Sparse Approximation: from Image Restoration to High Dimensional Classification

Bin Dong

Beijing International Center for Mathematical Research
Beijing Institute of Big Data Research
Peking University

Workshop on Frame Theory and Sparse Representation for Complex Data

Institute of Mathematical Science

National University of Singapore

29 May - 2 June 2017

Outlines

I. Brief review of image restoration models

II. Wavelet frame transforms and differential operators under variational and PDE framework

III. Sparse approximation for high-dimensional data classification

IV. Conclusions and Future work

Image Restoration Model

- Image Restoration Problems

$$f = Au + \eta$$

  - Denoising, when A is identity operator
  - Deblurring, when A is some blurring operator
  - Inpainting, when A is some restriction operator
  - CT/MR Imaging, when A is partial Radon/Fourier transform

- Chal ... ill-posed

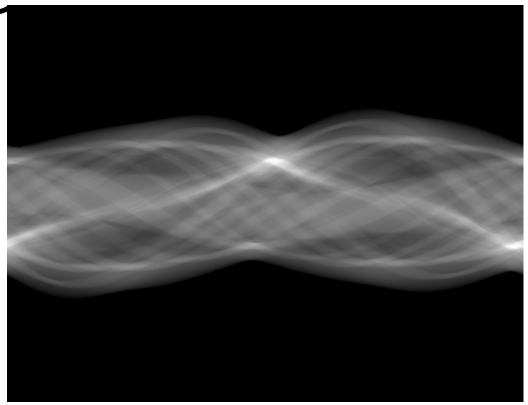


Image Restoration Models: A Quick Review

- Image restoration: $f = Au + \eta$
- Variational and Optimization Models

$$\min_u \lambda R(u) + \|Au - f\|^2$$

  - Total variation (TV) and generalizations: $R(u) = \|\nabla u\|_1 \quad \text{or} \quad \|\boldsymbol{D}u\|_1$
  - Wavelet frame based: $R(u) = \|Wu\|_1 \quad \text{or} \quad \|Wu\|_0$
  - Others: total generalized variation, low rank, NLM, BM3D, dictionary learning, etc.
- PDEs and Iterative Algorithms
  - Perona-Malik equation, shock-filtering (Rudin & Osher), etc

$$u_t = \sum_{\ell=1}^{L} \frac{\partial^{\boldsymbol{\alpha}_\ell}}{\partial x^{\boldsymbol{\alpha}_\ell}} \Phi_\ell(\boldsymbol{D}u, u) - A^*(Au - f), \qquad \text{with } \boldsymbol{D} = \left(\frac{\partial^{\boldsymbol{\beta}_1}}{\partial x^{\boldsymbol{\beta}_1}}, \ldots, \frac{\partial^{\boldsymbol{\beta}_L}}{\partial x^{\boldsymbol{\beta}_L}}\right)$$

  - Iterative shrinkage algorithm

$$\boldsymbol{u}^k = \widetilde{\boldsymbol{W}}^\top \boldsymbol{S}_{\boldsymbol{\alpha}^{k-1}}(\boldsymbol{W}\boldsymbol{u}^{k-1}) - \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{u}^{k-1} - \boldsymbol{f}), \quad k = 1, 2, \cdots$$

- What do they have in common?

“Dong and Shen, *Image restoration: a data-driven perspective*, Proceedings of the International Congress of Industrial and Applied Mathematics (ICIAM), 2015”

<u>Shrinkage</u> in <u>sparse</u> domain under <u>transformation</u>!

Bridging discrete and continuum

WAVELET FRAME TRANSFORMS AND DIFFERENTIAL OPERATORS

❖ J. Cai, B. Dong, S. Osher and Z. Shen, *Image restoration: total variation; wavelet frames; and beyond*, **Journal of American Mathematical Society**, 25(4), 1033-1089, 2012.

❖ B. Dong, Q. Jiang and Z. Shen, *Image restoration: wavelet frame shrinkage, nonlinear evolution PDEs, and beyond*, MMS, 15(1)， 606-660, 2017.

