TECHNISCHE UNIVERSITÄT DARMSTADT

A PARALLEL SUCCESSIVE CONVEX APPROXIMATION FRAMEWORK WITH SMOOTHING MAJORIZATION FOR PHASE RETRIEVAL

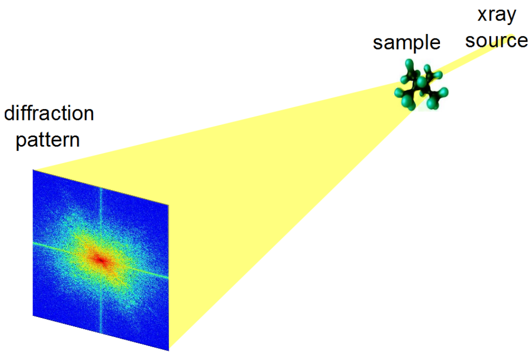
Tianyi Liu

September 26, 2024

INTRODUCTION

Motivation

Diffraction Imaging



Source: [Candes et al., 2015][1]

[1]Emmanuel J. Candès et al. “Phase retrieval via matrix completion”. In: *SIAM review* 57.2 (2015), pp. 225–251.

INTRODUCTION

Contributions

Phase Retrieval

- Challenge: Nonsmooth and Nonconvex Optimization

Proposed Algorithms for PR

- Compact Formulation
- Efficient Implementation
- Analysis of Tuning Parameters

Smoothing Successive Convex Approximation

- Smoothing Majorization
- Smoothing Successive Convex Approximation
- Convergence Analysis

TECHNISCHE
UNIVERSITÄT
DARMSTADT

OUTLINE

TECHNISCHE UNIVERSITÄT DARMSTADT

SIGNAL MODEL

Phase Retrieval

- In classic phase retrieval problem, estimate $\boldsymbol{X}$ from magnitude of linear measurements:

$$\boldsymbol{Y} = |\mathcal{F}(\boldsymbol{X})| + \boldsymbol{N}$$

  - $\boldsymbol{X} \in \mathbb{C}^{N \times I}$ is the signal to estimate
  - $\mathcal{F} : \mathbb{C}^{N \times I} \to \mathbb{C}^{M_1 \times M_2}$ is the linear measurement operator
  - $\boldsymbol{N} \in \mathbb{R}^{M_1 \times M_2}$ is a matrix of i.i.d. noise entries
  - $\boldsymbol{Y} \in \mathbb{R}_+^{M_1 \times M_2}$ is the magnitude-only measurements
  - Absolute value operation performed elementwise

TECHNISCHE UNIVERSITÄT DARMSTADT

SIGNAL MODEL

Phase Retrieval with Dictionary Learning

- Additional prior information of signal $\boldsymbol{X}$, e.g., sparsity
- Original signal $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_I]$ admits a sparse representation:

$$\boldsymbol{x}_i = \boldsymbol{D}\boldsymbol{z}_i$$

  - $\boldsymbol{D} = [\boldsymbol{d}_1, \ldots, \boldsymbol{d}_P] \in \mathbb{C}^{N \times P}$ is an unknown dictionary
  - $\boldsymbol{Z} = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_I] \in \mathbb{C}^{P \times I}$ is a sparse matrix of codes

Problem statement

Given magnitude-only measurements $\boldsymbol{Y}$ and linear operator $\mathcal{F}$, jointly learn an unknown dictionary $\boldsymbol{D}$ and a sparse code matrix $\boldsymbol{Z}$ such that $\boldsymbol{Y} \approx |\mathcal{F}(\boldsymbol{D}\boldsymbol{Z})|$

TECHNISCHE UNIVERSITÄT DARMSTADT

PROBLEM FORMULATION

- Proposed compact formulation:

$$\min_{\boldsymbol{D}\in\mathcal{D},\boldsymbol{Z}} \quad \underbrace{\frac{1}{2}\|\boldsymbol{Y} - |\mathcal{F}(\boldsymbol{DZ})|\|_{\mathsf{F}}^2}_{\text{loss}} + \underbrace{\lambda\|\boldsymbol{Z}\|_{1,1}}_{\text{sparsity}}$$

