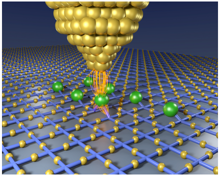
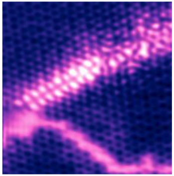
Sparse Blind Deconvolution: Nonconvex Geometry and Algorithm

Yuqian Zhang
Rutgers University

Application: Scanning Tunneling Microscopy

Abhay Pasupathy
Columbia Physics

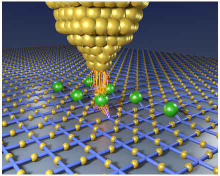



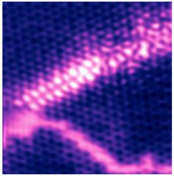
Scanning Tunneling Microscope

.

Application: Scanning Tunneling Microscopy

Abhay Pasupathy
Columbia Physics

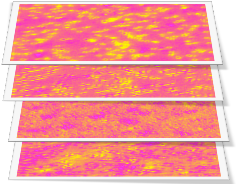




Scanning Tunneling Microscope

.

Scanning tunneling spectroscopy:
Interrogate material at different points in space $\times$ energy

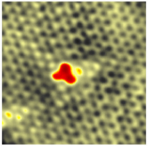


STS Data Cube
Space $\times$ Voltage

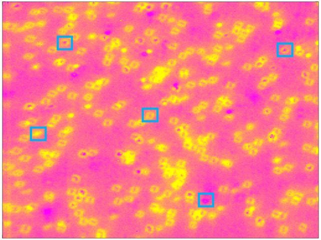
Defects in the Crystal Lattice

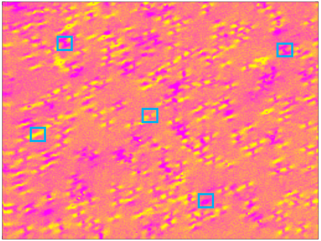
Defects in the crystal lattice encode electronic / material properties:

Defects have a characteristic signature (**motif**):



Doped Graphene

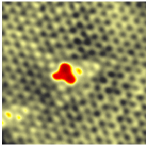




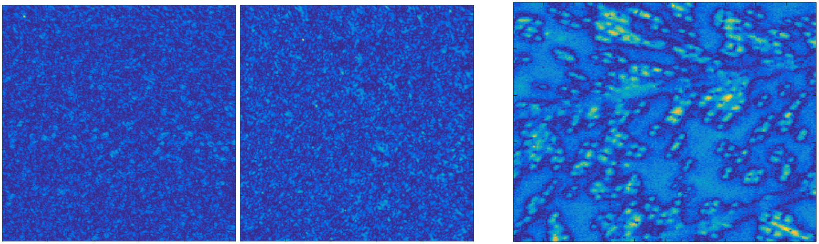
Can we determine the basic motifs and their locations?

Defects in the Crystal Lattice

Defects in the crystal lattice encode electronic / material properties: Defects have a characteristic signature (**motif**):

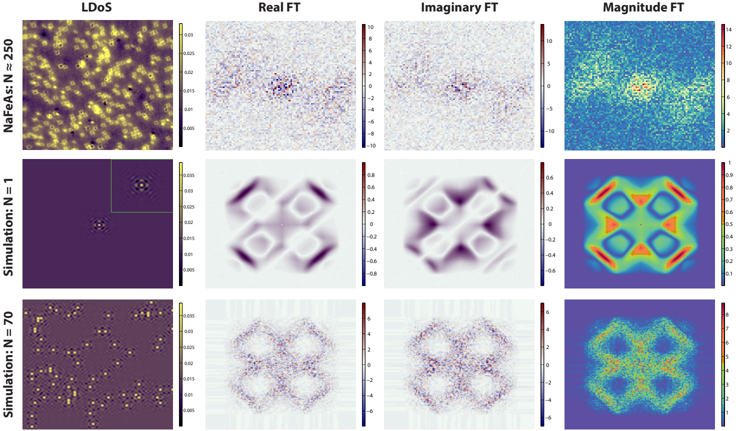


Doped Graphene



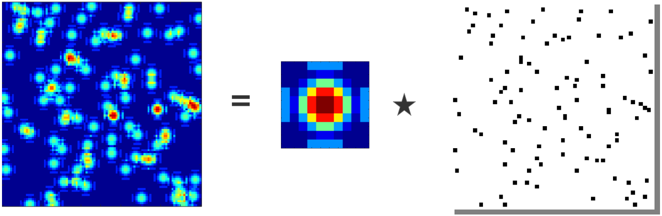
Can we determine the basic motifs and their locations?

Current Approach: Fourier Transform STS



$$\hat{\boldsymbol{y}}(\boldsymbol{\omega}) = \underbrace{\sum_{i=1}^{L} \exp\left\{-j\left\langle \boldsymbol{w}, \boldsymbol{x}_i \right\rangle\right\}}_{\text{Frequency-variant ``phase noise''}} \times \underbrace{\hat{\boldsymbol{a}}(\boldsymbol{\omega})}_{\text{Defect signature (Fourier)}}. \qquad (1)$$

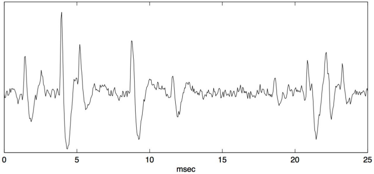
$$\boldsymbol{y} \;=\; \boldsymbol{a}_0 \circledast \boldsymbol{x}_0 \in \mathbb{R}^m$$



- $\boldsymbol{a}_0$ is short;
- $\boldsymbol{x}_0$ has a sparse and random support.

