Exact Recovery of Multichannel Sparse Blind Deconvolution via Gradient Descent

Qing Qu

Center for Data Science

New York University

Joint with Xiao Li (CUHK) and Zhihui Zhu (JHU)

December 12, 2019

Multichannel Sparse Blind Deconvolution

Given multiple measurement

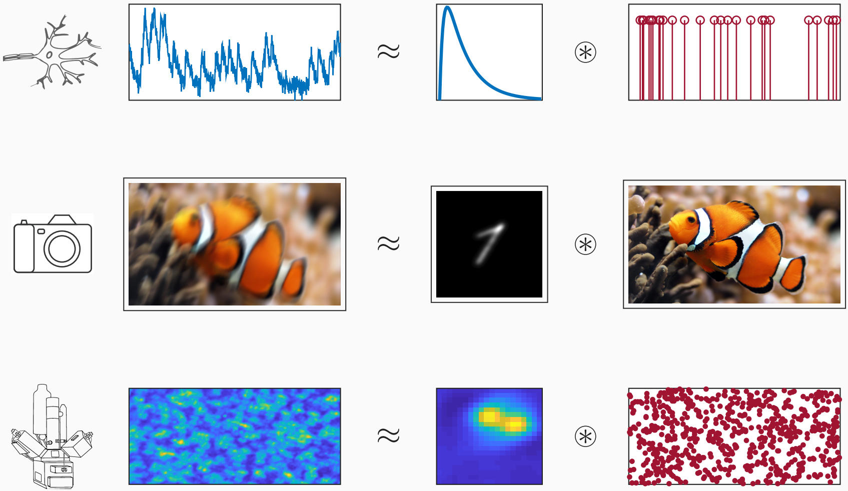
$$\boxed{\boldsymbol{y}_i = \boldsymbol{a} \circledast \boldsymbol{x}_i, \quad (1 \leq i \leq p),}$$

can we recover both $\boldsymbol{a}$ and $\{\boldsymbol{x}_i\}_{i=1}^{p}$ simultaneously?

- We assume $\boldsymbol{y}_i, \ \boldsymbol{a}, \ \boldsymbol{x}_i \in \mathbb{R}^n$.
- **Invertible** kernel $\boldsymbol{a}$.
- **Sparse** signal $\boldsymbol{x}_i$

$$\boldsymbol{x}_i \sim_{i.i.d.} \text{Bernoulli} - \text{Gaussian}(\theta)$$

Motivating Applications

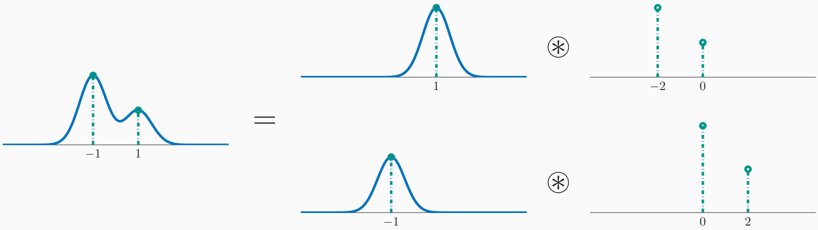


Symmetry Leads to Nonconvex Problems

- **Scaling Symmetry**: $\boldsymbol{y}_i = \boldsymbol{a} \circledast \boldsymbol{x}_i = \alpha \boldsymbol{a} \circledast \alpha^{-1} \boldsymbol{x}_i$
  - easy to handle, $\|\boldsymbol{a}\| = 1$;

Symmetry Leads to Nonconvex Problems

- **Scaling Symmetry**: $\boldsymbol{y}_i = \boldsymbol{a} \circledast \boldsymbol{x}_i = \alpha \boldsymbol{a} \circledast \alpha^{-1} \boldsymbol{x}_i$
  - easy to handle, $\|\boldsymbol{a}\| = 1$;
- **Shift Symmetry**: $\boldsymbol{y}_i = \boldsymbol{a} \circledast \boldsymbol{x}_i = \mathrm{s}_\ell[\boldsymbol{a}] \circledast \mathrm{s}_{-\ell}[\boldsymbol{x}_i]$

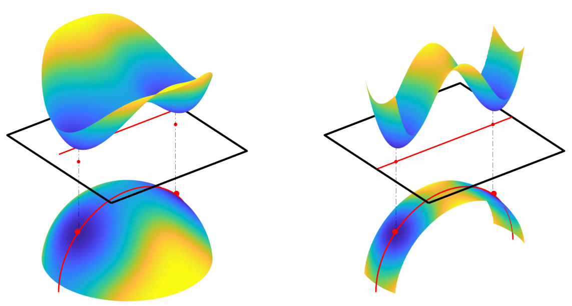


$$\boldsymbol{y}_i = \alpha\, s_\ell[\boldsymbol{a}] \circledast (1/\alpha)\, s_{-\ell}[\boldsymbol{x}_i]$$