❖ Jian-Feng Cai, B. Dong and Zuowei Shen, *Image restorations: a wavelet frame based model for piecewise smooth functions and beyond*, **Applied and Computational Harmonic Analysis**, 41(1), 94-138, 2016.

❖ Bin Dong, Zuowei Shen and Peichu Xie, Image restoration: a general wavelet frame based model and its asymptotic analysis, **SIAM Journal on Mathematical Analysis**, 49(1), 421-445, 2017.

MRA-Based Tight Wavelet Frames

- Refinable and wavelet functions

$$\phi = 2^d \sum \boldsymbol{a}_0[\boldsymbol{k}]\phi(2\cdot - \boldsymbol{k}) \quad \psi_\ell = 2^d \sum \boldsymbol{a}_\ell[\boldsymbol{k}]\phi(2\cdot - \boldsymbol{k}), \quad \ell = 1, 2, \ldots, q.$$

- Unitary extension principle (UEP)

$$\sum_{\ell=0}^{q} |\widehat{\boldsymbol{a}}_\ell(\xi)|^2 = 1 \quad \text{and} \quad \sum_{\ell=0}^{q} \widehat{\boldsymbol{a}}_\ell(\xi)\overline{\widehat{\boldsymbol{a}}_\ell(\xi+\nu)} = 0,$$

$$\nu \in \{0,\pi\}^d \setminus \{\boldsymbol{0}\} \text{ and } \xi \in [-\pi,\pi]^d$$

- Discrete 2D transformation: $\boldsymbol{W}\boldsymbol{u} = \{\boldsymbol{W}_{l,\boldsymbol{i}}\boldsymbol{u} : 0 \le l \le L-1, 0 \le i_1, i_2 \le r\}$

$$\boxed{\boldsymbol{W}_{l,\boldsymbol{i}}\boldsymbol{u} := \boldsymbol{a}_{l,\boldsymbol{i}}[-\cdot] \circledast \boldsymbol{u},}$$

$$\boldsymbol{a}_{\boldsymbol{i}}[\boldsymbol{k}] := \boldsymbol{a}_{i_1}[k_1]\boldsymbol{a}_{i_2}[k_2], \quad 0 \le i_1, i_2 \le r; \ (k_1,k_2) \in \mathbb{Z}^2.$$

$$\boldsymbol{a}_{l,\boldsymbol{i}} = \tilde{\boldsymbol{a}}_{l,\boldsymbol{i}} \circledast \tilde{\boldsymbol{a}}_{l-1,0} \circledast \ldots \circledast \tilde{\boldsymbol{a}}_{0,0} \quad \text{with} \quad \tilde{\boldsymbol{a}}_{l,\boldsymbol{i}}[\boldsymbol{k}] = \begin{cases} \boldsymbol{a}_{\boldsymbol{i}}[2^{-l}\boldsymbol{k}], & \boldsymbol{k} \in 2^l\mathbb{Z}^2; \\ 0, & \boldsymbol{k} \notin 2^l\mathbb{Z}^2. \end{cases}$$

- Perfect reconstruction: $\boldsymbol{W}^\top \boldsymbol{W} = \boldsymbol{I}$

- Further reading: **[Dong and Shen, *MRA-Based Wavelet Frames and Applications*, IAS Lecture Notes Series,2011]**

Connections: Motivation

- Difference operators in wavelet frame transform:

Haar Filters $h_{0,1} = \frac{1}{4}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, h_{1,0} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, h_{1,1} = \frac{1}{4}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

Transform $Wu = \{h_{0,1}[-\cdot] \circledast u; h_{1,0}[-\cdot] \circledast u; h_{1,1}[-\cdot] \circledast u\}$

Approximation

$$h_{0,1}[-\cdot] \circledast u \approx \frac{1}{2}\delta u_x, \quad h_{1,0}[-\cdot] \circledast u \approx \frac{1}{2}\delta u_y, \quad h_{1,1}[-\cdot] \circledast u \approx \frac{1}{4}\delta^2 u_{xy}$$

