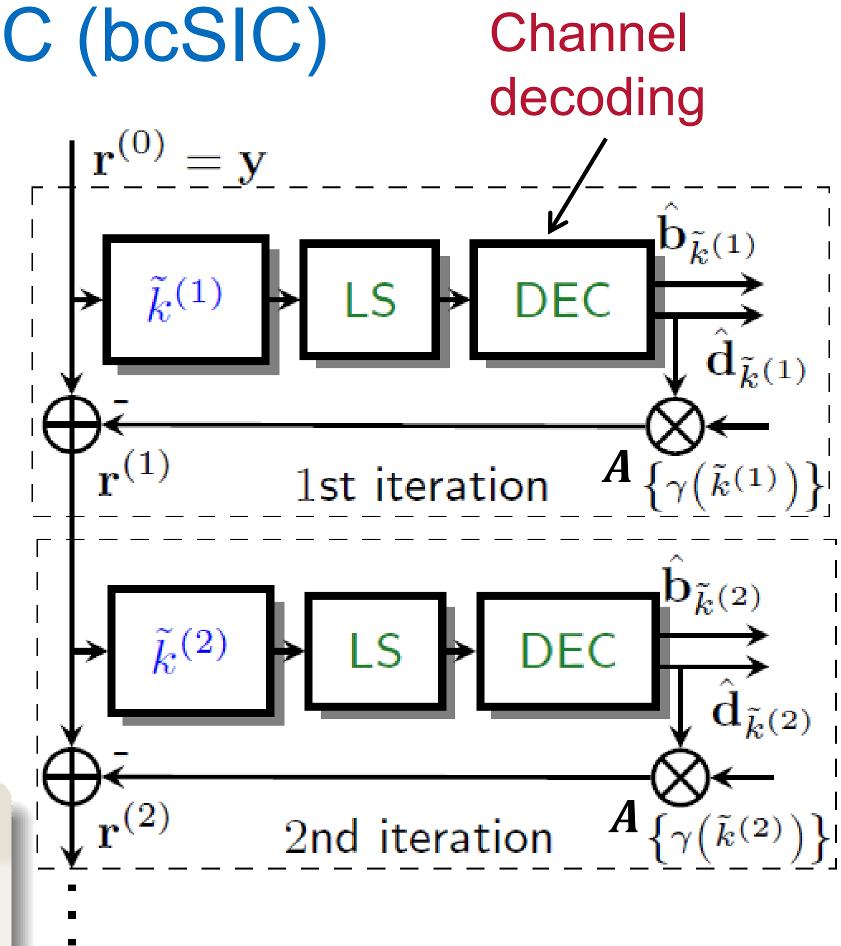
- Baseline / Classic Detection:
  Receiver knows active nodes, only data has to be estimated
- Result:
  CS-MUD nearly achieves performance of scheduled system (known activity)



Base CS Algorithm: GOMP, K=128 users, N=32 spreading sequence length, Activity probability 2%, BPSK symbols, Frames with 50 information symbols, [5;7] convolutional code, 3 Feedback iterations

Block correlation SIC (bcSIC)

