Exploiting Structure in Compressed Sensing Using Side Constraints – from Analysis to System Design (EXPRESS II)

CoSIP SPP 1798 Workshop

Prof. Martin Haardt Dr. Khaled Ardah FG Nachrichtentechnik, TU Ilmenau

Prof. Marius Pesavento Minh Trinh Hoang FG Nachrichtentechnische Systeme, TU Darmstadt

Prof. Marc Pfetsch Frederic Matter FG Diskrete Optimierung, TU Darmstadt

http://www.projekt-express.tu-darmstadt.de







TECHNISCHE UNIVERSITÄT



Overview



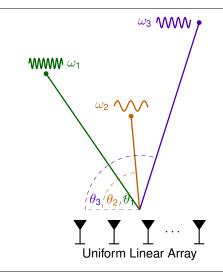


- Project Overview
- WP 0: Nonlinear Measurement Systems
- WP 1: Design of new Compression Matrices
 - Multidimensional Training Design
 - Identifiability
- ▶ WP 2 + 3: General Null Space Properties

Application Example Multidimensional Frequency Estimation



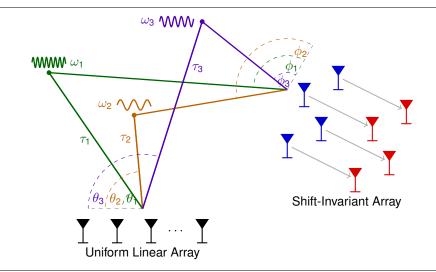




Application Example Multidimensional Frequency Estimation

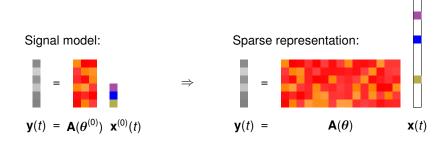


TECHNISCHE UNIVERSITÄT

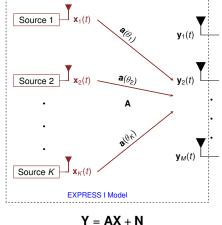


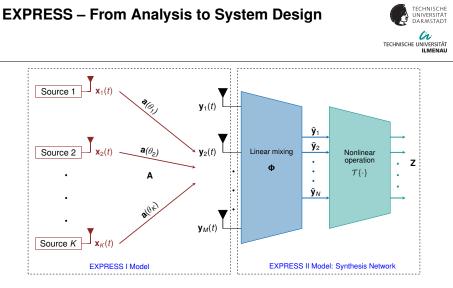
Low Rank Model and Sparse Representation – CS using Model Structure as Side Constraints





EXPRESS – From Analysis to System Design





$\mathbf{Z} = \mathcal{T} \left\{ \boldsymbol{\Phi} \mathbf{A} \, \mathbf{X} \right\} + \mathbf{N}$

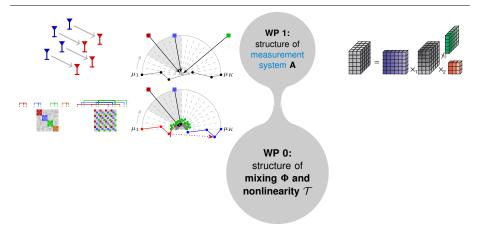




WP 0: structure of mixing Φ and nonlinearity \mathcal{T}

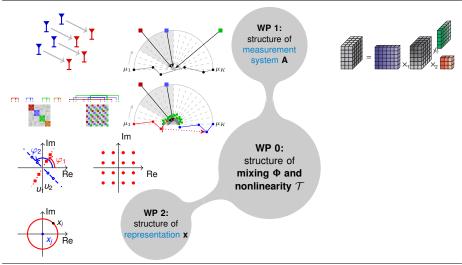


TECHNISCHE UNIVERSITÄT



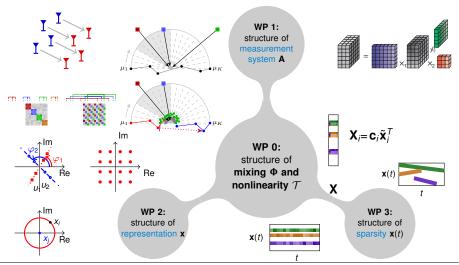


TECHNISCHE UNIVERSITÄT





CC TECHNISCHE UNIVERSITÄT ILMENAU



February 13, 2020 | SPP 1798 DFG - CoSIP - Workshop, Aachen | 6



TECHNISCHE UNIVERSITÄT

Nonlinear Measurement Systems

WP 0: structure of mixing Φ and nonlinearity ${\cal T}$

Nonlinear Measurement Systems



[1] Y. Yang, M. Pesavento, Z.-Q. Luo, und B. Ottersten, "Inexact Block Coordinate Descent Algorithms for Nonsmooth Nonconvex Optimization," IEEE Transactions on Signal Processing, 2019.

