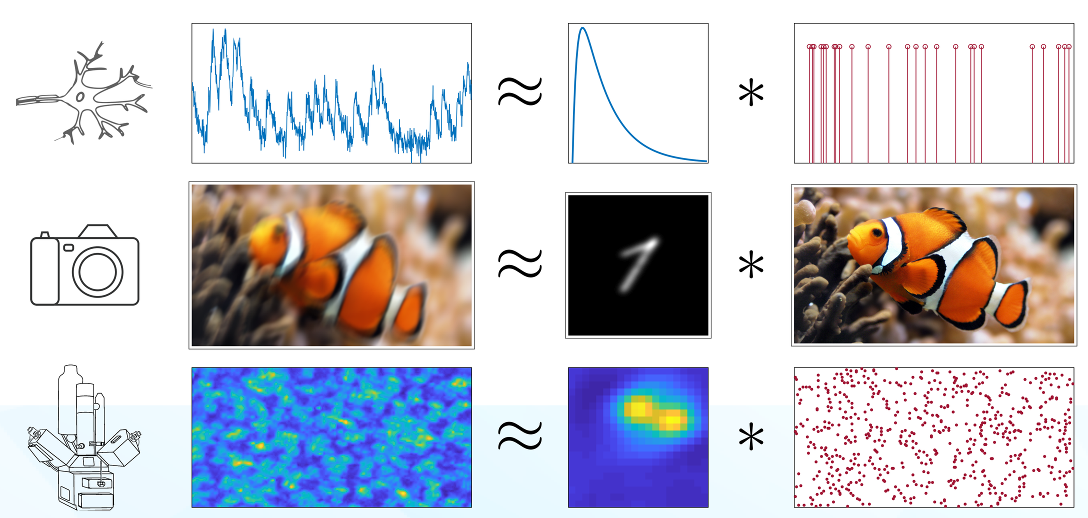
Geometry and Symmetry in Short-and-Sparse Deconvolution

Han-Wen (Henry) Kuo, Yuqian Zhang, Yenson Lau and John Wright
Columbia University

Short-and-Sparse Model

- Model signals containing repeated (short) motifs:

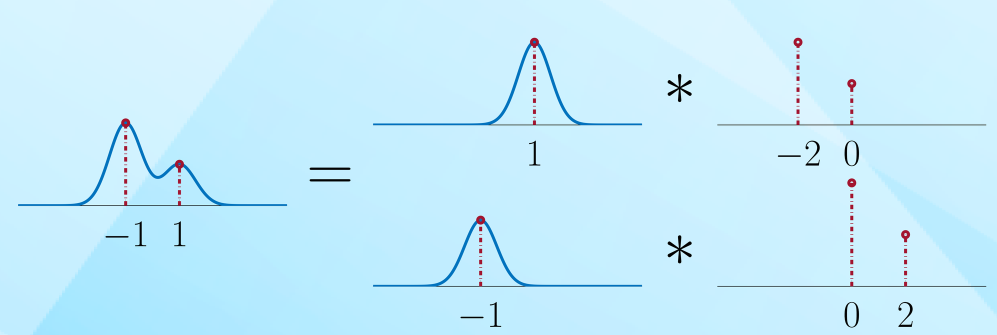


Problem: SaS Deconvolution

Given the cyclic convolution $\boldsymbol{y} = \boldsymbol{a}_0 * \boldsymbol{x}_0 \in \mathbb{R}^n$ of $\boldsymbol{a}_0 \in \mathbb{R}^{p_0}$ short ($p_0 \ll n$), and $\boldsymbol{x}_0 \in \mathbb{R}^n$ sparse, recover $\boldsymbol{a}_0$ and $\boldsymbol{x}_0$, up to a scaled shift.

Symmetric Solutions in SaSD

- All **scaled** & **shifts** of ($\boldsymbol{a}_0$,$\boldsymbol{x}_0$) are solutions to SaSD



$$\boldsymbol{y} \;=\; \alpha s_\ell[\boldsymbol{a}_0] \;*\; (1/\alpha) s_{-\ell}[\boldsymbol{x}_0]$$

- We fix the scale $\|\bar{\boldsymbol{a}}\|_2 = 1$.
- Signed shifts $\pm\{s_\ell[\boldsymbol{a}_0] : \ell = -p_0+1, \dots, p_0-1\}$ are solutions.

