Robust Recovery of Jointly-Sparse Signals Using Minimax Concave Loss Function

Kyohei Suzuki*, Masahiro Yukawa*

*Department of Electronics and Electrical Engineering, Keio University, Japan

International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2022)

May 9, 2022

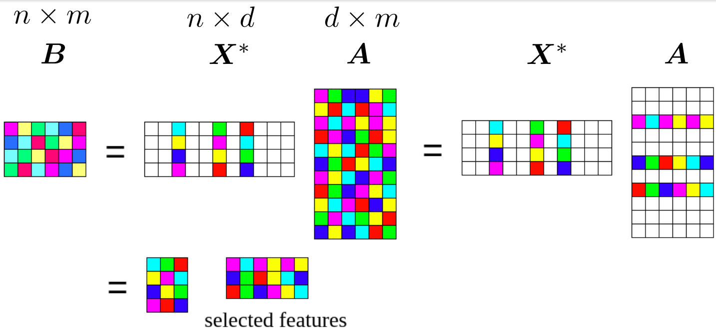
Outline

Outline

1. Background
2. Main Results
3. Numerical Examples
   - Experiment A: Robustness in Overdetermined Case
   - Experiment B: Support Recovery in Underdetermined Case
   - Experiment C: MECG Signal Recovery
4. Conclusion

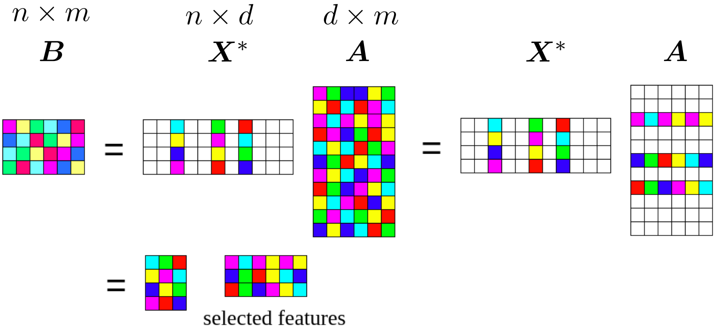
Feature Selection Problem

Select an important subset of features
→ enhancing the performances (classification accuracy, training time, *etc*)



Feature Selection Problem

Select an important subset of features
→ enhancing the performances (classification accuracy, training time, *etc*)



Without the condition $d > m$, this can be applied to multiple measurement vector (MMV) problem.

Applications: MECG, DNA microarrays, source localization, *etc*

Mathematical Model and Robust Feature Selection

$$\boldsymbol{B} = \boldsymbol{X}^* \boldsymbol{A} + \underset{\text{noise} \,\lrcorner\!\!\uparrow}{\boldsymbol{E}} + \underset{\uparrow\!\!\llcorner \, \text{outlier matrix (column sparse)}}{\boldsymbol{O}} \tag{1}$$

Mathematical Model and Robust Feature Selection

$$\boldsymbol{B} = \boldsymbol{X}^* \boldsymbol{A} + \underset{\text{noise}\,\uparrow}{\boldsymbol{E}} + \underset{\uparrow\ \text{outlier matrix (column sparse)}}{\boldsymbol{O}} \tag{1}$$

Robust Feature Selection (RFS, Nie *et al.*, '10 [1]):

$$(\mathrm{P}_0) \qquad \min_{\boldsymbol{X} \in \mathcal{X}} \underbrace{\|\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}\|_{2,1}}_{\uparrow\ \boxed{\text{outlier robustness}}} + \underbrace{\lambda \|\boldsymbol{X}\|_{2,1}}_{\uparrow\ \boxed{\text{column sparsity}}}$$

$\|\boldsymbol{X}\|_{2,1} := \sum_{i=1}^{d} \|\boldsymbol{x}_i\|_2$ (sum of the column norms)

---

[1] F. Nie, H. Huang, X. Cai, and C. H. Ding, "Efficient and robust feature selection via joint $\ell_{2,1}$-norms minimization," in Proc. Adv. Neural Inf. Process. Syst., 2010, pp. 1813–1821.