  - $\|\boldsymbol{Z}\|_{1,1} = \sum_{p=1}^{P}\sum_{i=1}^{I}|z_{pi}|$
  - $\lambda > 0$: sparsity regularization parameter
  - To avoid scaling ambiguities, restrict $\boldsymbol{D} \in \mathcal{D} = \{\boldsymbol{D} \mid \|\boldsymbol{d}_p\|_2 \leq 1, p = 1, \ldots, P\}$
  - Loss function is nonsmooth and nonconvex

PROBLEM FORMULATION

- Proposed compact formulation:

$$\min_{\boldsymbol{D}\in\mathcal{D},\boldsymbol{Z}} \quad \underbrace{\frac{1}{2}\|\boldsymbol{Y} - |\mathcal{F}(\boldsymbol{DZ})|\|_\mathsf{F}^2}_{\text{loss}} + \underbrace{\lambda\|\boldsymbol{Z}\|_{1,1}}_{\text{sparsity}}$$

- Alternative formulation [QP17][2]:

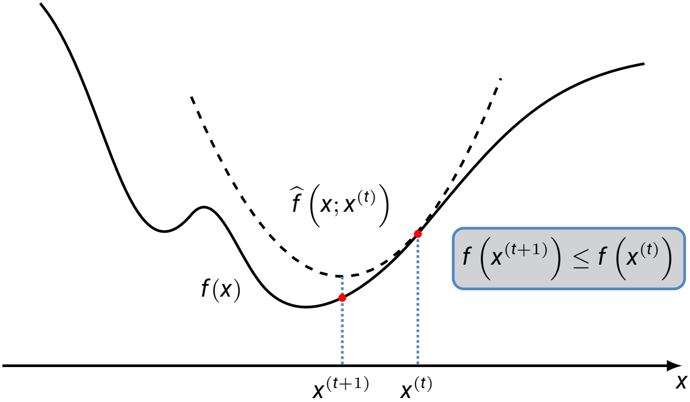
$$\min_{\boldsymbol{X},\boldsymbol{D}\in\mathcal{D},\boldsymbol{Z}} \quad \underbrace{\frac{1}{2}\|\boldsymbol{Y} - |\mathcal{F}(\boldsymbol{X})|\|_\mathsf{F}^2}_{\text{data fitting}} + \underbrace{\frac{\mu}{2}\|\boldsymbol{X} - \boldsymbol{DZ}\|_\mathsf{F}^2}_{\text{approximation quality}} + \underbrace{\rho\|\boldsymbol{Z}\|_{1,1}}_{\text{sparsity}}$$

  - Nonconvex and nonsmooth
  - Additional auxiliary variable $\boldsymbol{X}$ and two regularization parameters

[2] Tianyu Qiu and Daniel P. Palomar. "Undersampled sparse phase retrieval via majorization–minimization". In: *IEEE Transactions on Signal Processing* 65.22 (Nov. 2017), pp. 5957–5969. ISSN: 1053-587X.

OUTLINE

MAJORIZATION-MINIMIZATION



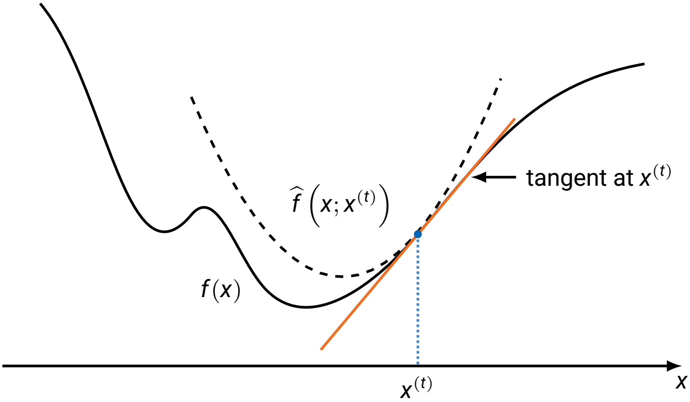
Majorization

- Tangent to the original function at $x^{(t)}$
- Global upper bound

TECHNISCHE UNIVERSITÄT DARMSTADT

MAJORIZATION-MINIMIZATION

Existing Convergence Result



Differential Consistency

$$\widehat{f}'_{\boldsymbol{d}}(\boldsymbol{x}^{(t)}; \boldsymbol{x}^{(t)}) = f'_{\boldsymbol{d}}(\boldsymbol{x}^{(t)}) \; \forall \boldsymbol{d}$$