Algorithm: Approximate Bilinear Lasso

- Natural, effective method to SaSD: *bilinear Lasso* [1].

$$\min_{\boldsymbol{a}\in\mathbb{S}^{p-1},\, \boldsymbol{x}\in\mathbb{R}^n} \lambda \|\boldsymbol{x}\|_1 + \tfrac{1}{2}\|\boldsymbol{a} * \boldsymbol{x} - \boldsymbol{y}\|_2^2. \qquad (1)$$

- To understand (1), we study a simplification: *"approximate bilinear Lasso"*:

$$\min_{\boldsymbol{a}\in\mathbb{S}^{p-1}} \left(\min_{\boldsymbol{x}\in\mathbb{R}^n} \lambda \boldsymbol{\rho}(\boldsymbol{x}) + \tfrac{1}{2}\|\boldsymbol{x}\|_2^2 + \langle \boldsymbol{a} * \boldsymbol{x}, \boldsymbol{y}\rangle \right)$$

$$=: \boxed{\min_{\boldsymbol{a}} \;\; \varphi_{\mathrm{ABL}}(\boldsymbol{a}) \quad s.t. \quad \boldsymbol{a} \in \mathbb{S}^{p-1}} \qquad (2)$$

- $\boldsymbol{\rho}$ smoothed approximates ℓ^1-sparsity surrogate.
- $\frac{1}{2}\|\boldsymbol{x}\|_2^2 + \langle \boldsymbol{a} * \boldsymbol{x}, \boldsymbol{y}\rangle$ approximates least square.
- Marginal minimize $\boldsymbol{a}$ over sphere.
- Domain dimension $p \approx 3p_0$ contains support of all shifts.

Geometry of Objective Landscape

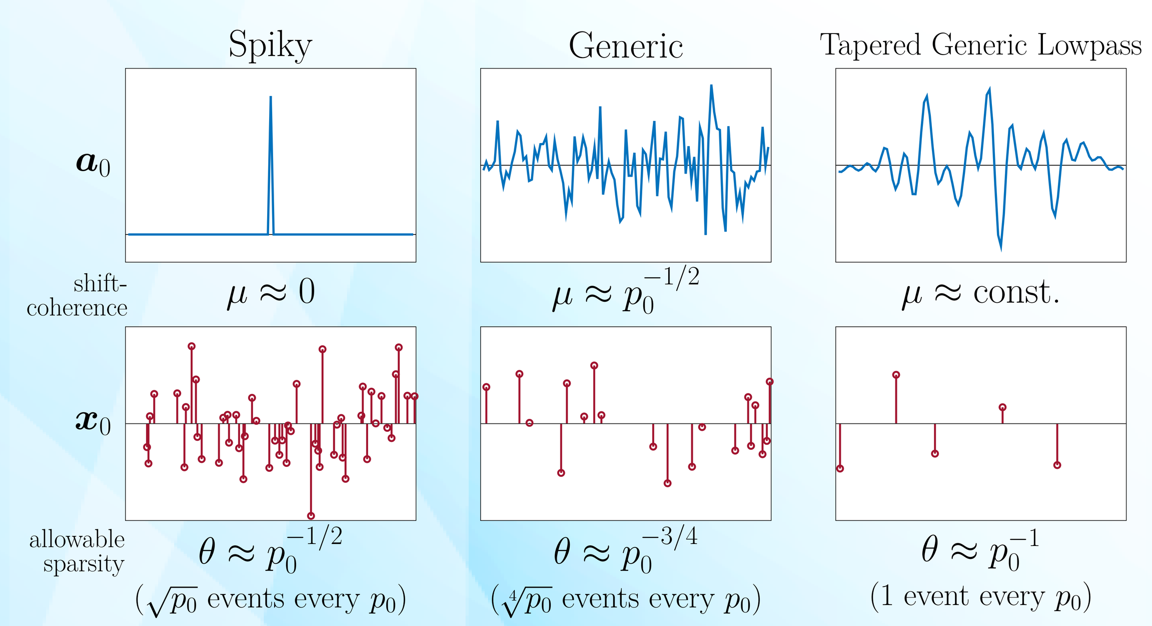
The geometry of φ_{ABL} over the **sphere** $\mathbb{S}^{p-1}$ is determined by the **shifts of** $\boldsymbol{a}_0$ (the solutions of SaSD).
φ_{ABL} is convex near every signed shift, and exhibits negative curvature at points that are superpositions of a few shifts. This regional geometry holds for every combination of shifts, whenever $\boldsymbol{x}_0/\boldsymbol{a}_0$ satisfy sparsity/coherence conditions.