Minimax Concave Function

Huber function: $\phi_\gamma^{\mathrm{HB}}(x) := \begin{cases} \frac{1}{2\gamma}x^2 & \text{if } |x| \leq \gamma \\ |x| - \frac{1}{2}\gamma & \text{if } |x| > \gamma \end{cases} \quad (\gamma > 0)$

MC function (Zhang 2010 [2], Selesnick 2017 [3])

$$\phi_\gamma^{\mathrm{MC}}(x) \coloneqq |x| - \phi_\gamma^{\mathrm{HB}}(x) = \begin{cases} |x| - \frac{1}{2\gamma}x^2 & \text{if } |x| \leq \gamma \\ \frac{1}{2}\gamma & \text{if } |x| > \gamma \end{cases} \quad (\gamma > 0)$$

Constant over $[\gamma, +\infty)$ → remarkable robustness against outliers

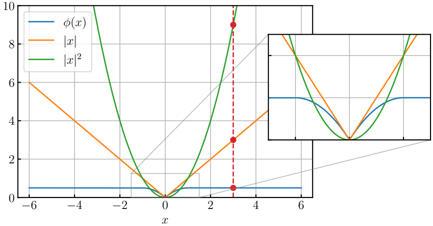
Minimax Concave Function

Huber function: $\phi_\gamma^{\mathrm{HB}}(x) := \begin{cases} \frac{1}{2\gamma}x^2 & \text{if } |x| \leq \gamma \\ |x| - \frac{1}{2}\gamma & \text{if } |x| > \gamma \end{cases} \quad (\gamma > 0)$

MC function (Zhang 2010 [2], Selesnick 2017 [3])

$$\phi_\gamma^{\mathrm{MC}}(x) := |x| - \phi_\gamma^{\mathrm{HB}}(x) = \begin{cases} |x| - \frac{1}{2\gamma}x^2 & \text{if } |x| \leq \gamma \\ \frac{1}{2}\gamma & \text{if } |x| > \gamma \end{cases} \quad (\gamma > 0)$$

Constant over $[\gamma, +\infty)$ → remarkable robustness against outliers



[2] C. H. Zhang, “Nearly unbiased variable selection under minimax concave penalty,” The Annals of Statistics, vol. 38, no. 2, pp. 894—942, 2010.
[3] I. Selesnick, “Sparse regularization via convex analysis,” IEEE Transactions on Signal Processing, vol. 65, no. 17, pp. 4481–4494, 2017.

Key Ideas

	robustness	convexity of loss	mathematical tractability (overall convexity)
Huber [4]	△	✓	✓
Tukey [5]	✓	×	×
Proposed	✓	(weakly convex)	✓

Key Ideas

	robustness	convexity of loss	mathematical tractability (overall convexity)
Huber [4]	△	✓	✓
Tukey [5]	✓	×	×
Proposed	✓	(weakly convex)	✓

Ideas

1. MC loss
   → **Remarkable Outlier robustness** due to the nonconvexity
2. The squared Frobenius norm as an additional penalty
   → **Global optimality**
3. Split the cost function into convex terms
   → The reformulated problem can be solved by the efficient primal-dual splitting method
   → **Scalability**

---

[4] P. J. Huber and E. Ronchetti, Robust Statistics, JohnWiley & Sons, 2009.
[5] R. A. Maronna, et al., Robust Statistics: Theory and Methods (with R), John Wiley & Sons, 2019.

Outline

1. Background
2. Main Results
3. Numerical Examples
   - Experiment A: Robustness in Overdetermined Case
   - Experiment B: Support Recovery in Underdetermined Case
   - Experiment C: MECG Signal Recovery
4. Conclusion

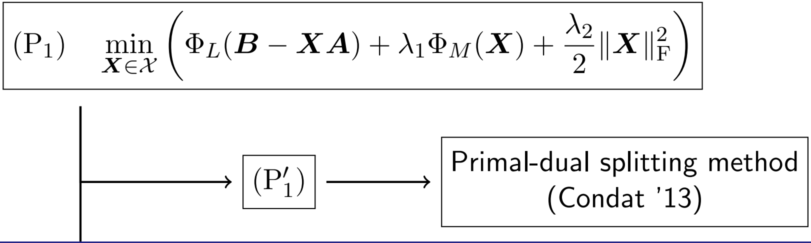
Proposed Formulation

$$(\mathrm{P}_1) \quad \min_{\boldsymbol{X}\in\mathcal{X}} \left(\Phi_L(\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}) + \lambda_1 \Phi_M(\boldsymbol{X}) + \frac{\lambda_2}{2}\|\boldsymbol{X}\|_{\mathrm{F}}^2 \right)$$