Symmetry Leads to Nonconvex Problems

- **Scaling Symmetry**: $\boldsymbol{y}_i = \boldsymbol{a} \circledast \boldsymbol{x}_i = \alpha \boldsymbol{a} \circledast \alpha^{-1} \boldsymbol{x}_i$
  - easy to handle, $\|\boldsymbol{a}\| = 1$;
- **Shift Symmetry** creates equivalent solutions:

$$\left(\boldsymbol{a}, \{\boldsymbol{x}_i\}_{i=1}^{p}\right) \quad = \quad \left(\mathrm{s}_\ell\left[\boldsymbol{a}\right], \{\mathrm{s}_{-\ell}\left[\boldsymbol{x}_i\right]\}_{i=1}^{p}\right)$$



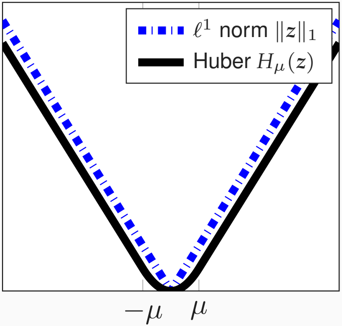
Nonconvex Formulation

Finding a shift of the filter $\boldsymbol{a}$ by solving

$$\boxed{\min_{\boldsymbol{q}} \ \frac{1}{np}\sum_{i=1}^{p} H_\mu\left(\boldsymbol{C}_{\boldsymbol{y}_i}\boldsymbol{P}\boldsymbol{q}\right), \quad \text{s.t.} \quad \boldsymbol{q} \in \mathbb{S}^{n-1}.}$$

Huber loss: 1st-order smooth & sparsity promoting

$$H_\mu(z) := \begin{cases} |z| & |z| \geq \mu \\ \frac{z^2}{2\mu} + \frac{\mu}{2} & |z| < \mu \end{cases}$$



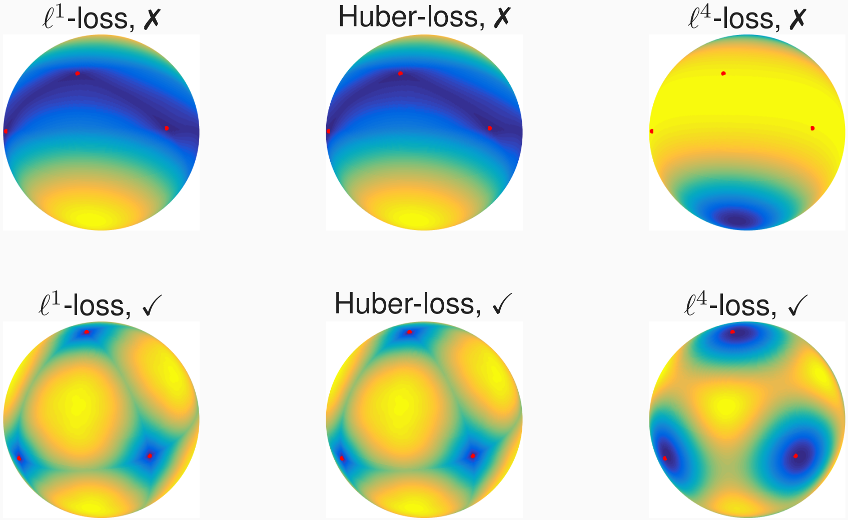
Finding a shift of the filter $\boldsymbol{a}$ by solving

$$\min_{\boldsymbol{q}} \frac{1}{np}\sum_{i=1}^{p} H_{\mu}\left(\boldsymbol{C}_{\boldsymbol{y}_i}\boldsymbol{P}\boldsymbol{q}\right), \quad \text{s.t.} \quad \boldsymbol{q} \in \mathbb{S}^{n-1}.$$

- Preconditioning leads to better landscape

$$\boldsymbol{P} = \left(\frac{1}{\theta np}\sum_{i=1}^{p} \boldsymbol{C}_{\boldsymbol{y}_i}^{\top}\boldsymbol{C}_{\boldsymbol{y}_i}\right)^{-1/2} \approx \left(\boldsymbol{C}_{\boldsymbol{a}}^{\top}\boldsymbol{C}_{\boldsymbol{a}}\right)^{-1/2},$$

Landscape with/without Preconditioning



Main Result

With random init., gradient descent solves sparse blind deconvolution in a linear rate.

Main Result

With random init., gradient descent solves sparse blind deconvolution in a linear rate.

- Study the geometry properties of optimization landscape.
  - regularity condition, implicit regularization, sharpness.