Other Scientific Data

- Neural Spike Sorting



- Astrophysical Data (LIGO)

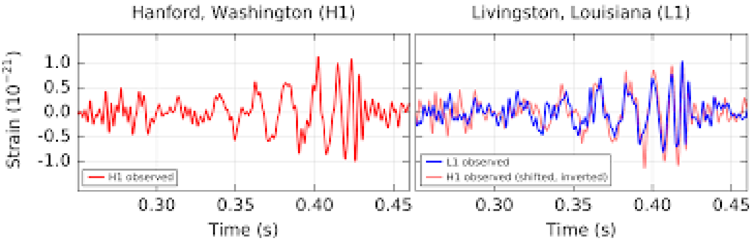
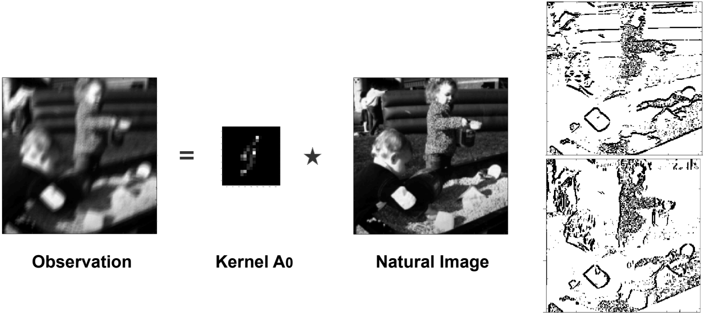


Image Deblurring

Natural images are **sparse** in the *gradient* domain:



Problem Formulation

Short and Sparse Blind Deconvolution

Given observation

$$\boldsymbol{y} = \boldsymbol{a}_0 \circledast \boldsymbol{x}_0 \in \mathbb{R}^m,$$

can we recover both unknown signals $\boldsymbol{a}_0 \in \mathbb{R}^k$ and $\boldsymbol{x}_0 \in \mathbb{R}^m$?

- $\boldsymbol{a}_0$ is short: $k << m$;
- $\boldsymbol{x}_0$ has a sparse and random support.

Symmetries

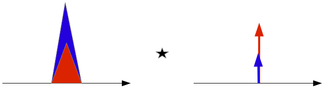
The major difficulty comes from the symmetric solutions!

Symmetries

The major difficulty comes from the symmetric solutions!

- Scale and Sign Symmetry

$$\boldsymbol{a} = \pm\alpha\boldsymbol{a}_0, \quad \boldsymbol{x} = \pm\frac{1}{\alpha}\boldsymbol{x}_0$$

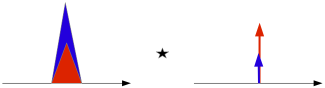


Symmetries

The major difficulty comes from the symmetric solutions!

- Scale and Sign Symmetry

$$\boldsymbol{a} = \pm\alpha\boldsymbol{a}_0, \quad \boldsymbol{x} = \pm\frac{1}{\alpha}\boldsymbol{x}_0$$

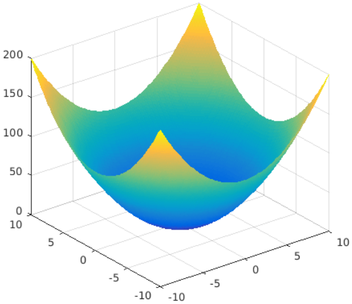


- Shift Symmetry

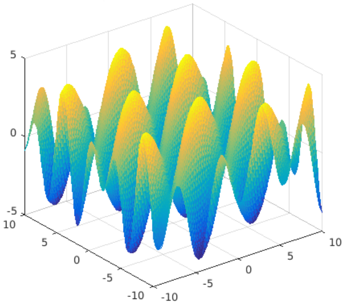
$$\boldsymbol{a} = s_\tau\left[\boldsymbol{a}_0\right], \quad \boldsymbol{x} = s_{-\tau}\left[\boldsymbol{x}_0\right]$$

Nonconvexity in Sparse Blind Deconvolution

Each symmetric solution creates a local optima.

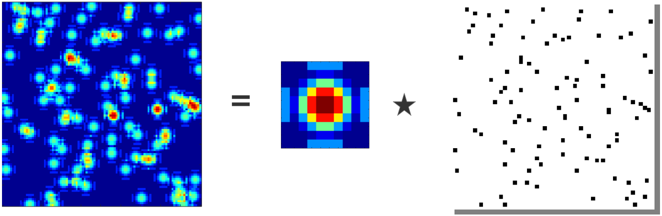


“easy”



“hard”

Nonconvex Formulation



$$\min_{\boldsymbol{a},\boldsymbol{x}} \quad \underbrace{\tfrac{1}{2}\left\|\boldsymbol{y}-\boldsymbol{a}\circledast\boldsymbol{x}\right\|_F^2}_{\text{data fidelity}} + \underbrace{\lambda\left\|\boldsymbol{x}\right\|_1}_{\text{sparsity}}$$

$$\text{s.t.} \quad \underbrace{\boldsymbol{a}\in\mathbb{R}^k,\ \left\|\boldsymbol{a}\right\|_F = 1}_{\boldsymbol{a}\in\mathbb{S}^{k-1}}$$

$$\min_{\boldsymbol{a},\boldsymbol{x}} \ \frac{1}{2}\left\|\boldsymbol{y}-\boldsymbol{a}\circledast\boldsymbol{x}\right\|_F^2+\lambda\left\|\boldsymbol{x}\right\|_1 \quad \text{s.t.} \quad \boldsymbol{a}\in\mathbb{S}^{k-1}$$