- $\Phi_L(\boldsymbol{Y}) := \|\boldsymbol{Y}\|_{2,1} - \min_{\boldsymbol{Z}\in\mathcal{Y}} \left(\|\boldsymbol{Z}\|_{2,1} + \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{Z}\|_L^2\right),\ \boldsymbol{Y} \in \mathcal{Y}$ $(\|\boldsymbol{Y}\|_L := \|L^{1/2}\boldsymbol{Y}\|_{\mathrm{F}})$
- $\Phi_M(\boldsymbol{X}) := \|\boldsymbol{X}\|_{2,1} - \min_{\boldsymbol{\Xi}\in\mathcal{X}} \left(\|\boldsymbol{\Xi}\|_{2,1} + \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{\Xi}\|_M^2\right),\ \boldsymbol{X} \in \mathcal{X}$
- $L\boldsymbol{B} := \boldsymbol{B}\mathrm{diag}(l_1, \ldots, l_m),\ l_i > 0,\ \forall i = 1, \ldots, m$
- $M\boldsymbol{X} := \boldsymbol{X}\mathrm{diag}(\mu_1, \ldots, \mu_n),\ \mu_j > 0,\ \forall j = 1, \ldots, n$
- $\lambda_1 \geq 0$
- $\lambda_2 \geq 0$

Φ_L and Φ_M are slight extensions of the MC function to group-sparse matrices

Flow of the Derivation



$$\boxed{(\mathrm{P}_1) \quad \min_{\boldsymbol{X}\in\mathcal{X}} \left(\Phi_L(\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}) + \lambda_1 \Phi_M(\boldsymbol{X}) + \frac{\lambda_2}{2}\|\boldsymbol{X}\|_{\mathrm{F}}^2 \right)}$$

$$\longrightarrow \boxed{(\mathrm{P}_1')} \longrightarrow$$

Primal-dual splitting method
(Condat '13)

Moreau's decomposition

Let $f \in \Gamma_0(\mathcal{X})$. Then,

$$ {}^{1}f + {}^{1}f^{*} = \frac{1}{2}\|\cdot\|_{\mathrm{F}}^2, \tag{2}$$

where the Moreau envelope of f of index γ is

$$ {}^{\gamma}f : \boldsymbol{X} \mapsto \min_{\boldsymbol{Y}\in\mathcal{Y}} \left\{ f(\boldsymbol{Y}) + \frac{1}{2\gamma}\|\boldsymbol{X} - \boldsymbol{Y}\|_{\mathrm{F}}^2 \right\}, \tag{3}$$

and f^{*} is the convex conjugate of f.

Reformulation Based on Moreau's Decomposition

$$(\mathrm{P}_1) \quad \min_{\boldsymbol{X} \in \mathcal{X}} \left(\Phi_L(\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}) + \lambda_1 \Phi_M(\boldsymbol{X}) + \frac{\lambda_2}{2} \|\boldsymbol{X}\|_{\mathrm{F}}^2 \right)$$

Reformulation

$$(\mathrm{P}_1) \Leftrightarrow (\mathrm{P}_1') \;\; \min_{\boldsymbol{X} \in \mathcal{X}} [F(\boldsymbol{X}) + G(\boldsymbol{X}) + H(L_1\boldsymbol{X})] \tag{4}$$

- $F(\boldsymbol{X}) := \frac{\lambda_2}{2}\|\boldsymbol{X}\|_{\mathrm{F}}^2 - \frac{1}{2}\|L_1\boldsymbol{X}\|_L^2 - \frac{\lambda_1}{2}\|\boldsymbol{X}\|_M^2 + \langle L_1\boldsymbol{X}, \boldsymbol{B}\rangle_L$
  $+ {}^{1}(\iota_C \circ L^{1/2})(L^{1/2}(\boldsymbol{B} - L_1\boldsymbol{X})) + \lambda_1 {}^{1}(\iota_C \circ M^{1/2})(M^{1/2}\boldsymbol{X})$

  (a certain condition → convexity)
- $G(\boldsymbol{X}) := \lambda_1 \|\boldsymbol{X}\|_{2,1}$ (automatically convex)
- $H(\boldsymbol{Y}) := \|\boldsymbol{B} - \boldsymbol{Y}\|_{2,1}$ (automatically convex)
- $L_1\boldsymbol{X} := \boldsymbol{X}\boldsymbol{A}$