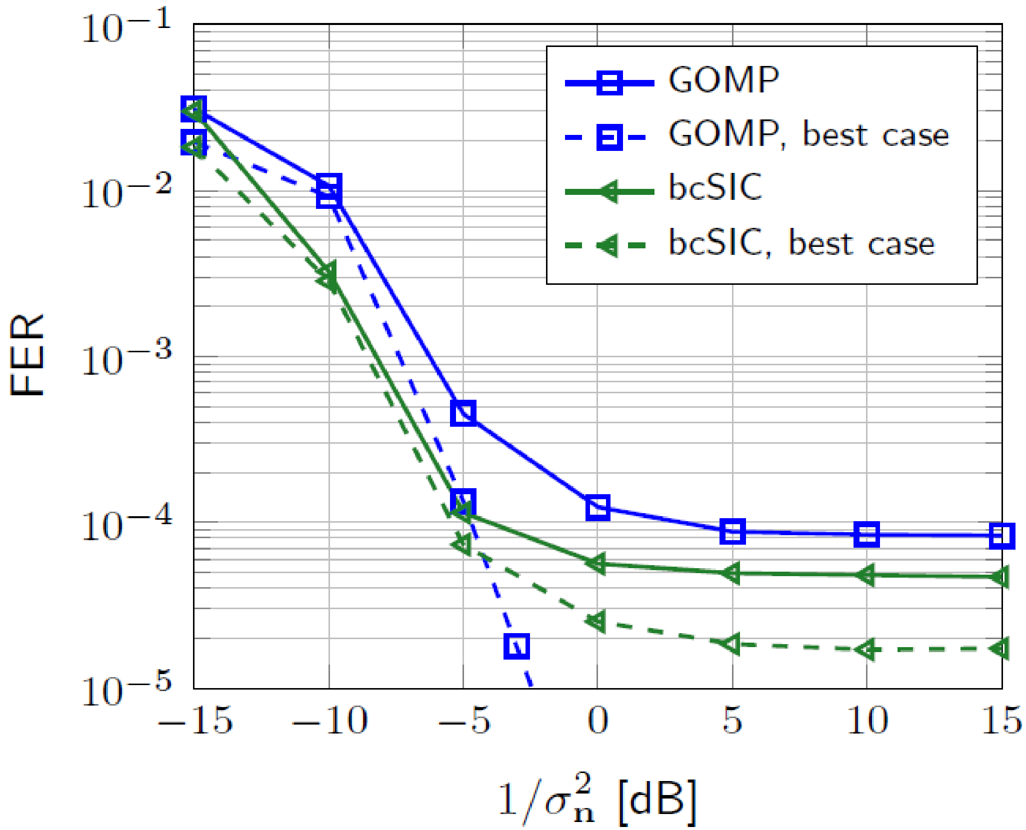
- Inspired by both SIC and GOMP
- During iteration u:
  1. **Activity decision:** Select node $\tilde{k}$ with highest average correlation to residual $\mathbf{r}^{(u-1)}$
  2. **Data detection:** Least-Square estimation for node $\tilde{k}$, followed by decoding
  3. **Update residual:** Subtract interference of node $\tilde{k}$ from residual $\mathbf{r}^{(u-1)}$



Difference to GOMP Iteration

GOMP: Estimate entire set of active nodes
bcSIC: Estimate and decode one node
$\Rightarrow$ bcSIC has lower complexity