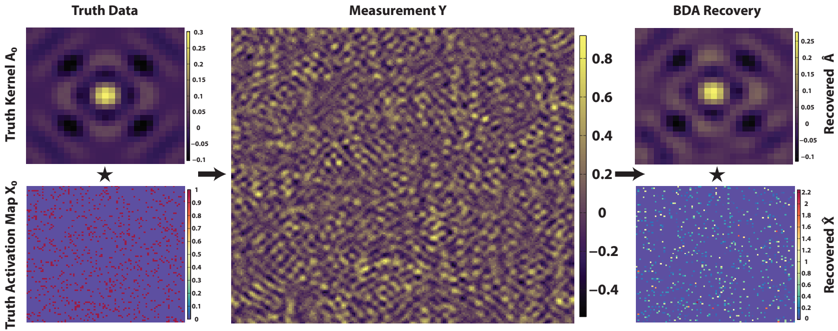


Figure 1: Synthetic Microscopy Data

Microscopy Image Analysis - Real Data I

$$\min_{\boldsymbol{a},\boldsymbol{x}} \ \frac{1}{2}\left\|\boldsymbol{y}-\boldsymbol{a}\circledast\boldsymbol{x}\right\|_F^2+\lambda\left\|\boldsymbol{x}\right\|_1 \quad \text{s. t.} \ \boldsymbol{a}\in\mathbb{S}^{k-1}$$

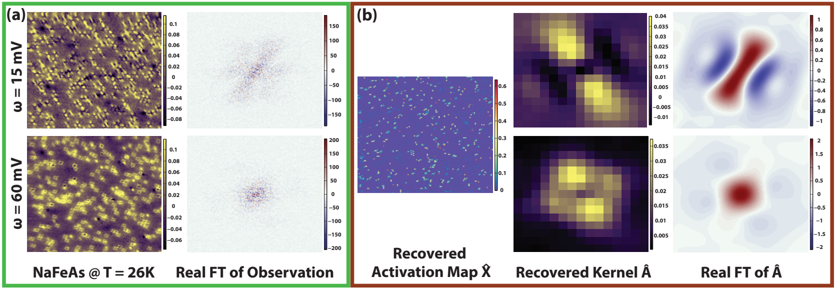


Figure 2: Real Microscopy Data

Microscopy Image Analysis - Real Data II

$$
\begin{aligned}
\min_{\boldsymbol{a}_n, \boldsymbol{x}_n} \quad & \frac{1}{2} \left\| \boldsymbol{y} - \sum_{n=1}^{N} \boldsymbol{a}_n \circledast \boldsymbol{x}_n \right\|_F^2 + \sum_{n=1}^{N} \lambda \left\| \boldsymbol{x}_n \right\|_1 \\
\text{s. t.} \quad & \boldsymbol{a}_n \in \mathbb{S}^{k-1}
\end{aligned}
$$

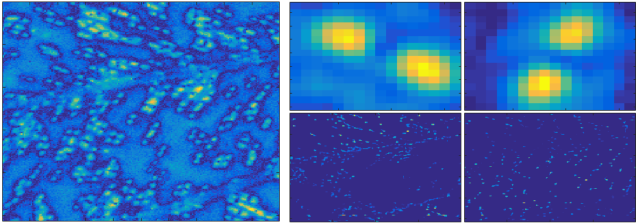
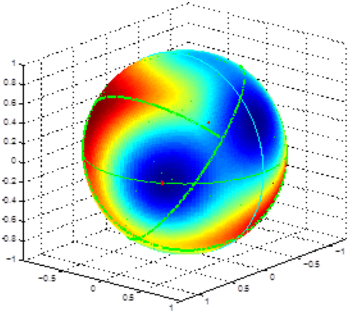
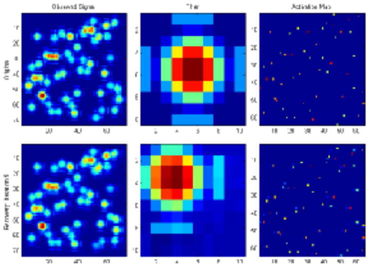


Figure 3: Multiple Defects Patterns

Local Optima are Good — Geometry

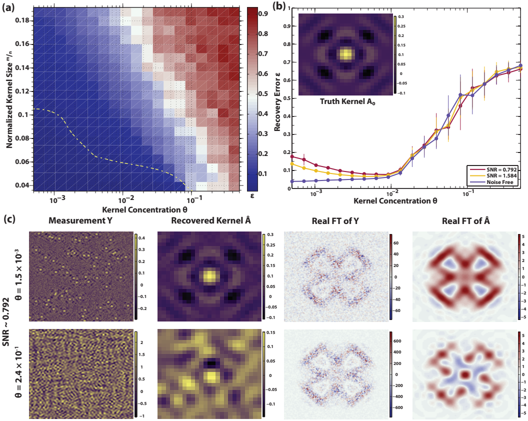




$$\varphi(\boldsymbol{a}) = \min_{\boldsymbol{x}} \; \frac{1}{2} \left\| \boldsymbol{y} - \boldsymbol{a} \circledast \boldsymbol{x} \right\|_F^2 + \lambda \left\| \boldsymbol{x} \right\|_1$$

Local minima are near signed shift-truncations.

Microscopy – easy vs. hard problems



Empirical observation: Whenever the target x_0 is sufficiently long and sparse, recover a_0 up to signed shift truncation.

Theory question: When and why does this occur?

Main Result

Guaranteed **SHORT-AND-SPARSE** deconvolution w.h.p., when

$$\frac{1}{k} \le \theta \le \frac{c}{\left(\sqrt{k} + \mu^{1/2} k\right) \log k}, \qquad m \ge \mathrm{poly}(k)$$

with shift-incoherence $\mu \doteq \max_{i \neq 0} \left| \langle \boldsymbol{a}_0, s_i\left[\boldsymbol{a}_0\right] \rangle \right|$.