Convexity Results

$$(\mathrm{P}_1') \quad \min_{\boldsymbol{X}\in\mathcal{X}} [F(\boldsymbol{X}) + G(\boldsymbol{X}) + H(L_1\boldsymbol{X})] \tag{5}$$

$$\begin{aligned} F(\boldsymbol{X}) := {} & \frac{\lambda_2}{2}\|\boldsymbol{X}\|_{\mathrm{F}}^2 - \frac{1}{2}\|L_1\boldsymbol{X}\|_L^2 - \frac{\lambda_1}{2}\|\boldsymbol{X}\|_M^2 + \langle L_1\boldsymbol{X}, \boldsymbol{B}\rangle_L \\ & + {}^{1}(\imath_C \circ L^{1/2})(L^{1/2}(\boldsymbol{B} - L_1\boldsymbol{X})) + \lambda_1 {}^{1}(\imath_C \circ M^{1/2})(M^{1/2}\boldsymbol{X}) \end{aligned} \tag{6}$$

Convexity Results

$$(\mathrm{P}_1') \quad \min_{\boldsymbol{X} \in \mathcal{X}} [F(\boldsymbol{X}) + G(\boldsymbol{X}) + H(L_1 \boldsymbol{X})] \tag{5}$$

$$F(\boldsymbol{X}) := \frac{\lambda_2}{2}\|\boldsymbol{X}\|_{\mathrm{F}}^2 - \frac{1}{2}\|L_1\boldsymbol{X}\|_L^2 - \frac{\lambda_1}{2}\|\boldsymbol{X}\|_M^2 + \langle L_1\boldsymbol{X}, \boldsymbol{B}\rangle_L$$
$$+ {}^{1}(\iota_C \circ L^{1/2})(L^{1/2}(\boldsymbol{B} - L_1\boldsymbol{X})) + \lambda_1 {}^{1}(\iota_C \circ M^{1/2})(M^{1/2}\boldsymbol{X}) \tag{6}$$

Proposition 1

1. The function F is convex ($\rightarrow$ **Global optimality**) if

$$\lambda_2 \geq \lambda_{\max}\{\boldsymbol{A}\mathrm{diag}(l_1, \ldots, l_m)\boldsymbol{A}^{\mathsf{T}} + \lambda_1 \mathrm{diag}(\mu_1, \ldots, \mu_n)\}. \tag{7}$$

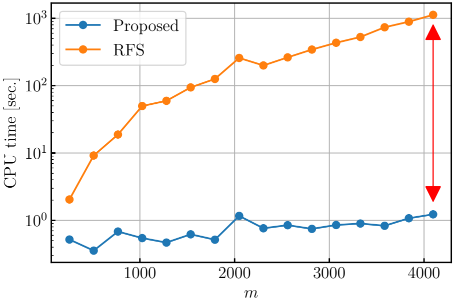
2. The condition (7) is also necessary when
$$K := \{\boldsymbol{X} \in \mathcal{X} \mid \|L(\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A})\|_{2,\infty} \leq 1, \|M\boldsymbol{X}\|_{2,\infty} \leq 1\}$$
has a nonempty interior. ($\|\boldsymbol{X}\|_{2,\infty}$: maximal column norm)

Note: $K := \{\boldsymbol{X} \in \mathcal{X} \mid \left(\iota_C \circ L^{1/2} \boxdot q\right)(L^{1/2}(\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A})) = 0\} \cap \{\boldsymbol{X} \in \mathcal{X} \mid \left(\iota_C \circ M^{1/2} \boxdot q\right)(M^{1/2}\boldsymbol{X}) = 0\}$ in the original paper.

Scalability of Proposed Method

Table: Computational complexity (q: the number of iterations)

Algorithm	Computational complexity
Proposed	$\mathcal{O}(\max\{qnmd, \min\{d^2, m^2\} \max\{d, m, p_\beta, p_\sigma\}\})$
RFS	$\mathcal{O}(qm(m+d) \max\{m+d, n\})$



18 min.

1 sec.

Scalability

Difficulties in Applying PDS or ADMM

$$(\mathrm{P}_1) \quad \min_{\boldsymbol{X}\in\mathcal{X}} \left(\Phi_L(\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}) + \lambda_1 \Phi_M(\boldsymbol{X}) + \frac{\lambda_2}{2}\|\boldsymbol{X}\|_{\mathrm{F}}^2 \right)$$

- Applying the primal-dual splitting method directly:
  - Each term should be convex
  - Proximity operator of the conjugate function of the first term is the zero operator
- Applying ADMM directly:
  - A certain proximity operator needs to be firmly nonexpansive, but... it is not even nonexpansive in the present case

Please refer to Section III-B of the following paper for detailed discussions:
K. Suzuki, and M. Yukawa, "Robust recovery of jointly-sparse signals using minimax concave loss function." IEEE Transactions on Signal Processing, pp. 669–681, 2021.