- Thus, $\boxed{\frac{2}{\delta} Wu \approx \nabla u}$

$$|\nabla u| \approx \left(\frac{1}{4}\left[(D_x^+ u_{i,j})^2 + (D_x^+ u_{i,j+1})^2 + (D_y^+ u_{i,j})^2 + (D_y^+ u_{i+1,j})^2\right] + \left[\frac{(D_x^+ u_{i,j} + D_y^- u_{i,j+1})^2}{4} + \frac{(D_x^+ u_{i,j} + D_y^+ u_{i+1,j})^2}{4}\right]\right)^{\frac{1}{2}}$$

- More rigorously **[Choi, Dong and Zhang, preprint, 2017]**

Proposition 2.2. *Let a tensor product framelet function $\psi_{\boldsymbol{\alpha}} \in L_2(\mathbb{R}^2)$ have vanishing moments of order $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}| \le s$, and let $\mathrm{supp}(\psi_{\boldsymbol{\alpha}}) = [a_1, a_2] \times [b_1, b_2]$. For $n \in \mathbb{N}$ and $\boldsymbol{k} \in \mathbb{Z}^2$ with $\mathrm{supp}(\psi_{\boldsymbol{\alpha},n-1,\boldsymbol{k}}) \subseteq \overline{\Omega}$, we have*

$$\langle u, \psi_{\boldsymbol{\alpha},n-1,\boldsymbol{k}} \rangle = (-1)^{|\boldsymbol{\alpha}|} 2^{|\boldsymbol{\alpha}|(1-n)} \langle \partial^{\boldsymbol{\alpha}} u, \varphi_{\boldsymbol{\alpha},n-1,\boldsymbol{k}} \rangle$$

for every $u \in W_1^s(\Omega)$.

$\int_{\mathbb{R}^2} \varphi_{\boldsymbol{\alpha}} \mathrm{d}\boldsymbol{x} \neq 0, \quad \mathrm{supp}(\varphi_{\boldsymbol{\alpha}}) = \mathrm{supp}(\psi_{\boldsymbol{\alpha}})$

Connections: Analysis Based Model and Variational Model

- **[Cai, Dong, Osher and Shen, *JAMS*, 2012]**:

$$\|\lambda W u\|_1 + \frac{1}{2}\|Au - f\|_2^2 \xrightarrow{\text{Converges}} \lambda\|D(u)\|_1 + \frac{1}{2}\|Au - f\|_{L_2(\Omega)}^2$$

For any differential operator when proper parameter is chosen.

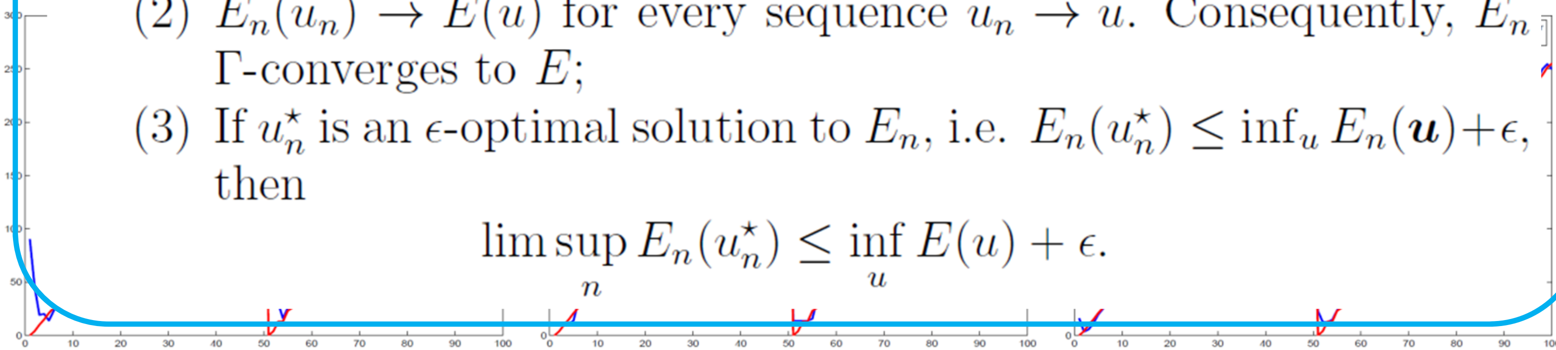
Theorem. Let the objective functionals of the analysis based model and the variational model be $E_n(u)$ and $E(u)$ respectively, then:

(1) $E_n(u) \to E(u)$ for each $u \in W_1^s(\Omega)$;

(2) $E_n(u_n) \to E(u)$ for every sequence $u_n \to u$. Consequently, E_n Γ-converges to E;

(3) If $u_n^\star$ is an ϵ-optimal solution to E_n, i.e. $E_n(u_n^\star) \le \inf_u E_n(\boldsymbol{u}) + \epsilon$, then

$$\limsup_n E_n(u_n^\star) \le \inf_u E(u) + \epsilon.$$



- Image segmentation: **[Dong, Chien and Shen, 2010]**
- Surface reconstruction from point clouds: **[Dong and Shen, 2011]**

Relations: Wavelet Shrinkage and Nonlinear PDEs

- **[Dong, Jiang and Shen, MMS, 2017]**

$$\boldsymbol{u}^k = \widetilde{\boldsymbol{W}}^\top \boldsymbol{S}_{\boldsymbol{\alpha}^{k-1}}(\boldsymbol{W}\boldsymbol{u}^{k-1}), \quad k = 1, 2, \cdots$$

$$u_t = \sum_{\ell=1}^{L} \frac{\partial^{\boldsymbol{\alpha}_\ell}}{\partial x^{\boldsymbol{\alpha}_\ell}} \Phi_\ell(\boldsymbol{D}u, u), \quad \text{with } \boldsymbol{D}u = \left(\frac{\partial^{\boldsymbol{\beta}_1}}{\partial x^{\boldsymbol{\beta}_1}}, \ldots, \frac{\partial^{\boldsymbol{\beta}_L}}{\partial x^{\boldsymbol{\beta}_L}}\right)$$

- Theoretical justification available for quasilinear parabolic equations.
- Lead to new PDE models such as:

$$u_{tt} + Cu_t = \sum_{\ell=1}^{L} (-1)^{1+|\boldsymbol{\beta}_\ell|} \frac{\partial^{\boldsymbol{\beta}_\ell}}{\partial \boldsymbol{x}^{\boldsymbol{\beta}_\ell}} \left[g_\ell\left(u, \frac{\partial^{\boldsymbol{\beta}_1} u}{\partial \boldsymbol{x}^{\boldsymbol{\beta}_1}}, \cdots, \frac{\partial^{\boldsymbol{\beta}_L} u}{\partial \boldsymbol{x}^{\boldsymbol{\beta}_L}}\right) \frac{\partial^{\boldsymbol{\beta}_\ell}}{\partial \boldsymbol{x}^{\boldsymbol{\beta}_\ell}} u \right] - \kappa A^\top (Au - f)$$