Main Result

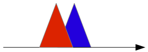
Guaranteed **SHORT-AND-SPARSE** deconvolution w.h.p., when

$$\frac{1}{k} \le \theta \le \frac{c}{\left(\sqrt{k} + \mu^{1/2} k\right) \log k}, \qquad m \ge \mathrm{poly}(k)$$

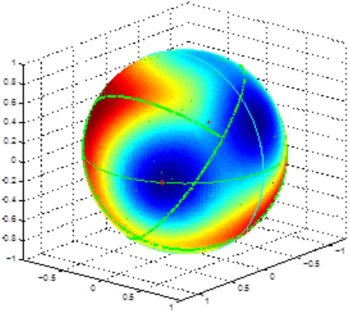
with shift-incoherence $\mu \doteq \max_{i \ne 0} |\langle \boldsymbol{a}_0, s_i[\boldsymbol{a}_0] \rangle|$.

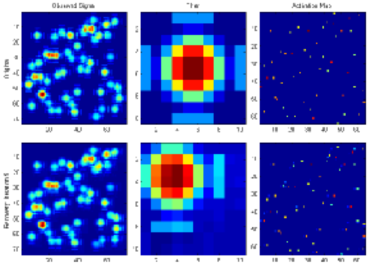
Comment on rates: For $\boldsymbol{a}_0 \sim \mathrm{uni}(\mathbb{S}^{k-1})$, success w.h.p. when:

$$\theta \lesssim \frac{1}{k^{3/4} \mathrm{polylog}(k)}.$$

$\approx k^{1/4}$ "copies" of $\boldsymbol{a}_0$ in each length-k window: 

Optimization Landscape





$$\varphi(\boldsymbol{a}) = \min_{\boldsymbol{x}} \; \frac{1}{2} \left\| \boldsymbol{y} - \boldsymbol{a} \circledast \boldsymbol{x} \right\|_F^2 + \lambda \left\| \boldsymbol{x} \right\|_1$$

There is no closed form expression for $\widehat{x}_{\textsf{Lasso}}$.

Objective Function – Approximations

$$\begin{aligned}\varphi(\boldsymbol{a}) &= \min_{\boldsymbol{x}} \ \tfrac{1}{2}\left\|\boldsymbol{y} - \boldsymbol{a} \circledast \boldsymbol{x}\right\|_F^2 + \lambda \left\|\boldsymbol{x}\right\|_1 \\ &= \min_{\boldsymbol{x}} \ \underbrace{\tfrac{1}{2}\left\|\boldsymbol{y}\right\|_F^2}_{\text{constant}} - \langle \boldsymbol{a} \circledast \boldsymbol{x}, \boldsymbol{y}\rangle + \tfrac{1}{2}\left\|\boldsymbol{a} \circledast \boldsymbol{x}\right\|_F^2 + \lambda \left\|\boldsymbol{x}\right\|_1\end{aligned}$$

Objective Function – Approximations

$$
\begin{aligned}
\varphi(\boldsymbol{a}) &= \min_{\boldsymbol{x}} \ \tfrac{1}{2}\left\|\boldsymbol{y} - \boldsymbol{a} \circledast \boldsymbol{x}\right\|_F^2 + \lambda \left\|\boldsymbol{x}\right\|_1 \\
&= \min_{\boldsymbol{x}} \ \underbrace{\tfrac{1}{2}\left\|\boldsymbol{y}\right\|_F^2}_{\text{constant}} - \langle \boldsymbol{a} \circledast \boldsymbol{x}, \boldsymbol{y}\rangle + \tfrac{1}{2}\left\|\boldsymbol{a} \circledast \boldsymbol{x}\right\|_F^2 + \lambda \left\|\boldsymbol{x}\right\|_1
\end{aligned}
$$

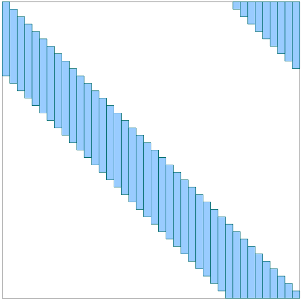


Figure 4: $C_{\boldsymbol{a}} \in \mathbb{R}^{m \times m}$

$$
\begin{aligned}
\boldsymbol{a} \circledast \boldsymbol{x} &= \boldsymbol{C_a x} \\
\left\|\boldsymbol{a} \circledast \boldsymbol{x}\right\|_F^2 &= \boldsymbol{x}^T \boldsymbol{C}_{\boldsymbol{a}}^T \boldsymbol{C_a x}
\end{aligned}
$$

$$\begin{aligned}\varphi(\boldsymbol{a}) &= \min_{\boldsymbol{x}} \tfrac{1}{2}\left\|\boldsymbol{y}-\boldsymbol{a}\circledast\boldsymbol{x}\right\|_F^2+\lambda\left\|\boldsymbol{x}\right\|_1\\ &= \min_{\boldsymbol{x}} \underbrace{\tfrac{1}{2}\left\|\boldsymbol{y}\right\|_F^2}_{\text{constant}}-\langle\boldsymbol{a}\circledast\boldsymbol{x},\boldsymbol{y}\rangle+\tfrac{1}{2}\left\|\boldsymbol{a}\circledast\boldsymbol{x}\right\|_F^2+\lambda\left\|\boldsymbol{x}\right\|_1\end{aligned}$$

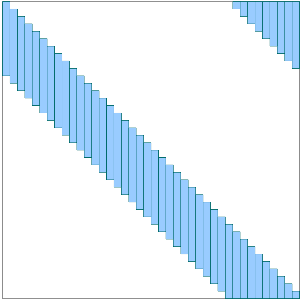


Figure 4: $\boldsymbol{C_a}\in\mathbb{R}^{m\times m}$

$$\begin{aligned}\boldsymbol{a}\circledast\boldsymbol{x} &= \boldsymbol{C_a x}\\ \left\|\boldsymbol{a}\circledast\boldsymbol{x}\right\|_F^2 &= \boldsymbol{x}^T\boldsymbol{C_a}^T\boldsymbol{C_a x}\end{aligned}$$

$$\begin{aligned}\operatorname{diag}\left(\boldsymbol{C_a}^T\boldsymbol{C_a}\right) &= \boldsymbol{1}\\ \boldsymbol{C_a}^T\boldsymbol{C_a}(i,j) &= \langle s_i[\boldsymbol{a}],s_j[\boldsymbol{a}]\rangle\end{aligned}$$