Outline

1. Background

2. Main Results

3. Numerical Examples
   - Experiment A: Robustness in Overdetermined Case
   - Experiment B: Support Recovery in Underdetermined Case
   - Experiment C: MECG Signal Recovery

4. Conclusion

Experiment A: Robustness in Overdetermined Case

- Signal model: $\boldsymbol{B} = \boldsymbol{X}^* \boldsymbol{A} + \boldsymbol{E} + \boldsymbol{O}$
  - $\boldsymbol{X}^* \in \mathbb{R}^{n \times d}$, $\boldsymbol{A} \in \mathbb{R}^{d \times m}$: i.i.d. $\mathcal{N}(0, 1)$
    ($\boldsymbol{X}^*$ is dense so that the pure effects of robustification can be seen.)
  - $\boldsymbol{E}$: SNR $10, 30$ dB
  - $\boldsymbol{O}$: column sparse, i.i.d. $\mathcal{N}(0, 100)$

Experiment A: Robustness in Overdetermined Case

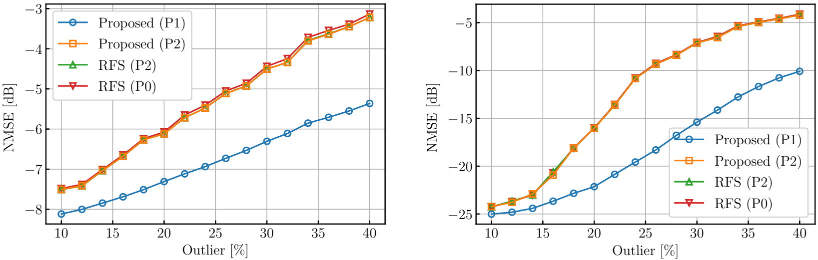
- Signal model: $\boldsymbol{B} = \boldsymbol{X}^* \boldsymbol{A} + \boldsymbol{E} + \boldsymbol{O}$
  - $\boldsymbol{X}^* \in \mathbb{R}^{n \times d}$, $\boldsymbol{A} \in \mathbb{R}^{d \times m}$: i.i.d. $\mathcal{N}(0, 1)$
    ($\boldsymbol{X}^*$ is dense so that the pure effects of robustification can be seen.)
  - $\boldsymbol{E}$: SNR 10, 30 dB
  - $\boldsymbol{O}$: column sparse, i.i.d. $\mathcal{N}(0, 100)$

- Approaches to be tested:
  1. (P_1): $\min_{\boldsymbol{X} \in \mathbb{R}^{n \times d}} \Phi_L(\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}) + \frac{\lambda_2}{2} \|\boldsymbol{X}\|_{\mathrm{F}}^2$ by the proposed approach
     $(\lambda_1 = 0)$
  2. (P_2): $\min_{\boldsymbol{X} \in \mathbb{R}^{n \times d}} \|\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}\|_{2,1} + \frac{\lambda_2}{2} \|\boldsymbol{X}\|_{\mathrm{F}}^2$ by the proposed approach
     $(\lambda_1 = 0,\ L = O)$
  3. (P_2): $\min_{\boldsymbol{X} \in \mathbb{R}^{n \times d}} \|\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}\|_{2,1} + \frac{\lambda}{2} \|\boldsymbol{X}\|_{\mathrm{F}}^2$ by RFS (Nie *et al.*, '10)
  4. (P_0): $\min_{\boldsymbol{X} \in \mathbb{R}^{n \times d}} \|\boldsymbol{B} - \boldsymbol{X}\boldsymbol{A}\|_{2,1} + \lambda \|\boldsymbol{X}\|_{2,1}$ by RFS

Experiment A: Results

- Evaluation metric: normalized mean squared errors (NMSE)

$$\mathrm{NMSE} := \frac{\|\boldsymbol{X}^* - \hat{\boldsymbol{X}}\|_{\mathrm{F}}^2}{\|\boldsymbol{X}^*\|_{\mathrm{F}}^2} \tag{8}$$