Exact Block Successive Convex Approximation

$$\underset{\mathbf{P},\mathbf{Q},\mathbf{S}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{P}\mathbf{Q} + \mathbf{D}\mathbf{S} - \mathbf{Y}\|_{\mathsf{F}}^2 + \frac{\lambda}{2} (\|\mathbf{P}\|_{\mathsf{F}}^2 + \|\mathbf{Q}\|_{\mathsf{F}}^2) + \mu \|\mathbf{S}\|_1$$

Inexact Block Successive Convex Approximation

minimize
$$\frac{1}{4} \sum_{n=1}^{N} \left((\mathbf{a}_k^T \mathbf{x})^2 - y_n \right)^2 + \mu \|\mathbf{x}\|_1$$

[2] Y. Yang, M. Pesavento, Y.C. Eldar, B. Ottersten, "Parallel Coordinate Descent Algorithms for Sparse Phase Retrieval," IEEE ICASSP 2019, May 2019.

minimize
$$\underbrace{\frac{1}{2}\sum_{n=1}^{N}\left(|\mathbf{a}_{n}^{H}\mathbf{x}|-y_{n}\right)^{2}}_{\text{loss, }f(\mathbf{x})} + \underbrace{\mu\left\|\mathbf{x}\right\|_{1}}_{\text{regularization, }g(\mathbf{x})}$$

Nonconvex Regularization

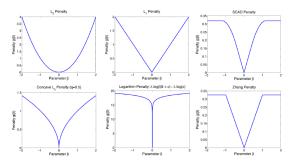


[3] Y. Yang, M. Pesavento, S. Chatzinotas, B. Ottersten, "Successive Convex Approximation Algorithms for Sparse Signal Estimation with Nonconvex Regularizations," IEEE JSTSP, Dec. 2018.

Phase retrieval is a special case of the following general formulation:

(smooth, nonconvex) + (nonsmooth, convex) - (nonsmooth, convex)

Difference-of-convex regularizer to promote sparse/unbiased estimate.



Graphical LASSO with Laplacian Constraints



[4] T. Liu, M.-T. Hoang, Y. Yang and M. Pesavento, "A block Coordinate Descent Algorithm for Sparse Gaussian Graphical Model Interference with Laplacian Constraints," IEEE CAMSAP 2019, Dec. 2019.

$$\begin{split} \min_{\mathbf{X} \succ 0, \mathbf{W}, \gamma} & f(\mathbf{X}) = -\log \det \mathbf{X} + \operatorname{tr}(\mathbf{S}\mathbf{X}) + \rho \|\mathbf{W}\|_{1} \\ \text{s.t.} & \mathbf{X} = \operatorname{diag}(\mathbf{W}\mathbf{1}) - \mathbf{W} + \gamma \mathbf{I} \\ & W_{ii} = 0, i = 1, ..., n \\ & W_{ij} = W_{ji} \geq 0, i = 1, ..., n; j = 1, ..., n \\ & \gamma > 0 \end{split}$$

• $\rho > 0$: Regularization parameter

2

- W: Adjacency matrix
- γ: Positive diagonal loading factor
- **X**: Precision matrix, only an auxiliary variable



TECHNISCHE UNIVERSITÄT

Design of new Sensing Matrices

WP 1: structure of measurement system A

Sensing Matrix Design via Mutual Coherence Minimization (MCM)



Question: how to design a sensing matrix with low mutual coherence?

mutual coherence
$$\mu(\mathbf{A}) = \max_{j \neq k} \frac{|\mathbf{a}_j^H \mathbf{a}_k|}{\|\mathbf{a}_j\|\|\mathbf{a}_k\|}$$

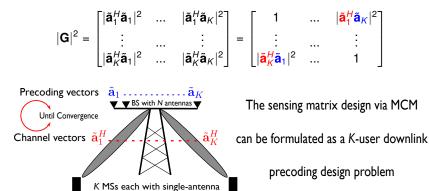
Mutual coherence minimization (MCM) can be formulated as

$$\begin{split} \min_{\mathbf{A} \in \mathbb{C}^{N \times \kappa}} \mu(\mathbf{A}) &= \min_{\substack{\mathbf{A} \in \mathbb{C}^{N \times \kappa} \\ \|\mathbf{a}_j\| = 1, \forall j}} \|\mathbf{A}^H \mathbf{A} - \mathbf{I}_{\kappa}\|_{\infty}^2 \\ &= \min_{\tilde{\mathbf{A}} \in \mathbb{C}^{N \times \kappa}} |\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}|_{\infty,off}^2 \end{split}$$