$f'_{\boldsymbol{d}}(\boldsymbol{x})$: directional derivative of f at $\boldsymbol{x}$ in direction $\boldsymbol{d}$

$$\Downarrow$$

$\{\boldsymbol{x}^{(t)}\}$ converges to a stationary point of f that satisfies

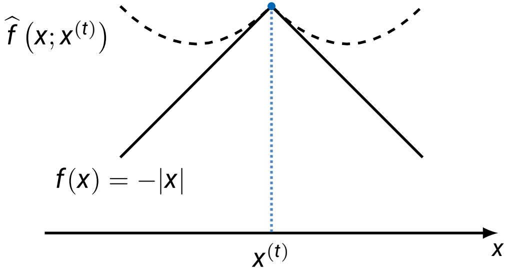
$$f'_{\boldsymbol{d}}(\boldsymbol{x}) \geq 0 \; \forall \boldsymbol{d}$$

- The stationarity is only *necessary* for local minimum

TECHNISCHE UNIVERSITÄT DARMSTADT

MAJORIZATION-MINIMIZATION

Existing Convergence Result



Differential Consistency

$$\widehat{f}'_{\boldsymbol{d}}(\boldsymbol{x}^{(t)}; \boldsymbol{x}^{(t)}) = f'_{\boldsymbol{d}}(\boldsymbol{x}^{(t)})\ \forall \boldsymbol{d}$$

$f'_{\boldsymbol{d}}(\boldsymbol{x})$: directional derivative of f at $\boldsymbol{x}$ in direction $\boldsymbol{d}$

- Limitation: Differential consistency restricts the majorizer at a nondifferentiable point to be nondifferentiable

SMOOTHING MAJORIZATION

Restriction

Smoothness of Majorizer

- Lower computational complexity
- Possibility of approximately minimizing the majorizer with SCA

Relaxation

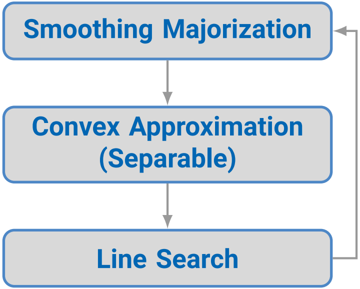
Almost "harmless" relaxation of the convergence set

– Convergence to a stationary point in a relaxed sense

+ Many nondifferentiable stationary points that are not local minima are excluded from the convergence set

TECHNISCHE UNIVERSITÄT DARMSTADT

SMOOTHING SUCCESSIVE CONVEX APPROXIMATION



- The convergence of SCA requires the smoothness of the original function
- Extend SCA to nonsmooth problems with smoothing majorization to enjoy the advantages:
  - Tighter approximation
  - Separable approximation

TECHNISCHE UNIVERSITÄT DARMSTADT

OUTLINE

TECHNISCHE UNIVERSITÄT DARMSTADT

PROPOSED ALGORITHMS FOR PHASE RETRIEVAL

- Proposed compact formulation:

$$\min_{\boldsymbol{D}\in\mathcal{D},\boldsymbol{Z}} \quad \underbrace{\frac{1}{2}\|\boldsymbol{Y}-|\mathcal{F}(\boldsymbol{D}\boldsymbol{Z})|\|_{\mathsf{F}}^2}_{\text{loss}} + \underbrace{\lambda\|\boldsymbol{Z}\|_{1,1}}_{\text{sparsity}} \qquad \text{compact-SCAphase}$$

  - Fast convergence in terms of number of iterations
- Alternative formulation [QP17][3]:

$$\min_{\boldsymbol{X},\boldsymbol{D}\in\mathcal{D},\boldsymbol{Z}} \quad \frac{1}{2}\|\boldsymbol{Y}-|\mathcal{F}(\boldsymbol{X})|\|_{\mathsf{F}}^2 + \frac{\mu}{2}\|\boldsymbol{X}-\boldsymbol{D}\boldsymbol{Z}\|_{\mathsf{F}}^2 + \rho\|\boldsymbol{Z}\|_{1,1} \qquad \text{SCAphase}$$