Main Result

With random init., gradient descent solves sparse blind deconvolution in a linear rate.

- Study the geometry properties of optimization landscape.
  - regularity condition, implicit regularization, sharpness.
- Benign geometry enables efficient optimization.

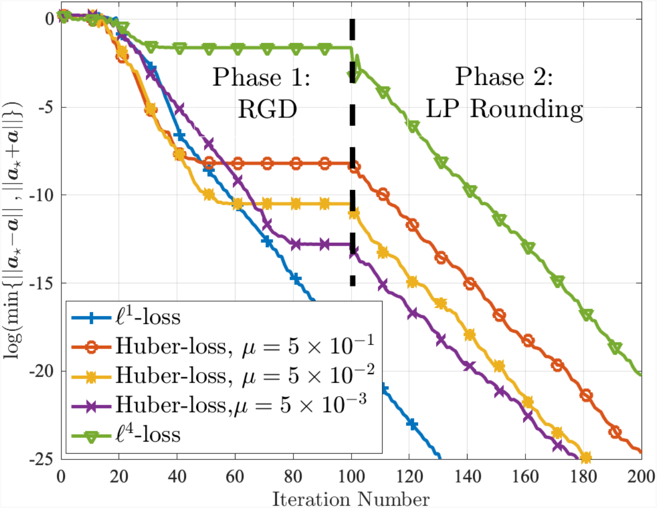
Comparison with Literature

Significant improvements in sample and time complexity.

Methods	Wang et al.[1]	Li et al.[2]	**Ours**
Assumptions	$\boldsymbol{a}$ spiky & invertible, $\boldsymbol{x}_i \sim_{i.i.d.} \mathcal{BG}(\theta)$	$\boldsymbol{a}$ invertible, $\boldsymbol{x}_i \sim_{i.i.d.} \mathcal{BR}(\theta)$	$\boldsymbol{a}$ invertible, $\boldsymbol{x}_i \sim_{i.i.d.} \mathcal{BG}(\theta)$
Formulation	$\min_{\lVert \boldsymbol{q} \rVert_\infty = 1} \lVert \boldsymbol{C_q Y} \rVert_1$	$\max_{\boldsymbol{q} \in \mathbb{S}^{n-1}} \lVert \boldsymbol{C_q P Y} \rVert_4^4$	$\min_{\boldsymbol{q} \in \mathbb{S}^{n-1}} H_\mu \left(\boldsymbol{C_q P Y} \right)$
Algorithm	interior point	*noisy* RGD	*vanilla* RGD
Recovery Condition	$\theta \in \mathcal{O}(1/\sqrt{n})$, $p \geq \widetilde{\Omega}(n)$	$\theta \in \mathcal{O}(1)$, $p \geq \widetilde{\Omega}(\max\left\{n, \kappa^8\right\} \frac{n^8}{\varepsilon^8})$	$\theta \in \mathcal{O}(1)$, $p \geq \widetilde{\Omega}(\max\left\{n, \frac{\kappa^8}{\mu^2}\right\} n^4)$
Time Complexity	$\widetilde{\mathcal{O}}(p^4 n^5 \log(1/\varepsilon))$	$\widetilde{\mathcal{O}}(pn^{13}/\varepsilon^8)$	$\widetilde{\mathcal{O}}(pn^5 + pn \log\left(1/\varepsilon\right))$

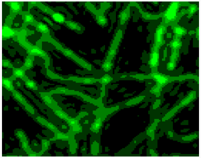
2. Wang et al., blind deconvolution from multiple sparse inputs, 2016.
3. Li et al., Multichannel sparse blind deconvolution on the sphere, 2018.

Experiment I: Convergence Comparison

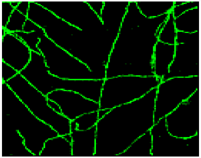


Experiment II: Super-resolution Microscopy

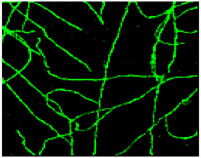
Observation



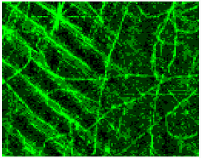
Ground truth



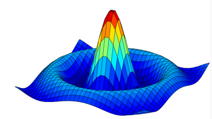
Huber-loss



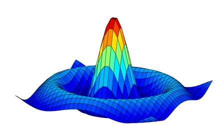
ℓ^4-loss[3]



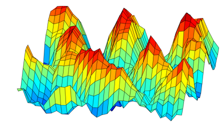
Ground truth



Huber-loss



ℓ^4-loss



Take home message

With random init., gradient descent solves sparse blind deconvolution in a linear rate.

Poster: Hall B + C #207

Acknowledgement

Xiao Li
(CUHK, EE)

Zhihui Zhu
(JHU, MINDS)

THANK YOU!

…AND