Proposed MCM Design Formulation



Expanding $|\mathbf{G}|^2 = |\mathbf{\tilde{A}}^H \mathbf{\tilde{A}}|^2$ (squared Gram-matrix), we have



Proposed Sequential MCM Approach (SMCM) [5]



- Step 0: $\tilde{\mathbf{A}}^{(0)} \in \mathbb{C}^{N \times K}$ and $\beta = \max\{\frac{K-N}{N(K-1)}, 0\}$
- Step 2: repeat until a convergence is reached
 - Step 2.1: Solve the *k*th sub-problem, $k = 1, ..., K (\tilde{\mathbf{A}}_{k}^{(n)} = \tilde{\mathbf{a}}_{k}^{(n)} (\tilde{\mathbf{a}}_{k}^{(n)})^{H})$

 $\mathbf{V}_{k}^{\star} = \max_{\mathbf{V}_{k} \in \mathbb{C}^{N \times N} \succeq 0} \operatorname{tr}\{\tilde{\mathbf{A}}_{k}^{(n)} \mathbf{V}_{k}\} \text{ s.t. } \operatorname{tr}\{\tilde{\mathbf{A}}_{j}^{(n)} \mathbf{V}_{k}\} \leq \beta, \forall j \neq k \quad (\operatorname{Precoding update step})$

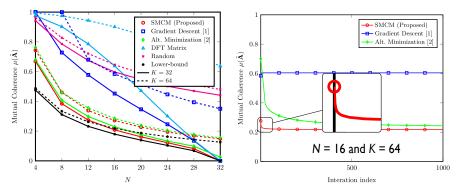
Step 2.2: Calculate EVD: V^{*}_k = **R**_kΣ_k**R**^H_k
Step 2.3: Update kth column of **Ã**⁽ⁿ⁾: **ã**⁽ⁿ⁾_k = [**R**_k]_{σmax} (Channel update step)
Step 2.4: if |μ(**Ã**⁽ⁿ⁾) − μ(**Ã**⁽ⁿ⁻¹⁾)|² ≤ ε, stop.

[5] K. Ardah, M. Pesavento, and M. Haardt, "A novel sensing matrix design for compressed sensing via mutual coherence minimization," IEEE CAMSAP 2019, Dec. 2019

Numerical Results



ILMENAU



[1] V. Abolghasemi, S. Ferdowsi, B. Makkiabadi, and S. Sanei, "On optimization of the measurement matrix for compressive sensing," in Proc. 18th European Signal Processing Conference, Aug. 2010, pp. 427–431.

[2] C. Lu, H. Li, and Z. Lin, "Optimized projections for compressed sensing via direct mutual coherence minimization," Signal Processing, vol. 151, pp. 45 – 55, 2018.



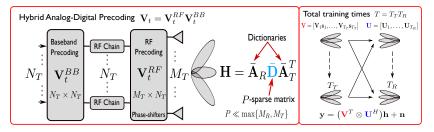
TECHNISCHE UNIVERSITÄT

Multidimensional Training Design

WP 1: structure of measurement system A

Channel Estimation in Hybrid Analog-Digital (HAD)

Channel estimation in HAD massive MIMO systems is a challenging problem
High channel dimension, low SNR before BF, and reduced No. of RF chains



► LS channel estimation: $\mathbf{h}_{LS} = (\mathbf{V}^T \otimes \mathbf{U}^H)^* \mathbf{y} \Rightarrow T_T T_R \ge \frac{M_R M_T}{N_R}$ ► Exploiting the low-rank (sparse) nature of the millimeter-wave channel

$$\mathbf{y} = \left(\mathbf{V}^{\mathsf{T}} \bar{\mathbf{A}}_{\mathsf{T}} \otimes \mathbf{U}^{\mathsf{H}} \bar{\mathbf{A}}_{\mathsf{R}}\right) \bar{\mathbf{d}} + \mathbf{n} = \mathbf{Q} \bar{\mathbf{d}} + \mathbf{n} \in \mathbb{C}^{T_{\mathsf{T}} T_{\mathsf{R}} N_{\mathsf{R}}}$$