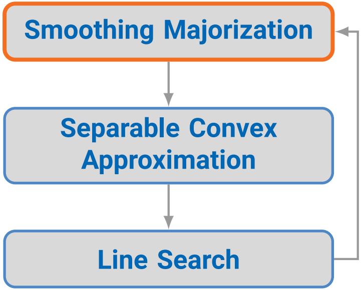
  - Lower per-iteration complexity

[3] Tianyu Qiu and Daniel P. Palomar. "Undersampled sparse phase retrieval via majorization–minimization". In: *IEEE Transactions on Signal Processing* 65.22 (Nov. 2017), pp. 5957–5969. ISSN: 1053-587X.

TECHNISCHE UNIVERSITÄT DARMSTADT

COMPACT-SCAPHASE

Smoothing Majorization



Smoothing majorization at $(\boldsymbol{D}^{(t)}, \boldsymbol{Z}^{(t)})$ in iteration t:

$$\frac{1}{2}\|\boldsymbol{Y} - |\mathcal{F}(\boldsymbol{DZ})|\|_\mathsf{F}^2 \leq \frac{1}{2}\left\|\boldsymbol{Y}^{(t)} - \mathcal{F}(\boldsymbol{DZ})\right\|_\mathsf{F}^2$$

- $\boldsymbol{Y}^{(t)}$: Magnitude measurements $\boldsymbol{Y}$ with phases of $\mathcal{F}(\boldsymbol{D}^{(t)}\boldsymbol{Z}^{(t)})$
- Challenge: Bilinear term $\boldsymbol{DZ}$
  - All elements of $\boldsymbol{D}$ and $\boldsymbol{Z}$ are coupled
  - Smooth but nonconvex

TECHNISCHE UNIVERSITÄT DARMSTADT

COMPACT-SCAPHASE

Separable Convex Approximation

- Exact minimization of the majorizing function:

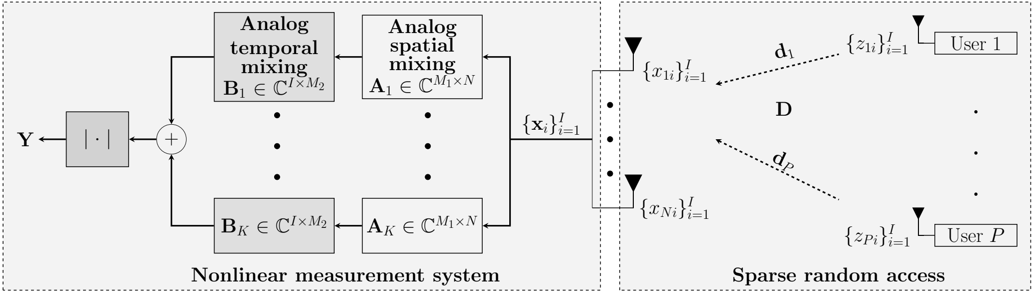
$$\min_{\boldsymbol{D}\in\mathcal{D},\boldsymbol{Z}} \quad \frac{1}{2}\|\boldsymbol{Y}^{(t)} - \mathcal{F}(\boldsymbol{D}\boldsymbol{Z})\|_{\mathsf{F}}^2 + \lambda\|\boldsymbol{Z}\|_{1,1}$$

- Block Coordinate Descent (BCD): Alternately minimize the function in different blocks of variables
  - The bilinear term becomes linear if one matrix is fixed
- Jacobi-type Approximation: Jointly consider all the convex partial minimizations in one iteration

OUTLINE

SIMULATION RESULTS

Multi-antenna Random Access Network



- Linear operator $\mathcal{F}(\boldsymbol{X}) = \sum_{i=1}^{K} \boldsymbol{A}_i \boldsymbol{X} \boldsymbol{B}_i$ with $\boldsymbol{A}_i \in \mathbb{C}^{M_1\times N}$ and $\boldsymbol{B}_i \in \mathbb{C}^{I\times M_2}$
- L users are active in each time-slot