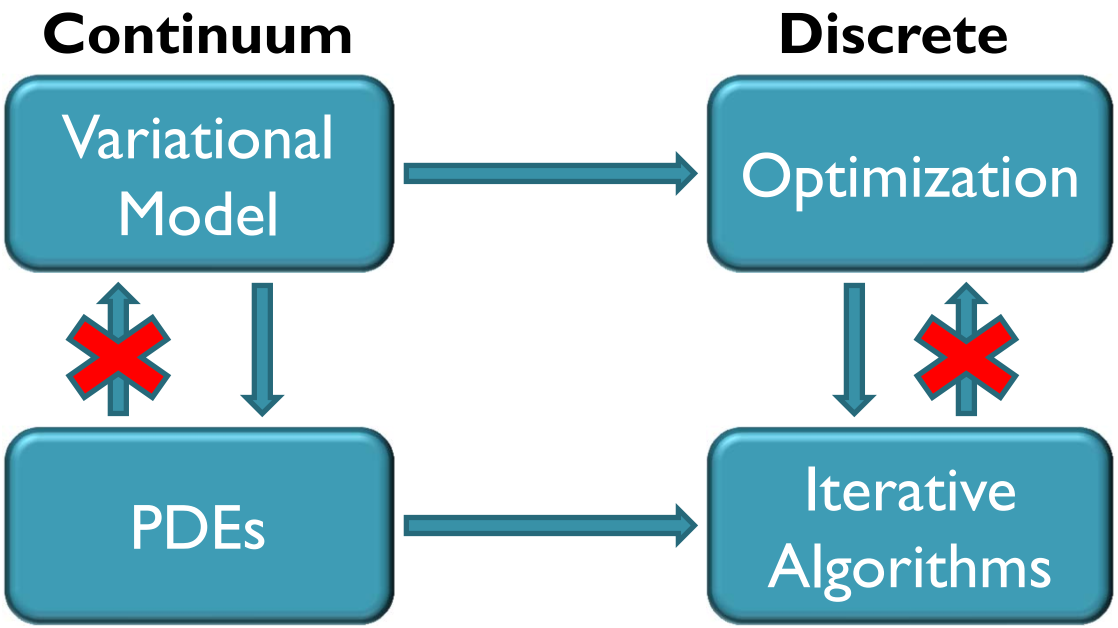
$$\boldsymbol{u}^k = (I - \mu \boldsymbol{A}^\top \boldsymbol{A}) \boldsymbol{W}^\top \boldsymbol{S}_{\boldsymbol{\alpha}^{k-1}}(\boldsymbol{W}\boldsymbol{u}^{k-1}) + \mu \boldsymbol{A}^\top \boldsymbol{f}$$

where

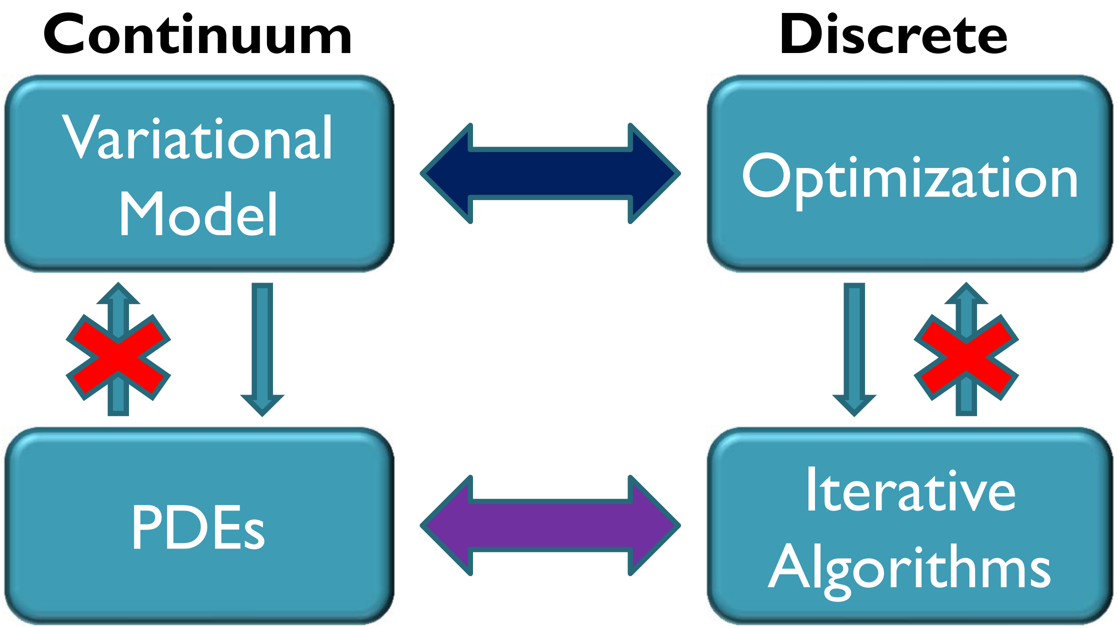
$$\boldsymbol{S}_{\boldsymbol{\alpha}^{k-1}}(\boldsymbol{W}\boldsymbol{u}^{k-1}) = \{S_{\alpha_{l,\ell,\boldsymbol{n}}(\boldsymbol{W}_l \boldsymbol{u}^{k-1})}(\boldsymbol{W}_l \boldsymbol{u}^{k-1}) : \ 0 \le l \le \text{Lev} - 1, 1 \le \ell \le L\}$$