▶ $\mu(\mathbf{Q}) = \max\{\mu(\mathbf{V}^T \bar{\mathbf{A}}_T), \mu(\mathbf{U}^H \bar{\mathbf{A}}_R)\} \Rightarrow \text{two independent sensing matrix design steps}$

Open-Loop Training Design for Hybrid Analog-Digital Massive MIMO Systems [6] TECHNISCHE UNIVERSITÄT

 $\mu(\mathbf{Q}) = \max\{\mu(\mathbf{V}^T \bar{\mathbf{A}}_T), \mu(\mathbf{U}^H \bar{\mathbf{A}}_R)\}$ $\mathbf{F}_T \leftarrow \mathsf{SMCM}$ (offline design) $\mathbf{F}_R \leftarrow \mathsf{SMCM}$ (offline design) $\mathbf{F}_{R} = \mathbf{U}^{H} \mathbf{\bar{A}}_{R} \Rightarrow \mathbf{U}_{LS} = (\mathbf{F}_{R} \mathbf{\bar{A}}_{R}^{+})^{H}$ $\mathbf{F}_T = \mathbf{V}^T \bar{\mathbf{A}}_T \Rightarrow \mathbf{V}_{LS} = (\mathbf{F}_T \bar{\mathbf{A}}_T^+)^T$ $\mathbf{V}_{LS} = [\mathbf{v}_1, \dots, \mathbf{v}_t, \dots, \mathbf{v}_{T_T}]$ $\mathbf{U}_{LS} = [\mathbf{U}_1, \dots, \mathbf{U}_r, \dots, \mathbf{U}_{T_R}]$ $\mathbf{v}_t = \mathbf{V}_t \mathbf{s}_t \Rightarrow \mathbf{V}_t = \mathbf{v}_t \mathbf{s}_t^+ = \mathbf{v}_t \mathbf{s}_t^H$ $\mathbf{V}_t = \mathbf{V}_t^{RF} \mathbf{V}_t^{BB}$ $\mathbf{U}_r = \mathbf{U}_r^{RF} \mathbf{U}_r^{BB}$ $\min_{\mathbf{U}^{RF}\mathbf{U}^{BB}} \|\mathbf{U}_r - \mathbf{U}_r^{RF}\mathbf{U}_r^{BB}\|_F^2$ $\min_{\mathbf{v}^{nr},\mathbf{v}^{ns}} \|\mathbf{V}_t - \mathbf{V}_t^{RF} \mathbf{V}_t^{BB}\|_F^2$

[6] K. Ardah, B. Sokal, A. L. F. de Almeida, and M. Haardt, "Compressed sensing based channel estimation And open-loop training design for hybrid analog-digital massive MIMO systems," ICASSP 2020, May 2020 (accepted).

LINIVERSITÄ^{*}

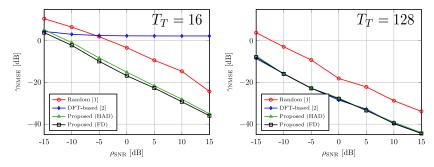
Ch.

ILMENAU

Numerical Results







 J. Zhang, I. Podkurkov, M. Haardt, and A. Nadeev, "Channel estimation and training design for hybrid analog-digital multicarrier single-user massive MIMO systems," in Proc. 20th International ITG Workshop on Smart Antennas (WSA), Mar. 2016

[2] J. Lee, G. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," IEEE Trans. Commun., vol. 64, no. 6, pp. 2370–2386, Jun. 2016.



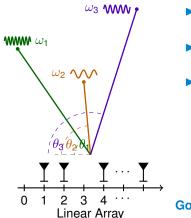
ILMENAU

Identifiability for Sparse Nonuniform Linear Arrays

WP 1: structure of measurement system A

Identifiability for Sparse Nonuniform Linear Arrays Signal Model





- Linear Array: set of *M* sensors located on the *x*-axis at positions r ∈ Z^M.
- ► *N* sources with Direction-Of-Arrivals (DOAs) $\Theta = \{\theta_1, \dots, \theta_N\}.$
- y ∈ ℂ^M is the signal that is received by the *M* sensors: y = A(Θ) x
 - ▶ $\mathbf{x} \in \mathbb{C}^{N}$: emitted signal array,
 - $\mathbf{A}(\Theta) \in \mathbb{C}^{M \times N}$: array steering matrix with columns

$$\mathbf{a}(\theta) = (\mathbf{e}^{-i\pi r_1 \cdot \cos(\theta)}, \dots, \mathbf{e}^{-i\pi r_M \cdot \cos(\theta)})^\top, \ \theta \in \Theta.$$

Goal: Given **y**, (uniquely) estimate $\theta_1, \ldots, \theta_N$.