Objective Function – Approximations

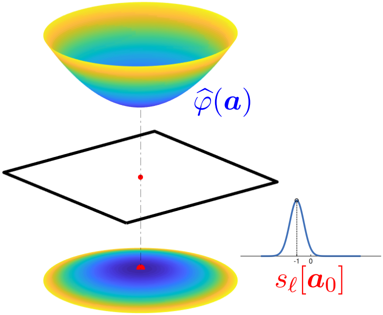
$$\varphi(\boldsymbol{a}) = \min_{\boldsymbol{x}} \left\{ \underbrace{\frac{1}{2} \ \|\boldsymbol{a} \circledast \boldsymbol{x} - \boldsymbol{y}\|_F^2}_{\text{data fidelity}} + \underbrace{\lambda \ \|\boldsymbol{x}\|_1}_{\text{sparsity}} \right\}$$

$$\approx \widehat{\varphi}(\boldsymbol{a}) = \min_{\boldsymbol{x}} \left\{ \underbrace{\frac{1}{2} \ \|\boldsymbol{x}\|_F^2 - \langle \boldsymbol{a} \circledast \boldsymbol{x}, \boldsymbol{y} \rangle + \frac{1}{2} \|\boldsymbol{y}\|_F^2}_{\text{approximating } \mathcal{C}_{\boldsymbol{a}}^* \mathcal{C}_{\boldsymbol{a}} \approx \mathcal{I}} + \underbrace{\lambda \ \|\boldsymbol{x}\|_1}_{\text{sparsity}} \right\}$$

Simplified Lasso:

$$\min \ \widehat{\varphi}(\boldsymbol{a}) \quad \text{s.t.} \quad \|\boldsymbol{a}\|_F = 1.$$

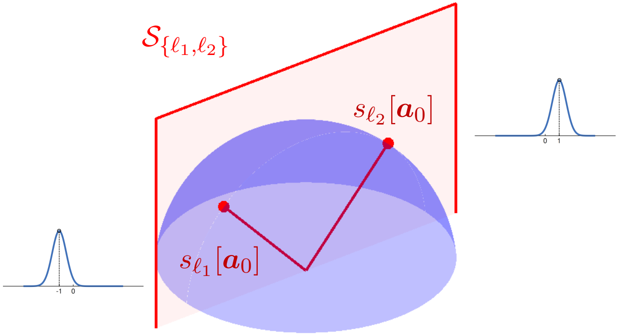
Objective Function – Near One Shift



$$\mathbb{S}^{p-1} \cap \left\{ \boldsymbol{a} \in \mathbb{S}^{p-1} \mid \|\boldsymbol{a} - s_\ell[\boldsymbol{a}_0]\|_2 \le r \right\}$$

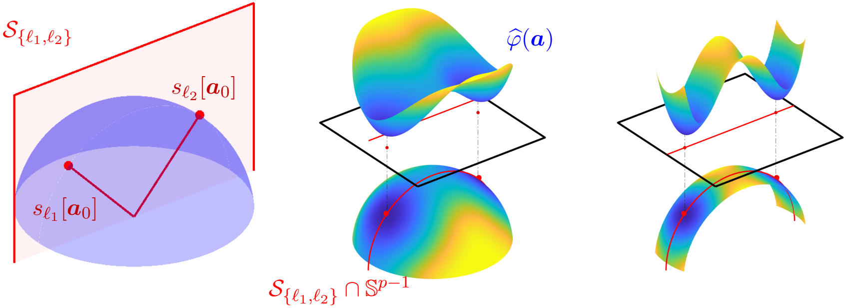
Objective function is **strongly convex** near a shift $s_\ell[\boldsymbol{a}_0]$ of the ground truth.

Objective Function – Linear Span of Two Shifts



Subspace $\mathcal{S}_{\ell_1,\ell_2} = \{\alpha_{\ell_1} s_{\ell_1}[\boldsymbol{a}_0] + \alpha_{\ell_2} s_{\ell_2}[\boldsymbol{a}_0] \mid \alpha_{\ell_1}, \alpha_{\ell_2} \in \mathbb{R}\}\,.$

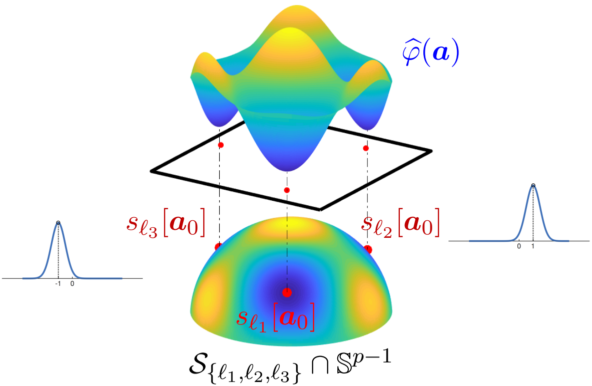
Objective Function – Linear Span of Two Shifts



Local minimizers are near signed shifts $\pm s_\ell[\boldsymbol{a}_0]$.

Negative curvature between two shifts $s_{\ell_1}[\boldsymbol{a}_0]$, $s_{\ell_2}[\boldsymbol{a}_0]$.