$$S_{\boldsymbol{\alpha}_{\ell,\boldsymbol{n}}(\boldsymbol{d})}(d_{1,\boldsymbol{n}}, d_{2,\boldsymbol{n}}) = d_{\ell,\boldsymbol{n}} \left(1 - \frac{4\tau}{h^2} g\left(\frac{4(d_{1,\boldsymbol{n}})^2 + 4(d_{2,\boldsymbol{n}})^2}{h^2}\right)\right)$$

Summary


Continuum
Discrete
Variational Model
Optimization
PDEs
Iterative Algorithms

Summary



❖ J. Cai, B. Dong, S. Osher and Z. Shen, *Image restoration: total variation; wavelet frames; and beyond*, **Journal of American Mathematical Society**, 25(4), 1033-1089, 2012.
❖Jian-Feng Cai, B. Dong and Z. Shen, *Image restorations: a wavelet frame based model for piecewise smooth functions and beyond*, **Applied and Computational Harmonic Analysis**, 41(1), 94-138, 2016.
❖ B. Dong, Z. Shen and P. Xie, Image restoration: a general wavelet frame based model and its asymptotic analysis, **SIAM Journal on Mathematical Analysis**, 49(1), 421-445, 2017.

❖B. Dong, Q. Jiang and Z. Shen, *Image restoration: wavelet frame shrinkage, nonlinear evolution PDEs, and beyond*, **Multiscale Modeling & Simulation**, 15(1)， 606-660, 2017.

SPARSE APPROXIMATION IN HIGH-DIMENSIONAL DATA CLASSIFICATION

❖ Ning Hao, Bin Dong and Jianqing Fan, *Sparsifying the Fisher Linear Discriminant by Rotation*, **Journal of the Royal Statistical Society Series B**, 2015.

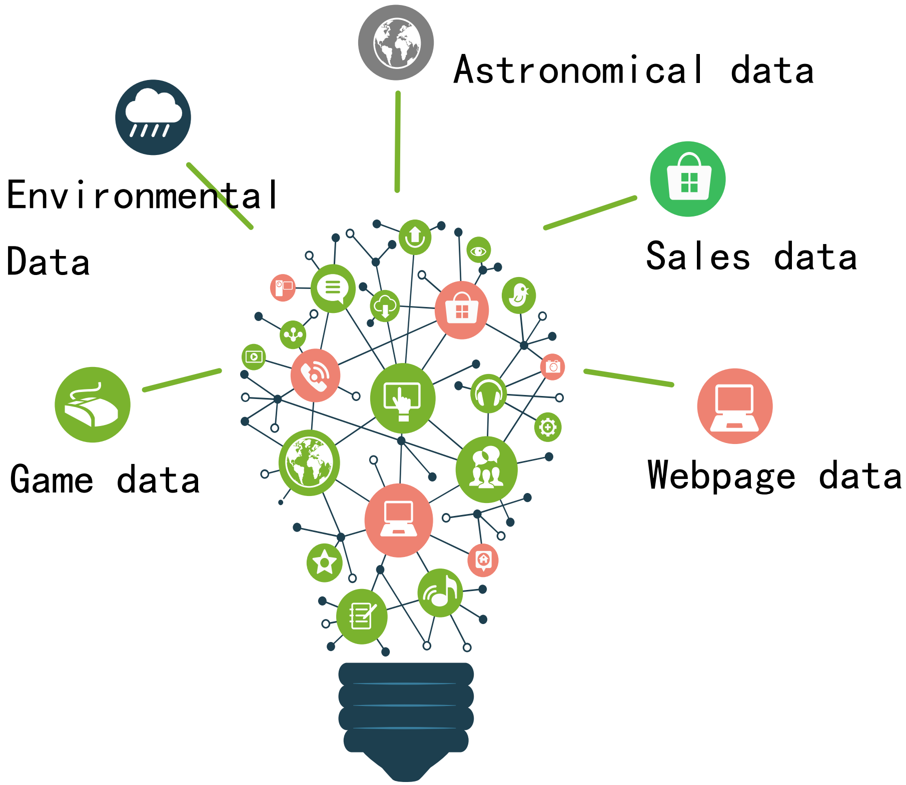
❖ Bin Dong, *Sparse Representation on Graphs by Tight Wavelet Frames and Applications*, **Applied and Computational Harmonic Analysis**, 2015.