Consider the linear array $\mathbf{r} = (0, 1, 3, 4)^{\top}$ and the DOAs $\Theta = [0^{\circ}, 60^{\circ}, 90^{\circ}, 120^{\circ}]$.

The array steering matrix $\mathbf{A}(\Theta)$ is rank deficient with rank $(\mathbf{A}(\Theta)) = 3$:

Identifiability for Sparse Nonuniform Linear Arrays The Ambiguity Problem in Subspace-based DOA-Estimation

$$\mathbf{A}(\Theta) = (\mathbf{a}(0^{\circ}), \, \mathbf{a}(60^{\circ}), \, \mathbf{a}(90^{\circ}), \, \mathbf{a}(120^{\circ})) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -i & 1 & i \\ -1 & i & 1 & -i \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Definition. For a linear array $\mathbf{r} \in \mathbb{R}^M$ with M sensors, the (sorted) vector of $n \leq M$ DOAs $\Theta = [\theta_1, ..., \theta_n]$ is ambiguous, if $\rho_a = \operatorname{rank} (A(\Theta)) < n$.

 \Rightarrow Ambiguities are roots of the $M \times M$ minors of the steering matrix $\mathbf{A}(\Theta)$.

Every $M \times M$ minor is a generalized Vandermonde determinant (GVD).

The quotient of a GVD and the classical Vandermonde determinant is given by the Schur polynomial. ⇒ Search roots of the Schur polynomial.



Identifiability for Sparse Nonuniform Linear Arrays Finding Ambiguities



(Semi-standard) Young tableaux yield the following representation of the Schur polynomial, where $\alpha_m^\ell \in \mathbb{Z}$.

$$s_{\lambda}(\Theta) = \sum_{\ell=1}^{n} e^{i\sigma_{\ell}}, \quad \sigma_{\ell} = \sum_{m=1}^{M} -\alpha_{m}^{\ell} \pi \cos(\theta_{m}) \pmod{2\pi} \quad \forall \ell \in [n].$$
(1)

 \Rightarrow Ambiguities \doteq Roots of $s_{\lambda}(\Theta)$.

Let $\mathbf{r} \in \mathbb{Z}^M$ be an integer linear array. In order to find ambiguities:

Enumerate all *n* SSYTs corresponding to $\mathbf{r} \in \mathbb{Z}^{M}$.

Consider e^{iσ_ℓ} to be roots of unity and use vanishing sums of roots of unity to search for roots of (1).

Can be formulated as a mixed-integer program (MIP)!

Lemma. Each solution of the MIP corresponds to an ambiguity of the linear array $\mathbf{r} \in \mathbb{Z}^{M}$.

[7] T. Fischer, F. Matter, M. Pesavento, M. E. Pfetsch, "Ambiguities in DOA Estimation for Linear Arrays," Preprint, unpublished



TECHNISCHE UNIVERSITÄT

General Null Space Properties

WP 2: structure of representation x

WP 3: structure of **sparsity** $\mathbf{x}(t)$



General Null Space Properties – Motivation Uniform Recovery using an NSP

Uniform recovery using a null space property (NSP):

$$\min\{\|x\|_0 : Ax = Ax^{(0)}, x \in \mathbb{R}^n\}, \qquad (P_0)$$

$$\min\{\|x\|_1 : Ax = Ax^{(0)}, x \in \mathbb{R}^n\}.$$
 (P₁)

Characterization when the optimal solution of the nonconvex exact recovery problem (P_0) and its convex relaxation (P_1) coincide.

Known NSPs:

- Sparse (nonnegative) vectors, low-rank (psd) matrices, block-sparse vectors.
- A general framework that subsumes most of the existing NSPs for settings without side constraints [Juditsky et al. '14].

Goal: A general framework for uniform (sparse) recovery under side constraints using NSPs.

General Null Space Properties

A Framework for Sparse Recovery under Side Constraints

- TECHNISCHE UNIVERSITÄT
- ▶ Let \mathcal{X} , \mathcal{E} be finite-dimensional Euclidean spaces and $\mathcal{C} \subseteq \mathcal{X}$ with $0 \in \mathcal{C}$.
- Consider a linear sensing map A: X → ℝ^m, and a linear representation map B: X → E.
- ▶ Define $\mathcal{D} := \{Bx : x \in \mathcal{C}\} \subseteq \mathcal{E}$, consider a norm $\|\cdot\|$ on \mathcal{E} .
- Let P be a set of linear maps on E, with a real weight ν(P) ≥ 0 and a linear map P: E → E for all P ∈ P. Define P_s = {P ∈ P : ν(P) ≤ s}.