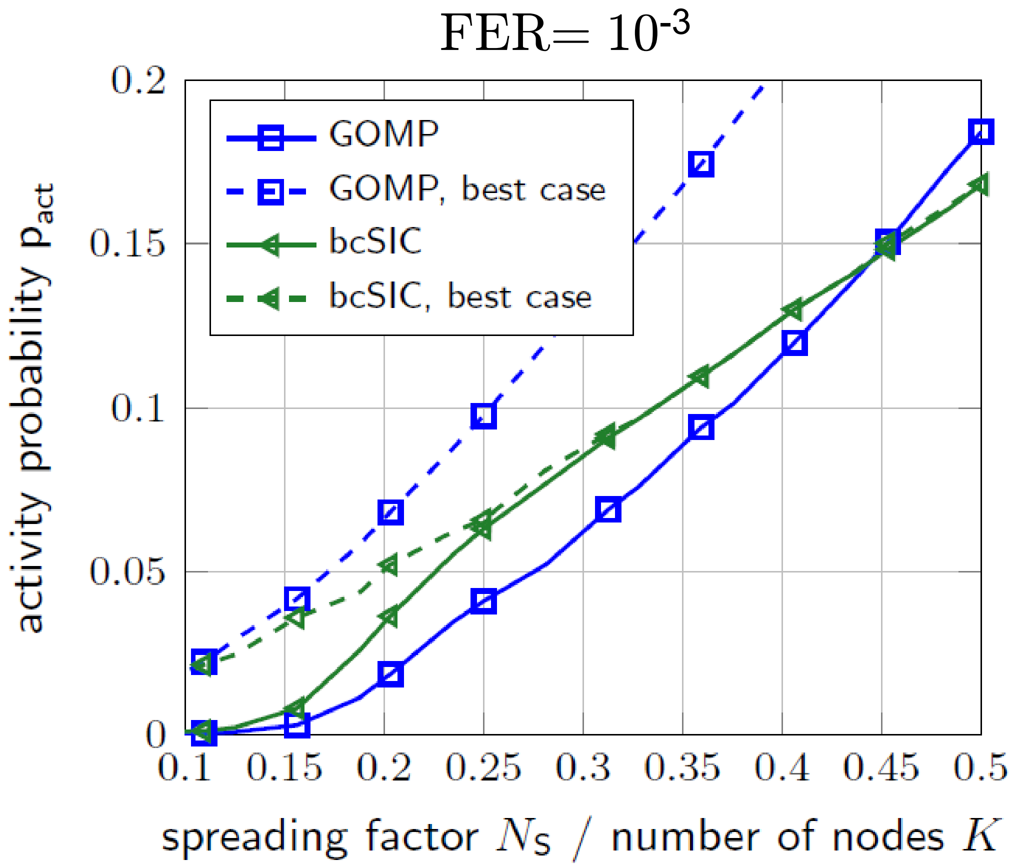
Comparison: GOMP vs. bcSIC



$N_S/K = 0.25$, $p_{act} = 0.02$, $\mathbf{H}_k = \mathbf{I}$,
$K = 64$ nodes, $L = 52$

- Best case: known activity
- Low SNR-range:
  - CS-MUD shows almost no loss
  - bcSIC outperforms GOMP
- High SNR-range:
  - bcSIC suffers from error propagation

Scalability of CS-MUD to keep FER performance

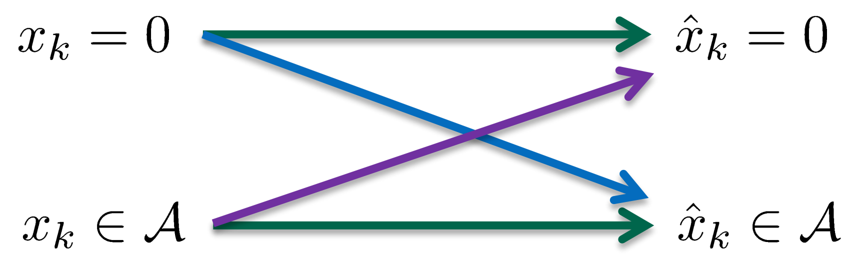


$1/\sigma_n^2 = 0\,\text{dB}$, $\mathbf{H}_k = \mathbf{I}$, $K = 64$ nodes, $L = 52$

- CS-MUD detectors scales gracefully with increased activity probability
- GOMP vs. bcSIC
  - bcSIC outperforms GOMP
  - bcSIC is of lower complexity

Activity Errors

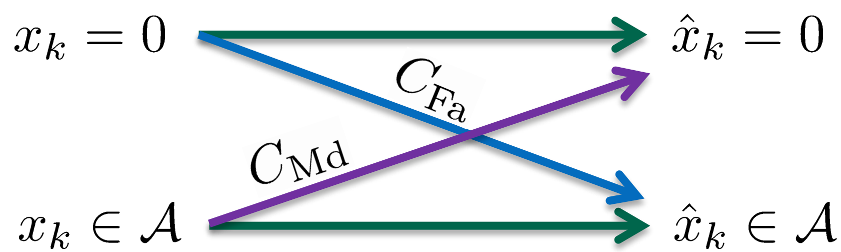
- SER / FER do not describe the complete behavior of the detector
  - Activity model leads to zeros in estimated user data
  - Symbol and Frame Errors include bit as well as activity errors
- Error events of activity at PHY:
  - Loss of data (Missed Detection/MD)
  - Pseudo data (False Alarm/FA)



- CS algorithms and MAP do not consider these classes
  - Impact on higher layers is generally not the same
  - Optimum trade-off may depend on higher layer processing

Bayes risk detector

- Approach: Minimize weighted risk of erroneous activity detection



$$R := C_{\mathrm{Fa}} \mathbf{Prob}\left(\hat{x} \in \mathcal{A} | x_k = 0\right) + C_{\mathrm{Md}} \mathbf{Prob}\left(\hat{x} = 0 | x_k \in \mathcal{A}\right)$$