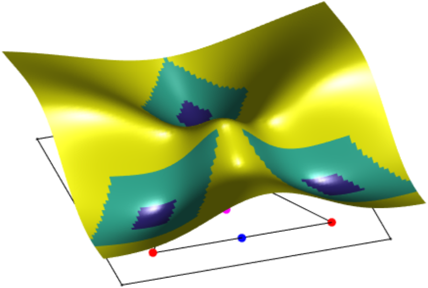
Objective Function – Multiple Shifts



Objective $\widehat{\varphi}$ over the linear span $\mathcal{S}_{\ell_1,\ell_2,\ell_3} = \left\{ \sum_{i=1}^{3} \alpha_{\ell_i} s_{\ell_i}[\boldsymbol{a}_0] \right\}$

Local minimizers are near signed shifts $\pm s_{\ell_i}[\boldsymbol{a}_0]$.

Objective function – Three Regions

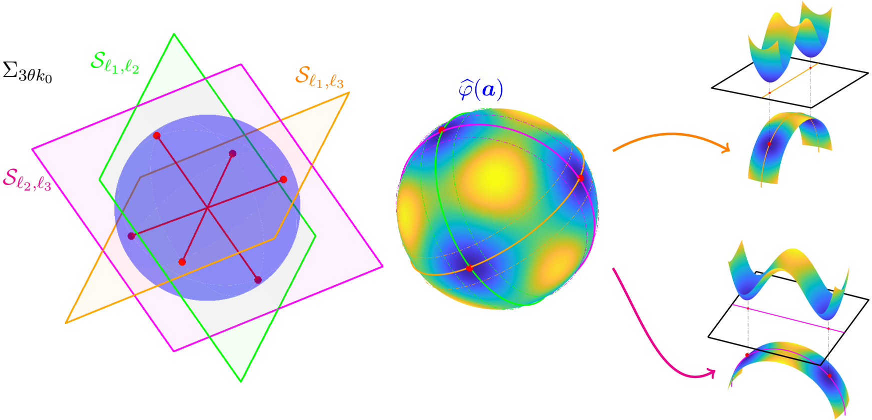


Negative curvature
Nonzero gradient
Strong convexity

The function $\widehat{\varphi}$ is **strict saddle**. At every point in the space, there is either a **negative curvature**, **strong gradient**, or **strong convexity** in the vicinity of a minimizer.

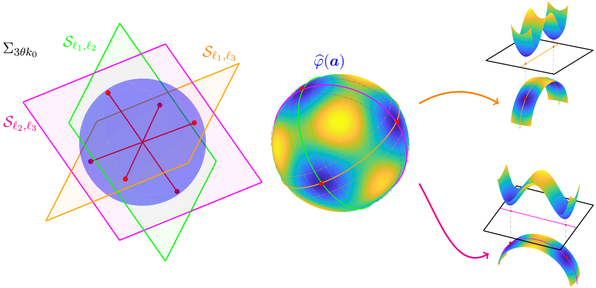
$\Longrightarrow$ a variety of methods efficiently find minimizers.

Objective Function – on a Union of Subspaces



Objective φ_ρ is "benign" over every subspace $\mathcal{S}_{\boldsymbol{\tau}}$ spanned by just a few shifts $\boldsymbol{\tau}$.

Objective Function – on a Union of Subspaces

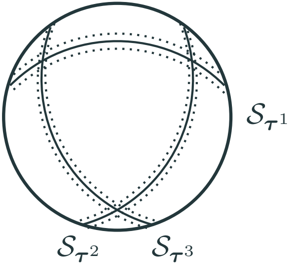


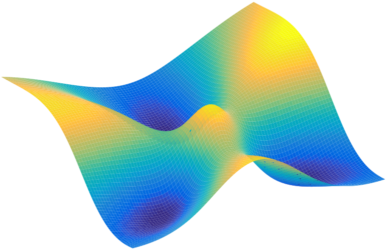
Theorem *When* $m > Ck^{4.5}\log k$ *and*

$$\frac{1}{k} < \theta < \frac{c}{(\sqrt{k} + k\mu^{1/2})\log k},$$

with high probability every local minimizer of φ_ρ *over* $\Sigma_{4\theta k_0}$ *is within distance* $C\mu\sqrt{1+k\mu} \times \theta^2 k^{3/2}$ *of some* $\pm s_\ell[\boldsymbol{a}_0]$.

Visually

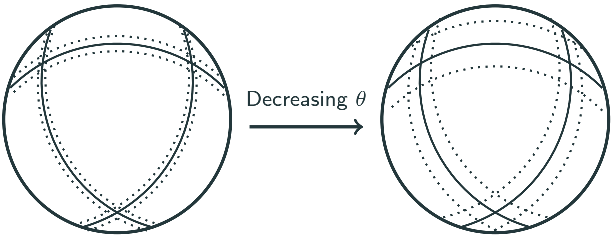




On each $\mathcal{S}_{\boldsymbol{\tau}}$

Globalization

Results characterize $\widehat{\varphi}(\boldsymbol{a})$ near a union of subspaces $\mathcal{S}_{\boldsymbol{\tau}}$:

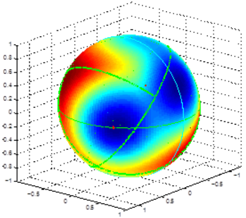


Can **globalize** in the "dilute limit" $\theta \searrow 0$.

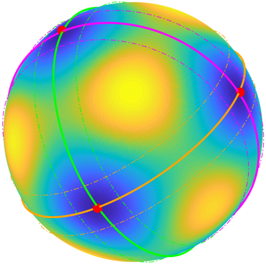
[Zhang, Lau, Kuo, Cheung, Pasupathy, Wright '17].

Main challenge for larger θ:

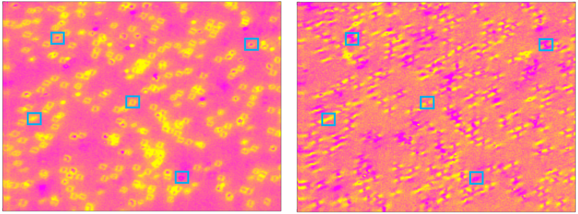
order-chaos boundary.