(a) SNR 10 dB

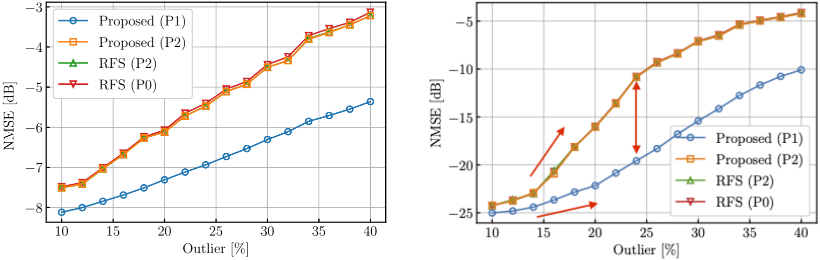
(b) SNR 30 dB

Figure: NMSE for different column-sparsity of outlier matrix under $d = 128$, $m = 256$, and $n = 128$.

Experiment A: Results

- Evaluation metric: normalized mean squared errors (NMSE)

$$\mathrm{NMSE} := \frac{\|\boldsymbol{X}^* - \hat{\boldsymbol{X}}\|_{\mathrm{F}}^2}{\|\boldsymbol{X}^*\|_{\mathrm{F}}^2} \tag{8}$$



Remarkable Outlier–Robustness

(a) SNR 10 dB

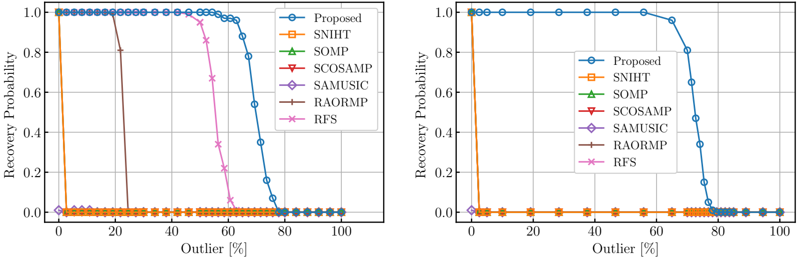
(b) SNR 30 dB

Figure: NMSE for different column-sparsity of outlier matrix under $d = 128$, $m = 256$, and $n = 128$.

Experiment B: Support Recovery in Underdetermined Case

- Signal model: $\boldsymbol{B} = \boldsymbol{X}^* \boldsymbol{A} + \boldsymbol{E} + \boldsymbol{O}$
  - $\boldsymbol{X}^* \in \mathbb{R}^{n \times d}$: column sparse ($k$ non-zero vectors), i.i.d. $\mathcal{N}(0, 1)$
  - $\boldsymbol{A} \in \mathbb{R}^{d \times m}$: i.i.d. $\mathcal{N}(0, 1)$
  - $\boldsymbol{E}$: SNR 30 dB
  - $\boldsymbol{O}$: column sparse (k' non-zero vectors) with $\mathrm{SNR_O} := \frac{\|\boldsymbol{X}^* \boldsymbol{A}\|_{\mathrm{F}}^2 / m}{\|\boldsymbol{O}\|_{\mathrm{F}}^2 / k'}$
- Algorithms to be tested:
  - Proposed approach
  - Subspace augmented MUSIC (SAMUSIC) (Tropp *et al.*, '06)
  - RFS (Nie *et al.*, '10)
  - Simultaneous OMP (SOMP) (Kim *et al.*, '12)
  - Rank aware order recursive matching pursuit (RAORMP) (Davies and Eldar, '12)
  - Simultaneous normalized IHT (SNIHT) (Blanchard *et al.*, '14)
  - Simultaneous COSAMP (SCOSAMP) (Blanchard *et al.*, '14)
- Evaluation metric: success probability of support recovery

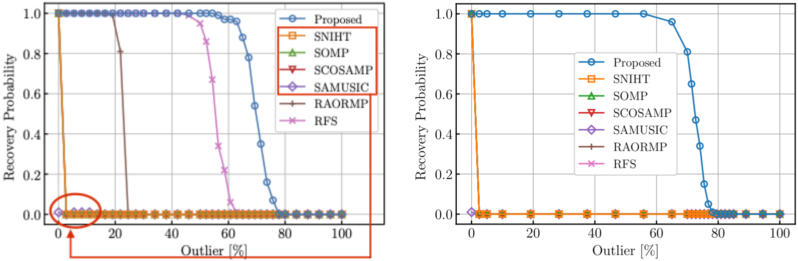
Experiment B: Results (1/2)



(a) $\mathrm{SNR_O}$ -30 dB (b) $\mathrm{SNR_O}$ -3000 dB

Figure: Recovery probability for different column sparsity of outlier matrix under $d = 256$, $m = 128$, $n = 32$, and $k = 16$.