# of non-zeros

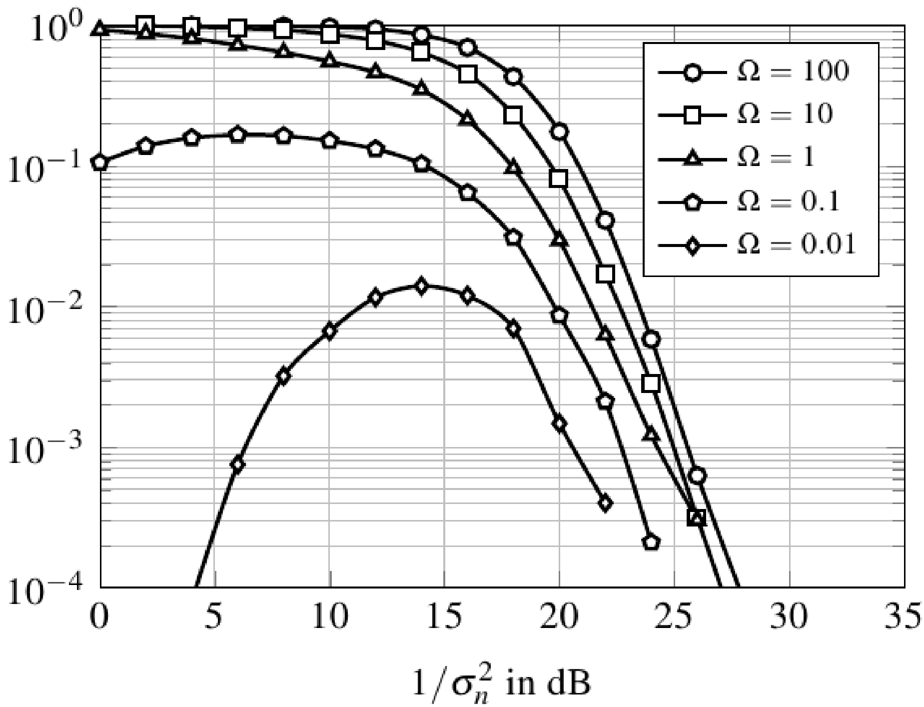
$$\hat{\mathbf{x}}^{\Omega} = \underset{\mathbf{x} \in \mathcal{A}_0}{\text{minimize}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \underbrace{2\sigma_n^2 \|\mathbf{x}\|_0 \ln\left(\Omega \frac{1 - p_a}{p_a / |\mathcal{A}|}\right)}_{\text{„Activity cost"}} \qquad \Omega = \frac{C_{\mathrm{Fa}}}{C_{\mathrm{Md}}}$$

- Tuning parameter Ω

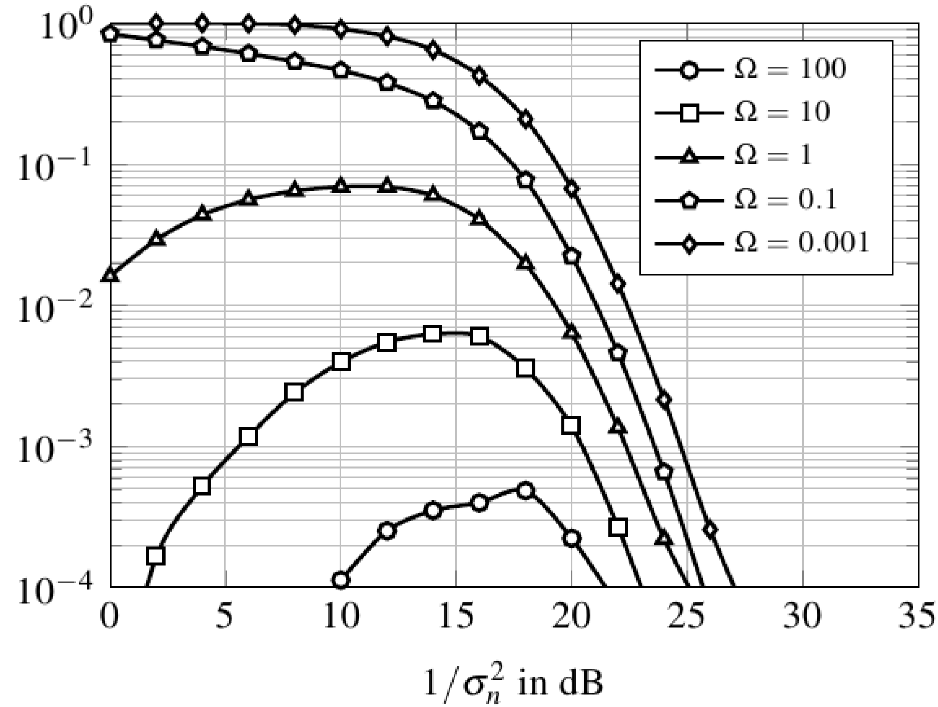
Prefer activity		Prefer inactivity
$\Omega < 1$	$\Omega = 1$	$\Omega > 1$
	MAP-Detector	