TECHNISCHE UNIVERSITÄT DARMSTADT

SIMULATION RESULTS

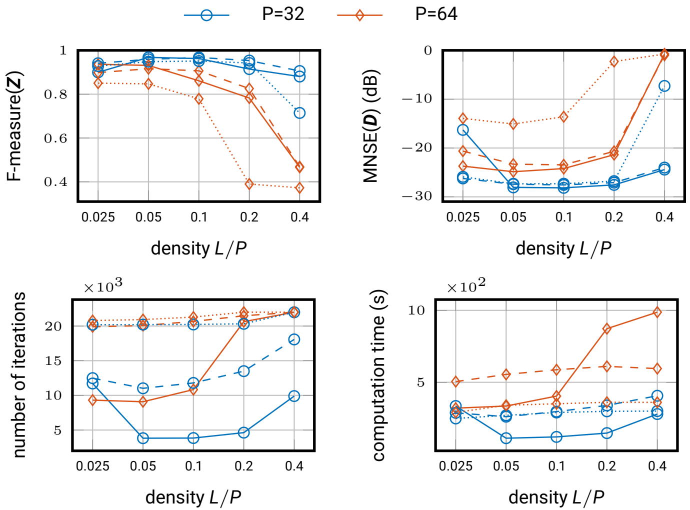
Simulation Setup

- Number of receive antennas $N = 64$
- Number of time-slots $I = 16N = 1024$
- Number of users $P \in \{N/2, N\}$
- *Time-invariant Gaussian spatial mixing* $\boldsymbol{A}$ with oversampling rate $M_1/N = 4$
- Ground-truth $\boldsymbol{D}$ and $\boldsymbol{Z}$ i.i.d. $\sim \mathcal{CN}(0, 1)$
- Gaussian noise with SNR $= 15$ dB
- 50 Monte-Carlo trials
- *Maximum number of iterations: 2000*

TECHNISCHE UNIVERSITÄT DARMSTADT

SIMULATION RESULTS

Influence of Sparsity Level

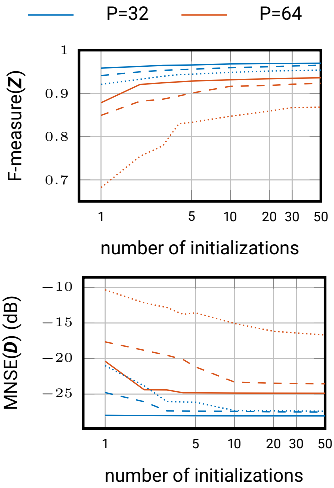


- 10 random initializations
- compact-SCAphase (solid)
  - Fast convergence in terms of no. of iterations
- SCAphase (dashed)
  - Low per-iteration complexity
- Benchmark: SC-PRIME [QP17] (dotted)
  - Block-coordinate-wise majorization

TECHNISCHE UNIVERSITÄT DARMSTADT

SIMULATION RESULTS

Robustness to Initialization



- density $L/P = 0.05$
- compact-SCAphase (solid)
    - Most robust to initialization
- SCAphase (dashed)
- Benchmark: SC-PRIME (dotted)

TECHNISCHE UNIVERSITÄT DARMSTADT

OUTLINE

TECHNISCHE UNIVERSITÄT DARMSTADT

SUMMARY OF CONTRIBUTIONS

On the optimization framework aspect

- *Smoothing Majorization:* Construct a smooth majorizer for a nonsmooth function
- *Smoothing SCA:* Extension of SCA to nonsmooth problems with smoothing majorization
- Block coordinate descent version
- Justification of the convergence to a stationary point in a relaxed sense (not published yet)

On the application aspect [Liu+22][4]

- Proposal of a compact formulation for PRDL with reduced variable size and number of regularization parameters
- Implementation of Smoothing SCA for PRDL: Efficient solution approach for subproblems
- Effective ranges for the regularization parameters

[4] Tianyi Liu et al. "Extended successive convex approximation for phase retrieval with dictionary learning". In: *IEEE Transactions on Signal Processing* 70 (2022), pp. 6300–6315. ISSN: 1941-0476.