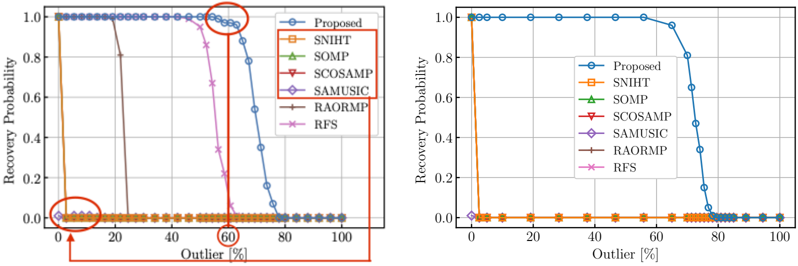
Experiment B: Results (1/2)



(a) $\mathrm{SNR}_{\mathrm{O}}$ -30 dB (b) $\mathrm{SNR}_{\mathrm{O}}$ -3000 dB

Figure: Recovery probability for different column sparsity of outlier matrix under $d = 256$, $m = 128$, $n = 32$, and $k = 16$.

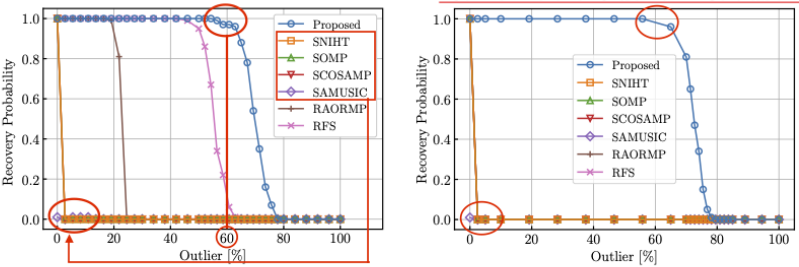
Experiment B: Results (1/2)



(a) $\mathrm{SNR_O}$ -30 dB (b) $\mathrm{SNR_O}$ -3000 dB

Figure: Recovery probability for different column sparsity of outlier matrix under $d = 256$, $m = 128$, $n = 32$, and $k = 16$.

Experiment B: Results (1/2)



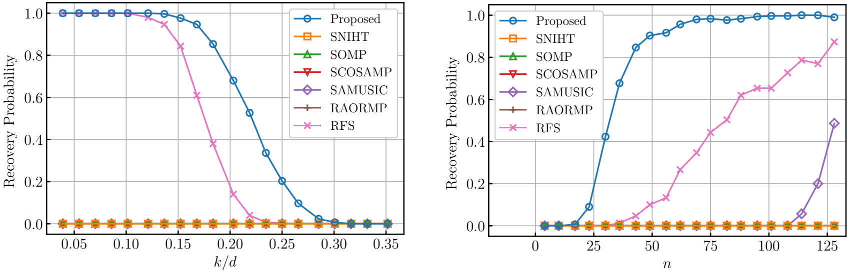
Remarkable Outlier–Robustness

(a) $\mathrm{SNR_O}$ -30 dB

(b) $\mathrm{SNR_O}$ -3000 dB

Figure: Recovery probability for different column sparsity of outlier matrix under $d = 256$, $m = 128$, $n = 32$, and $k = 16$.

Experiment B: Results (2/2)

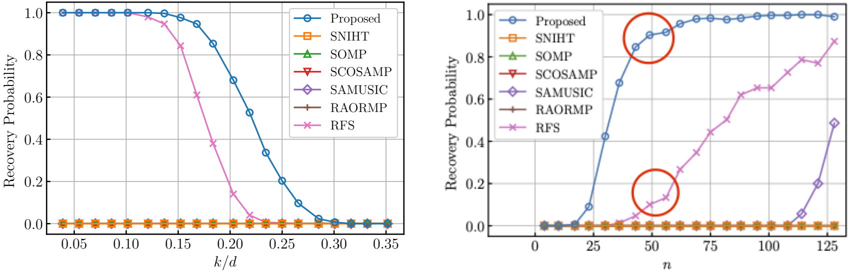


(a) $d = 256$, $m = 128$, $n = 32$.