Sparsity-Coherence Tradeoff

- **Shift-coherence** μ of $\boldsymbol{a}_0$:

$$\mu(\boldsymbol{a}_0) = \max_{i\neq j} |\langle s_i[\boldsymbol{a}_0], s_j[\boldsymbol{a}_0]\rangle| \qquad (3)$$

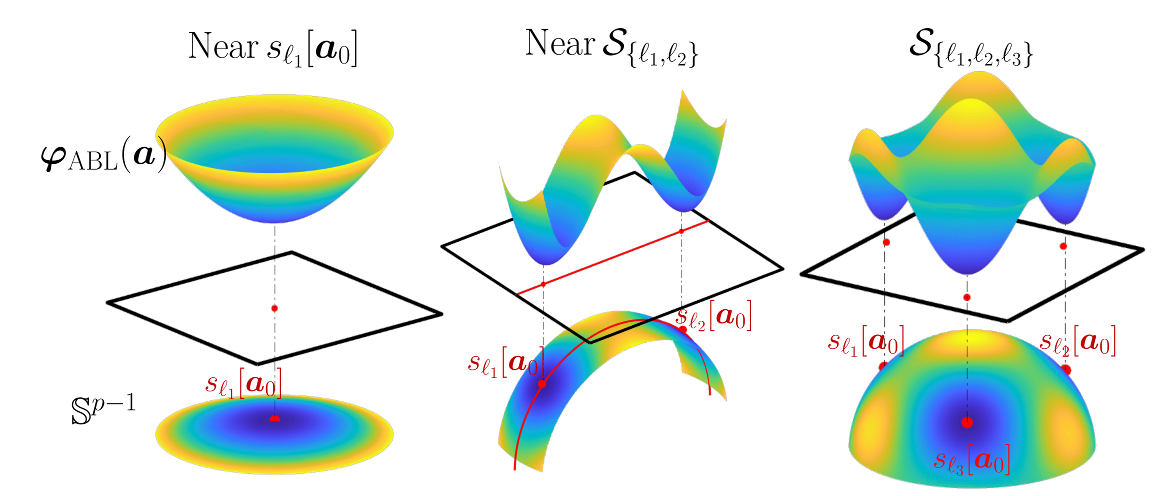
- **Sparsity rate** θ of $\boldsymbol{x}_0$: $\boldsymbol{x}_0 \sim_{\mathrm{i.i.d.}} \mathrm{BG}(\theta)$.
- SaSD is harder if $\boldsymbol{a}_0$ is more shift-coherent (solutions are closer on sphere) or $\boldsymbol{x}_0$ is denser (more unknowns).



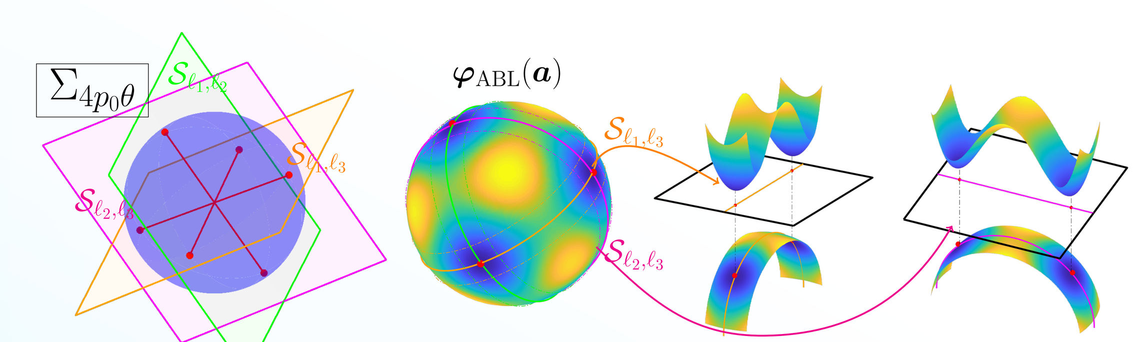
If shift-coherence of $\boldsymbol{a}_0$ increases from 0 to 1, then allowable sparsity of $\boldsymbol{x}_0$ decreases from $1/\sqrt{p_0}$ to $1/p_0$

Geometry over Subspace of Shifts

- φ_{ABL} has **local minimizer** near shifts and has **negative curvature** breaks symmetry in subspace $\mathcal{S}_{\{\ell_1,\cdots,\ell_3\}}$ spanned by shifts $\{s_{\ell_1}[\boldsymbol{a}_0],\cdots,s_{\ell_3}[\boldsymbol{a}_0]\}$.