TECHNISCHE UNIVERSITÄT DARMSTADT

PUBLICATIONS

Relevant to Dissertation

Journal Article

- **T. Liu**, A. M. Tillmann, Y. Yang, Y. C. Eldar, and M. Pesavento, “Extended successive convex approximation for phase retrieval with dictionary learning,” *IEEE Transactions on Signal Processing*, vol. 70, pp. 6300–6315, 2022.

Book Section

- K. Ardah, M. Haardt, **T. Liu**, F. Matter, M. Pesavento, and M. E. Pfetsch, “Recovery under side constraints,” in *Compressed sensing in information processing*, G. Kutyniok, H. Rauhut, and R. J. Kunsch, Eds., Cham: Springer International Publishing, 2022, pp. 213–246.

Conference Paper

- **T. Liu**, A. M. Tillmann, Y. Yang, Y. C. Eldar, and M. Pesavento, “A parallel algorithm for phase retrieval with dictionary learning,” in *Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Jun. 2021.

TECHNISCHE UNIVERSITÄT DARMSTADT

PUBLICATIONS

Others

Preprints

- **T. Liu**, S. P. Deram, K. Ardah, M. Haardt, M. E. Pfetsch, and M. Pesavento, "Gridless Parameter Estimation in Partly Calibrated Rectangular Arrays." submitted to the *IEEE Transactions on Signal Processing*, 2024. URL: `http://arxiv.org/abs/2406.16041`.
- **T. Liu**, F. Matter, A. Sorg, M. E. Pfetsch, M. Haardt, and M. Pesavento, "Joint sparse estimation with cardinality constraint via mixed-integer semidefinite programming." submitted to the *IEEE Transactions on Signal Processing*, 2023. URL: `http://arxiv.org/abs/2311.03501`.

TECHNISCHE UNIVERSITÄT DARMSTADT

PUBLICATIONS

Others

Conference Papers

- **T. Liu** and M. Pesavento, "Blind Phase-Offset Estimation in Sparse Partly Calibrated Arrays," in *Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jul. 2024. selected for the best student paper contest.
- **T. Liu**, S. P. Deram, K. Ardah, M. Haardt, M. E. Pfetsch, and M. Pesavento, "Gridless parameter estimation in partly calibrated rectangular arrays," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2024.
- **T. Liu**, F. Matter, A. Sorg, M. E. Pfetsch, M. Haardt, and M. Pesavento, "Joint sparse estimation with cardinality constraint via mixed-integer semidefinite programming," in *Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2023.
- Y. Zhang, **T. Liu**, and M. Pesavento, "Direction-of-arrival estimation for correlated sources and low sample size," in *Proceedings of European Signal Processing Conference (EUSIPCO)*, Sep. 2023.

TECHNISCHE UNIVERSITÄT DARMSTADT

PUBLICATIONS

Others

Conference Papers

- **T. Liu**, M. Trinh-Hoang, Y. Yang, and M. Pesavento, "A block coordinate descent algorithm for sparse Gaussian graphical model inference with laplacian constraints," in *Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2019.
- X. Wang, **T. Liu**, M. Trinh-Hoang, and M. Pesavento, "GPU-accelerated parallel optimization for sparse regularization," in *Proceedings of Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jun. 2020.
- **T. Liu**, M. Trinh-Hoang, Y. Yang, and M. Pesavento, "A parallel optimization approach on the infinity norm minimization problem," in *Proceedings of European Signal Processing Conference (EUSIPCO)*, Sep. 2019. selected for the best student paper contest.

TECHNISCHE UNIVERSITÄT DARMSTADT

Thank you for your attention!

REFERENCES I

[Can+15] Emmanuel J. Candès et al. “Phase retrieval via matrix completion”. In: *SIAM review* 57.2 (2015), pp. 225–251.

[Liu+22] Tianyi Liu et al. “Extended successive convex approximation for phase retrieval with dictionary learning”. In: *IEEE Transactions on Signal Processing* 70 (2022), pp. 6300–6315. ISSN: 1941-0476.

[QP17] Tianyu Qiu and Daniel P. Palomar. “Undersampled sparse phase retrieval via majorization–minimization”. In: *IEEE Transactions on Signal Processing* 65.22 (Nov. 2017), pp. 5957–5969. ISSN: 1053-587X.