Algorithmic Implications I – easy to start near a few shifts

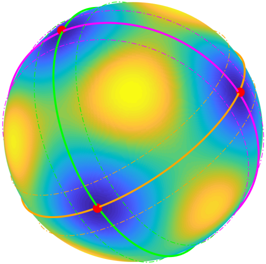


Good geometry near a few shifts

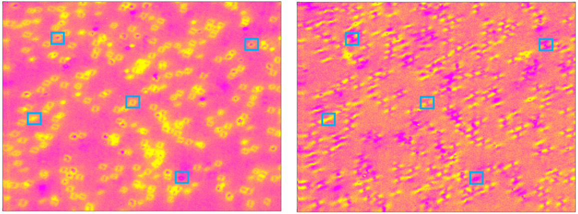


Data are a few shifts

Algorithmic Implications I – easy to start near a few shifts



Good geometry near a few shifts

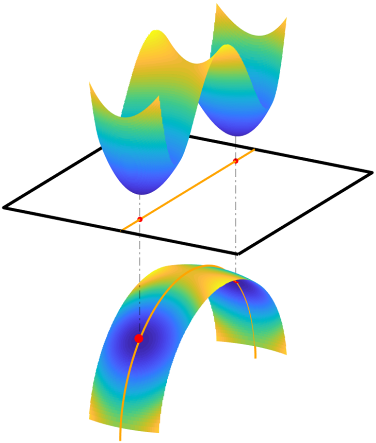


Data are a few shifts

Initialization: $\tilde{a}_{\text{init}} = \mathcal{P}_{\mathbb{S}}[\boldsymbol{y}_i, \boldsymbol{y}_{i+1}, \cdots, \boldsymbol{y}_{i+k-1}]^*$ is a superposition of about $2\theta k$ shifts of $\boldsymbol{a}_0$.

Zero pad to length $K = 3k-2$, set $\boldsymbol{a}_{\text{init}} = -\mathcal{P}_{\mathbb{S}^{K-1}} \nabla \widehat{\varphi}\,(\tilde{\boldsymbol{a}}_{\text{init}})$.

Algorithmic Implications II – easy to stay near a few shifts



Objective $\widehat{\varphi}$ grows away from $\mathcal{S}_{\boldsymbol{\tau}}$.

$\Longrightarrow$ Small stepping descent methods **stay** near this set.

Main Algorithmic Result (sketch)

Data-driven initialization:
$\boldsymbol{a}^{(0)} = \mathcal{P}_{\mathbb{S}} \nabla \widehat{\varphi}(\mathcal{P}_{\mathbb{S}}[0, \dots, 0, \boldsymbol{y}_0, \dots, \boldsymbol{y}_{k-1}, 0, \dots, 0])$

Minimization: of $\widehat{\varphi}$ over $\mathbb{S}^{3k-2}$ starting from $\boldsymbol{a}^{(0)}$ using small-stepping curvilinear search.

Rounding: to an exact solution $(\widehat{\boldsymbol{a}}, \widehat{\boldsymbol{x}})$ by locally minimizing

$$(\boldsymbol{a}, \boldsymbol{x}) = \tfrac{1}{2} \left\| \boldsymbol{a} * \boldsymbol{x} - \boldsymbol{y} \right\|_2^2 + \lambda \left\| \boldsymbol{x} \right\|_1 .$$

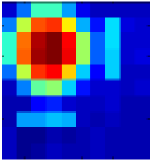
Theorem (sketch) *When* $C k^{4.5} \log k < m < c \exp(\theta k)$ *and*

$$\frac{1}{k} < \theta < \frac{c}{(\sqrt{k} + k\mu) \log k},$$

with high probability $(\widehat{\boldsymbol{a}}, \widehat{\boldsymbol{x}}) = \pm(s_\ell[\boldsymbol{a}_0], s_{-\ell}[\boldsymbol{x}_0])$ *for some* ℓ.

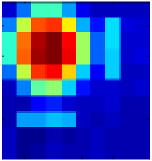
Sphere vs. Simplex Constraint

Imposing a sphere constraint for $\boldsymbol{a}_0$ leads to benign global geometry:
local minima are near signed shift truncations.



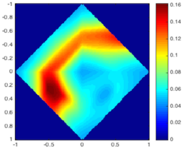
Sphere vs. Simplex Constraint

Imposing a sphere constraint for $\boldsymbol{a}_0$ leads to benign global geometry:
local minima are near signed shift truncations.



In image deblurring, a simplex constraint for $\boldsymbol{a}_0$ is natural model for the blurring process due to camera shake.

$$\begin{array}{ll} \min_{\boldsymbol{a},\boldsymbol{x}} & \frac{1}{2}\left\|\boldsymbol{y}-\boldsymbol{a}\circledast\boldsymbol{x}\right\|_F^2+\lambda\left\|\boldsymbol{x}\right\|_1 \\ \mathrm{s.\,t.} & \boldsymbol{a}\geq 0, \left\|\boldsymbol{a}\right\|_1=1 \end{array}$$

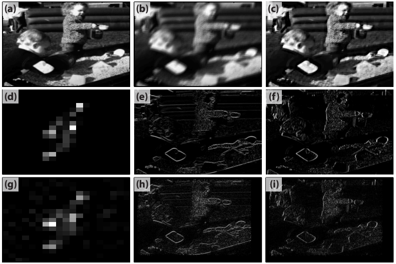


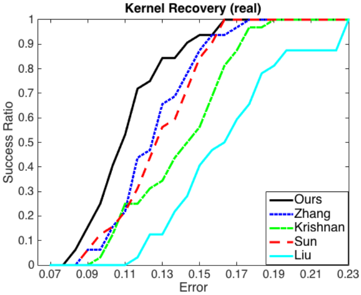
... but has spurious minimizers at spikes $\boldsymbol{a}=\boldsymbol{\delta}$.

e.g. Levin, Gribonval, Wipf.