For $s \ge 0$, an element $x \in C$ is s-sparse $\Leftrightarrow \exists P \in P$ with $\nu(P) \le s$ and PBx = Bx.

For a given right-hand side $b \in \mathbb{R}$, the generalized recovery problem now reads

$$\min\{\|Bx\|: Ax = b, x \in \mathcal{C}\}.$$

Importantly: By using the set C any side constraint (e.g., nonnegativity, positive semidefiniteness, integrality) can be incorporated into the recovery.

A Generalized Framework for Sparse Recovery under Side Constraints – Main Result



TECHNISCHE UNIVERSITÄT

Theorem

Suppose that some technical assumptions are satisfied. Let *A* be a linear sensing map and $s \ge 1$. Then the following statements are equivalent.

- (i) Every *s*-sparse $x^{(0)} \in C$ is the unique solution of $\min_{x \in C} \{ \|Bx\| : Ax = Ax^{(0)} \}$.
- (ii) For all $v \in (\mathcal{N}(A) \cap (\mathcal{C} + (-\mathcal{C})))$ with $Bv \neq 0$ and all $P \in \mathcal{P}_s$:

 $\overline{P}Bv \in \mathcal{D} \Rightarrow \forall v^{(1)}, v^{(2)} \in \mathcal{C} \text{ with } v = v^{(1)} - v^{(2)} : \|PBv^{(1)}\| - \|PBv^{(2)}\| + \|\overline{P}Bv\| > 0.$

 \Rightarrow General NSP for uniform (sparse) recovery under side constraints.

For example, new NSPs for the following settings can be obtained.

- (Positive semidefinite) Block-diagonal matrices, and
- Nonnegative Block-linear vectors.

[8] J. Heuer, F. Matter, M. E. Pfetsch, T. Theobald, "Block-sparse Recovery of Semidefinite Systems and Generalized Null Space Conditions," July 2019, [arXiv:1907.09442].

Conclusions





We have exploited structure in many ways:

- Design of parallel algorithms,
- System design, i.e., design of sensing matrices with small coherence,
- Ambiguity properties for linear arrays,
- Sparse recovery under side constraints with an NSP.

EXPRESS II Publications I



ILMENAU



Y. Yang, M. Pesavento, Z. Luo, and B. Ottersten,

"Inexact block coordinate descent algorithms for nonsmooth nonconvex optimization," *IEEE Transactions on Signal Processing*, 2019.



Y. Yang, M. Pesavento, Y. C. Eldar, and B. Ottersten,

"Parallel coordinate descent algorithms for sparse phase retrieval,"

in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2019), May 2019.

Y. Yang, M. Pesavento, S. Chatzinotas, and B. Ottersten,

"Successive convex approximation algorithms for sparse signal estimation with nonconvex regularizations,"

IEEE Journal of Selected Topics in Signal Processing, Dec. 2018.

EXPRESS II Publications II



TECHNISCHE UNIVERSITÄT



T. Liu, M.-T. Hoang, Y. Yang, and M. Pesavento,

"A block coordinate descent algorithm for sparse gaussian graphical model interference with laplacian constraints,"

in IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2019), Guadeloupe, West Indies, Dec. 2019.

K. Ardah, M. Pesavento, and M. Haardt,

"A novel sensing matrix design for compressed sensing via mutual coherence minimization,"

in IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2019), Guadeloupe, West Indies, Dec. 2019.



K. Ardah, B. Sokal, A. L. F. de Almeida, and M. Haardt,

"Compressed sensing based channel estimation and open-loop training design for hybrid analog-digital massive MIMO systems,"

in Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP), (Barcelona, Spain), May 2020, Dec. 2019.

EXPRESS II Publications III





Tobias Fischer, Frederic Matter, Marius Pesavento, and Marc E Pfetsch, "Ambiguities in DOA estimation for linear arrays," Tech. Rep., TU Darmstadt, 2020, unpublished.

Janin Heuer, Frederic Matter, Marc E Pfetsch, and Thorsten Theobald, "Block-sparse recovery of semidefinite systems and generalized null space conditions,"

Preprint arXiv:1907.09442, July 2019.