(b) $d = 192$, $m = 128$, $k = 0.4d$.

Figure: Recovery probability as a function of k/d and n under outlier 30%, and $\mathrm{SNR_O}$ -30 dB.

Experiment B: Results (2/2)



(a) $d = 256,\ m = 128,\ n = 32.$

(b) $d = 192,\ m = 128,\ k = 0.4d.$

Figure: Recovery probability as a function of k/d and n under outlier 30%, and $\mathrm{SNR_O}$ -30 dB.

Experiment C: MECG Signal Recovery

Signal model:

$$\begin{aligned} \boldsymbol{B} &= \boldsymbol{X}^{*}\boldsymbol{A} + \boldsymbol{E} + \boldsymbol{O} = \boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{A} + \boldsymbol{E} + \boldsymbol{O} \\ &= \boldsymbol{W}\boldsymbol{\Theta} + \boldsymbol{E} + \boldsymbol{O} \end{aligned} \tag{9}$$

- $\boldsymbol{B} \in \mathbb{R}^{n \times m}$: publicly available database PTB [6]

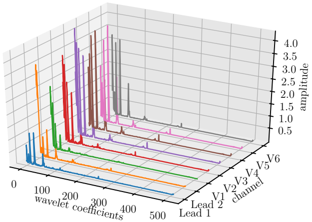
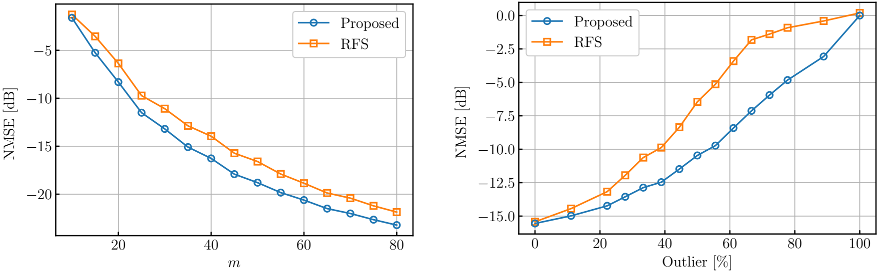


Figure: Amplitudes of ECG wavelet coefficients (**jointly sparse**)

- $\boldsymbol{A} \in \mathbb{R}^{d \times m}$: random sparse binary matrix
- $\boldsymbol{\Psi} \in \mathbb{R}^{d \times d}$: orthonormal wavelet basis
- $\boldsymbol{\Theta} := \boldsymbol{\Psi}\boldsymbol{A}$
- $\boldsymbol{W} \in \mathbb{R}^{n \times d}$: wavelet coefficient vectors
- SNR 30 dB
- $\boldsymbol{O}$: column sparse, $\mathrm{SNR_O} : -40$ dB

[6] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," Circulation, vol. 101, no. 23, pp. e215—e220, 2000.

Experiment C: Results



(a) (b)

Figure: NMSE as a function of m under outlier rate 30%, NMSE as a function of outlier rate under $m = 30$.

Outline

1. Background
2. Main Results
3. Numerical Examples
    - Experiment A: Robustness in Overdetermined Case
    - Experiment B: Support Recovery in Underdetermined Case
    - Experiment C: MECG Signal Recovery
4. Conclusion

Conclusion

- We proposed a robust algorithm to recover jointly-sparse signals in the presence of outliers.
- We showed the convexity condition for the proposed formulation → this led to **the global optimality**.
    - The problem was reformulated via Moreau's decomposition for splitting the cost function into convex terms.
- The proposed approach is **scalable** to high dimensional settings by the use of the efficient primal-dual splitting method.
- Extensive simulation studies showed the **remarkable robustness** of the proposed method to outliers.

Conclusion

- We proposed a robust algorithm to recover jointly-sparse signals in the presence of outliers.
- We showed the convexity condition for the proposed formulation → this led to **the global optimality**.
  - The problem was reformulated via Moreau's decomposition for splitting the cost function into convex terms.
- The proposed approach is **scalable** to high dimensional settings by the use of the efficient primal-dual splitting method.
- Extensive simulation studies showed the **remarkable robustness** of the proposed method to outliers.

To the best of our knowledge, this is the first work leveraging the MC function in a robust framework while maintaining the overall convexity of the whole cost function.

Thank you very much for your kind listening