- The geometry of φ_{ABL} is benign over $\Sigma_{4p_0\theta}$, which is the **union of subspaces** spanned by $4p_0\theta$ shifts.



Theorem 1: Geometry of φ_{ABL} over Union of Subspaces

Let $\boldsymbol{y} = \boldsymbol{a}_0 * \boldsymbol{x}_0$ with $\boldsymbol{a}_0 \in \mathbb{S}^{p_0-1}$ μ-shift coherent and $\boldsymbol{x}_0 \sim_{\mathrm{i.i.d.}} \mathrm{BG}(\theta) \in \mathbb{R}^n$ with sparsity rate

$$\theta \in \left[\frac{c_1}{p_0}, \frac{c_2}{p_0\sqrt{\mu} + \sqrt{p_0}}\right] \cdot \frac{1}{\log^2 p_0}. \qquad (4)$$

Set $\boldsymbol{\rho}(x) = \sqrt{x^2+\delta^2}$ and $\lambda = 0.1/\sqrt{p_0\theta}$ in φ_{ABL}. There exists $c, \delta > 0$ such that if $n \geq \mathrm{poly}(p_0)$, with high probability, every local minimizer $\bar{\boldsymbol{a}}$ of φ_{ABL} over $\Sigma_{4\theta p_0}$ satisfies $\|\bar{\boldsymbol{a}} - \sigma s_\ell[\boldsymbol{a}_0]\|_2 \leq c \max\{\mu, p_0^{-1}\}$.

From Geometry to Provable Algorithm

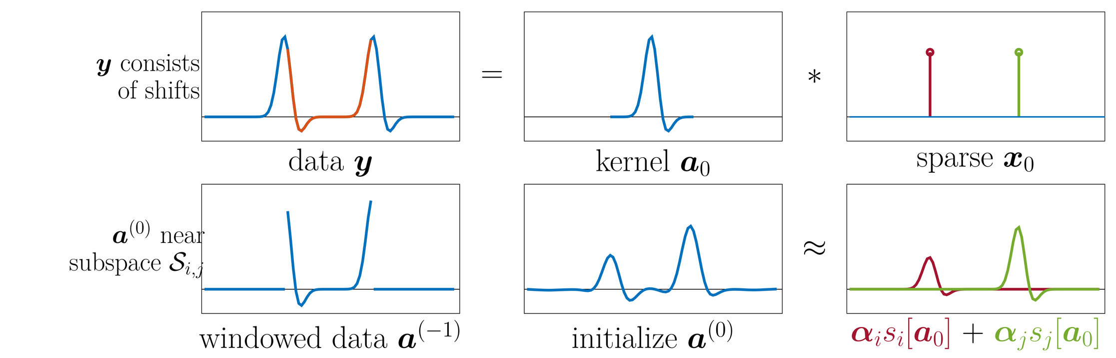
Design a **provable** algorithm for **exact recovery** based on the geometry of φ_{ABL}.
The algorithm initializes $\boldsymbol{a}^{(0)}$ near one of the subspaces in $\Sigma_{4\theta p_0}$; then the geometry of φ_{ABL} ensures small stepping descent method stay near subspace and converges toward the local minimizer close to a shift.

Theorem 2: Provable Algorithm of SaSD

Suppose $\boldsymbol{a}_0$ is μ-truncated shift coherent and $\boldsymbol{x}_0 \sim_{\mathrm{i.i.d.}} \mathrm{BG}(\theta) \in \mathbb{R}^n$ with θ, μ satisfying (4) and $\mu \leq \frac{c_3}{\log^2 n}$. If lengths n, p_0 satisfy $n > \mathrm{poly}(p_0)$ and $p_0 > \mathrm{polylog}(n)$, then with high probability, our algorithm produces $(\hat{\boldsymbol{a}}, \hat{\boldsymbol{x}})$ satisfies $\left\|(\hat{\boldsymbol{a}}, \hat{\boldsymbol{x}}) - \sigma\big(s_\ell[\boldsymbol{a}_0],\, s_{-\ell}[\boldsymbol{x}_0]\big)\right\|_2 \leq \varepsilon$ with running time $\mathcal{O}(\mathrm{poly}(n, p_0, \varepsilon^{-1}))$.