Image Deblurring

Sparsity + benign geometry ⇒ surprisingly competitive performance.





Surprisingly competitive performance with a relatively simple idea – optimize over the sphere, tailor the algorithm to the geometry.

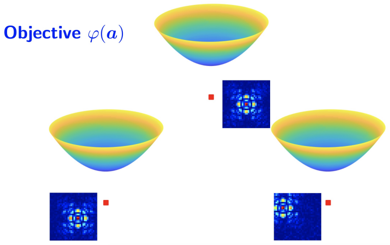
General Formulation

$$\begin{aligned} \min_{\boldsymbol{a}} \quad & \varphi(\boldsymbol{a}) \doteq \min_{\boldsymbol{x}} \frac{1}{p} \left\| \boldsymbol{y} - \boldsymbol{a} \circledast \boldsymbol{x} \right\|_p^p + \lambda \left\| \boldsymbol{x} \right\|_1 \\ \text{s. t.} \quad & \left\| \boldsymbol{a} \right\|_q = 1 \end{aligned}$$

General Formulation

$$\begin{aligned} \min_{\boldsymbol{a}} \quad & \varphi(\boldsymbol{a}) \doteq \min_{\boldsymbol{x}} \frac{1}{p} \|\boldsymbol{y} - \boldsymbol{a} \circledast \boldsymbol{x}\|_p^p + \lambda \|\boldsymbol{x}\|_1 \\ \text{s.t.} \quad & \|\boldsymbol{a}\|_q = 1 \end{aligned}$$

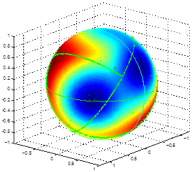
If $p = q \geq 2$, shift truncations of $\boldsymbol{a}_0$ persist as local minimizers.

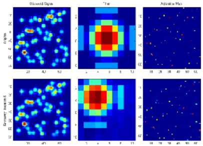


Summary

In certain region of the sphere,

- all local optima are near shift truncations of the ground truth;
- there exist reliable and efficient algorithms recovering the ground truth.

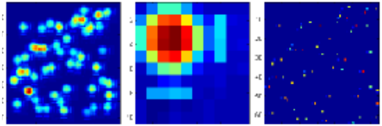




Geometry Inspired Algorithm

Phase I finds one local minimizer by solving

$$\boldsymbol{a}_*^{(0)} = \arg \min_{\|\boldsymbol{a}\|_F=1} \varphi_{\lambda_0}(\boldsymbol{a})$$

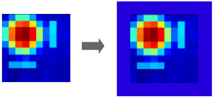


- with a random or sample-based initialization;
- with a reasonably large λ_0 to encourage sparsity.

Geometry Inspired Algorithm

Phase II tries to recover the global minimizer from the local minimizer generated via phase I:

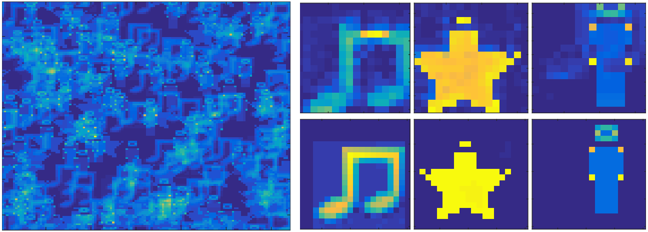
- **Zero-pad** $\boldsymbol{a}_*^{(0)}$ to $\boldsymbol{a}_*^{(1)}$;



- **Continuation:** Repeat solving $\boldsymbol{a}_*^{(k+1)} = \arg\min \varphi_{\lambda_{k+1}}(\boldsymbol{a})$ with $\lambda_{k+1} = \lambda_k/\beta$ and initialization $\boldsymbol{a}_*^{(k)}$ until $\lambda_k < \lambda_{end}$.

Geometry Inspired Algorithm

$$\begin{aligned} \min \varphi(\boldsymbol{a}) \quad &\doteq \quad \min_{\boldsymbol{x}} \ \frac{1}{2} \left\| \boldsymbol{y} - \sum_{n=1}^{N} \boldsymbol{a}_n \circledast \boldsymbol{x}_n \right\|_F^2 + \sum_{n=1}^{N} \lambda \left\| \boldsymbol{x}_n \right\|_1 \\ \text{s. t.} \quad & \quad \boldsymbol{a}_n \in \mathbb{S}^{k-1} \end{aligned}$$



Local minima $\bar{a}$ are near signed shift-truncations of a_0.

Comparison with the (Recent) Deconvolution Literature

Comparison with the (Recent) Deconvolution Literature

Random subspace model:

ala [Ahmed, Recht, Romberg '12]

Sign symmetry, no shift symmetry.

Topologically $\approx$ generalized phase retrieval.

$$\boldsymbol{y} = \underbrace{\left(\quad \boldsymbol{v} \right)}_{\boldsymbol{a}} \circledast \underbrace{\left(\quad \boldsymbol{w} \right)}_{\boldsymbol{x}}$$

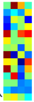
Comparison with the (Recent) Deconvolution Literature

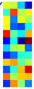
Random subspace model:

ala [Ahmed, Recht, Romberg '12]

Sign symmetry, no shift symmetry.

Topologically $\approx$ generalized phase retrieval.





$$\boldsymbol{y} = \underbrace{\left(\quad \boldsymbol{v} \right)}_{\boldsymbol{a}} \circledast \underbrace{\left(\quad \boldsymbol{w} \right)}_{\boldsymbol{x}}$$

Challenges for **SHORT-AND-SPARSE**:

Simultaneous structures: natural SDP relaxations break down.

Can't Avoid Symmetry: objective topology more complicated.

But ... very natural model for motif finding, image deblurring, . . .