Bayes-Risk based activity detection

Missed Detections



False Alarms



- MAP detection $\Omega = 1$ is not sufficient for CS-MUD
- Bayes-Risk approach allows to trade-off between both activity error events
- Low Missed Detection rate automatically increases False Alarm rate and vice versa

Neyman-Pearson Approach

- Main idea: Tightly control one activity error while minimizing the other
  - Activity errors: False Alarm and Missed Detection

$$\hat{\mathbf{x}}^{\eta} = \arg\min \Pr(\mathsf{Fa}) \qquad \text{s.t.} \ \Pr(\mathsf{Md}) \leq \eta$$

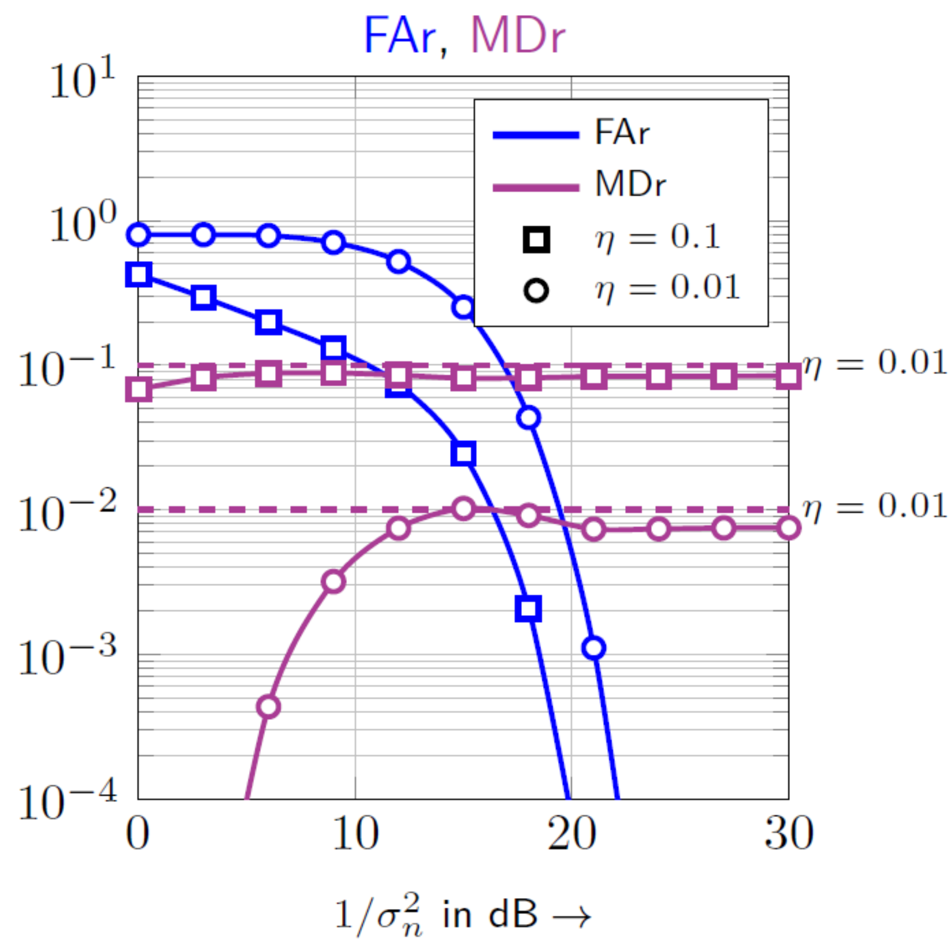
  - Bounds the probability of MD to η minimizes the probability of FA

- Problem: Neither probability can be formulated in closed form
  - Analytical solution not achievable
  - Approximation required

- Solution: Use soft-values for activity to estimate probabilites
  1. Soft-value calculation, e.g. by activity-MAP
  2. Estimation of probabilities (the more soft-values, the better)
  3. Minimization by threshold finding algorithm