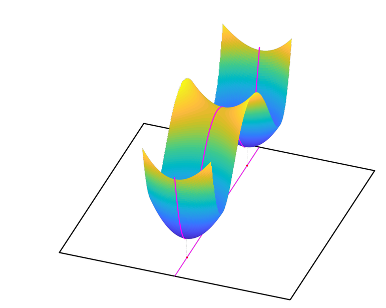
Provable Algorithm of SaSD

- *Initialize*: Use chunk of $\boldsymbol{y}$ (sum of truncated shifts)



where $\boldsymbol{a}^{(0)} = -\mathcal{P}_{\mathbb{S}^{p-1}}\nabla\varphi_{\mathrm{ABL}}\mathcal{P}_{\mathbb{S}^{p-1}}[\boldsymbol{a}^{(-1)}]$.

- *Minimization*: Small step descent method stays near subspace since φ_{ABL} *grows away from subspace.*

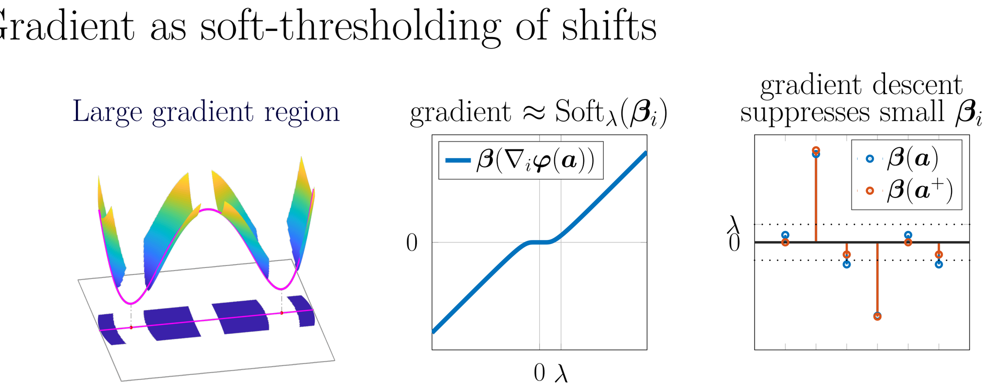


- *Refinement*: (sketch) Alternating minimize bilinear Lasso converges to exact solution at a linear rate.

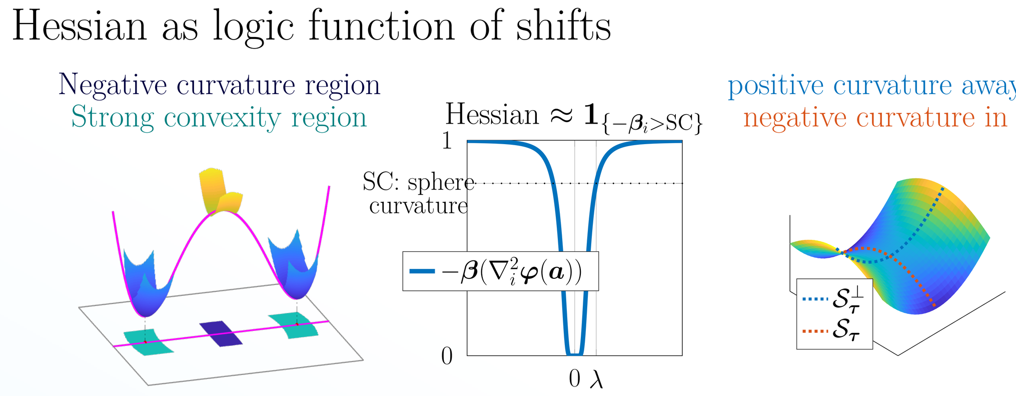
Analysis: Sparsifying in Shift Space

During minimization the summands of shifts in $\boldsymbol{a}^{(0)}$ sparsifies until one shift left. Write $\boldsymbol{\beta}(\boldsymbol{a})$ as "shift space coefficients" of $\boldsymbol{a}$:

- Gradient as soft-thresholding of shifts



- Hessian as logic function of shifts



Discussion

- Our main contribution is in *theory* (optimize φ_{ABL} is not recommended in practice), but the ideas are useful for developing practical algorithms [2].

References

[1] Y. Zhang, Y. Lau, H-W. Kuo, S.Cheung, A. Pasupathy and J. Wright, "On the global geometry of sphere-constrained sparse blind deconvolution", *CVPR*, 2017.
[2] Y. Lau, Q. Qu, H-W. Kuo, Y. Zhang, P. Zhou and J. Wright, "Short-and-Sparse Deconvolution-A Geometric Approach". *Submitted*, 2019.

Columbia University in the City of New York