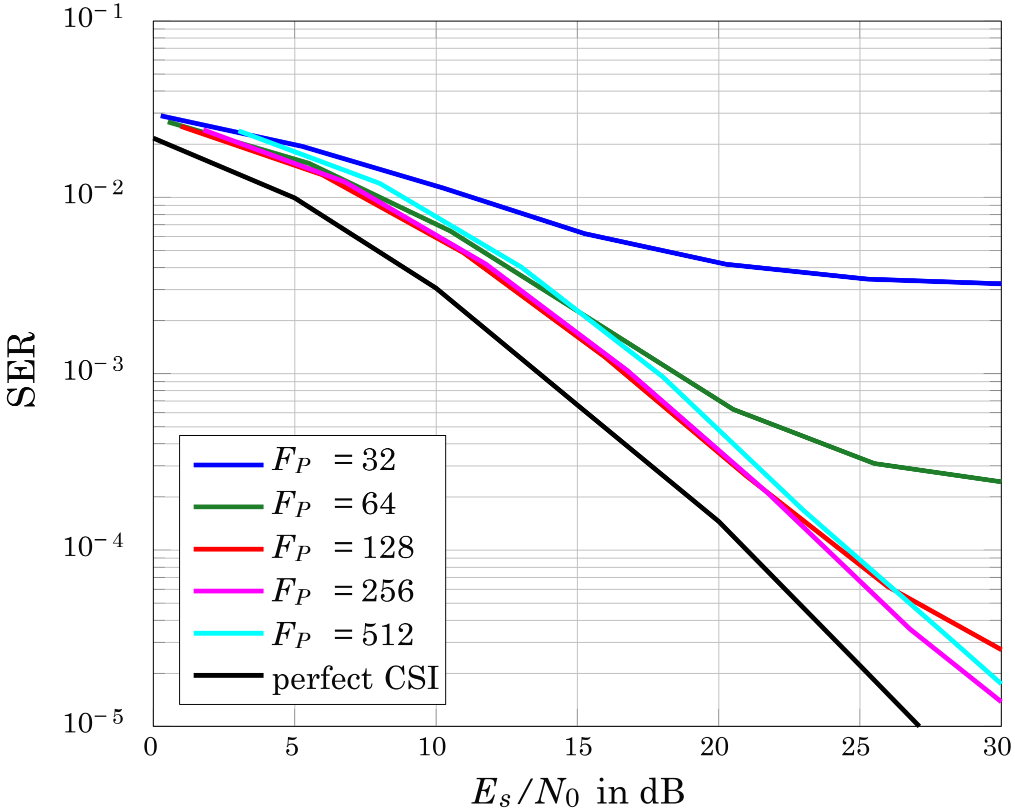
Neyman Pearson results

- Standard MAP
  - No control over False Alarm (FA) and Missed Detection (MD) rates
- Neyman-Pearson Approach:
  - No violation of MD constraint
  - MD constraint is over fullled for low target rates
  - Overfitting in low SNR range
  - FA rate increases for lower MD constraints
  - Unreliable LLRs lead to overfitting in low SNR range
- Conclusion
  - Allows perfect activity error control



CS-MUD Channel Estimation Results

- Main assumption:
  - Random pilots for each packet
  - E_s/N_0 loss included: $F / (F + F_P)$
- Observations:
  - Small F_P: SER increase dominates
  - Large F_P: E_s/N_0 loss dominates
  - Overall best choice here: $F_P = 128$
    - ¼ of frame length F
- Result
  - Joint channel and activity estimation by Compressed Sensing algorithms
  - Asynchronicity included by maximum channel delay $\tau_{max} = 20$ chips
  - Performance lost mostly SNR loss (pilot overhead)



Base CS Algorithm: modified GOMP, K=64 users, random spreading Activity probability 2%, BPSK symbols, Frames with 8 information symbols, 3-tap Rayleigh fading channel with random delays up to $\tau_{max} = 20$, exponentially decaying power → Asynchronous communication!

Outlook

- CS in (wireless) communication scenarios is a hot research topic
  - Joint Activity and Data detection by CS detectors is a promising approach for low overhead communication
  - Besides CDMA communication many other schemes can be applied (e.g. SC-FDMA like in LTE)
  - Many other areas are of interest: Decoding by CS, Channel Estimation, Radar Signal Processing, …