❖ Bin Dong and Ning Hao, *Semi-supervised high dimensional clustering by tight wavelet frames*, **Proceedings of SPIE**, 2015.

Introduction

What is **data science**? Extracting knowledge from data to make intelligent observations and decisions.

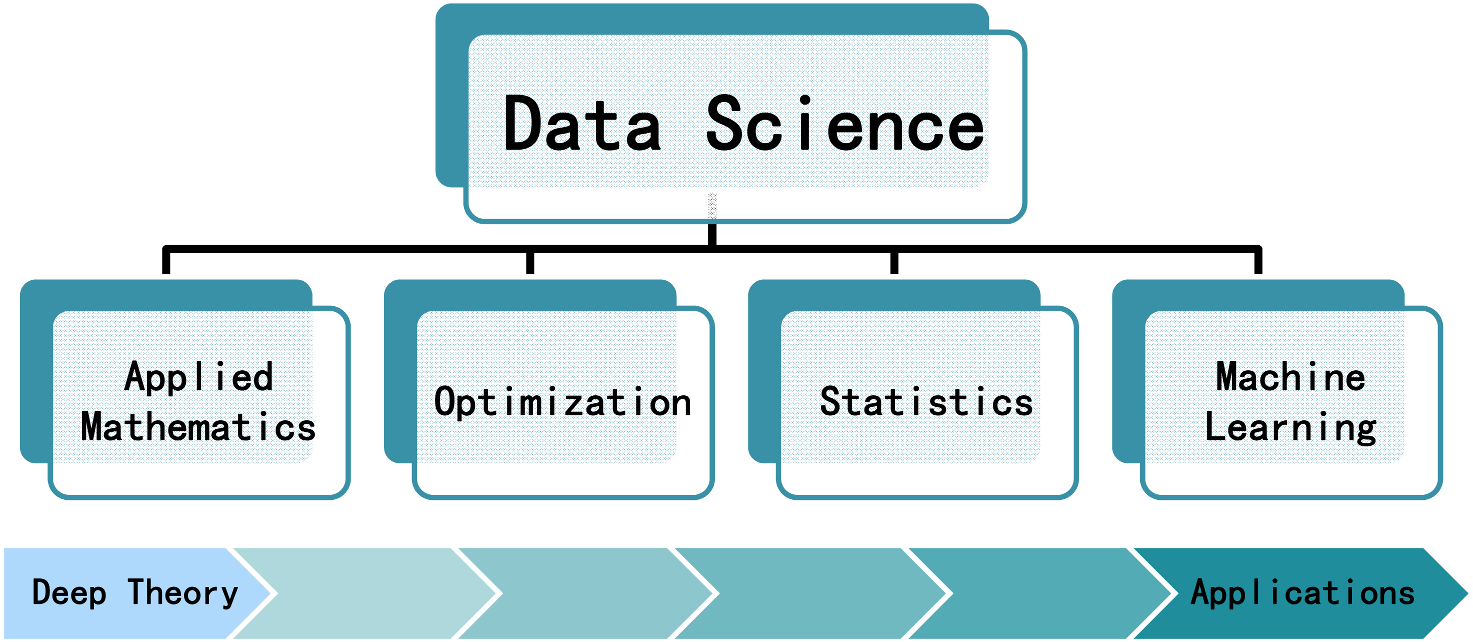
1. Broad applications



2. Variety

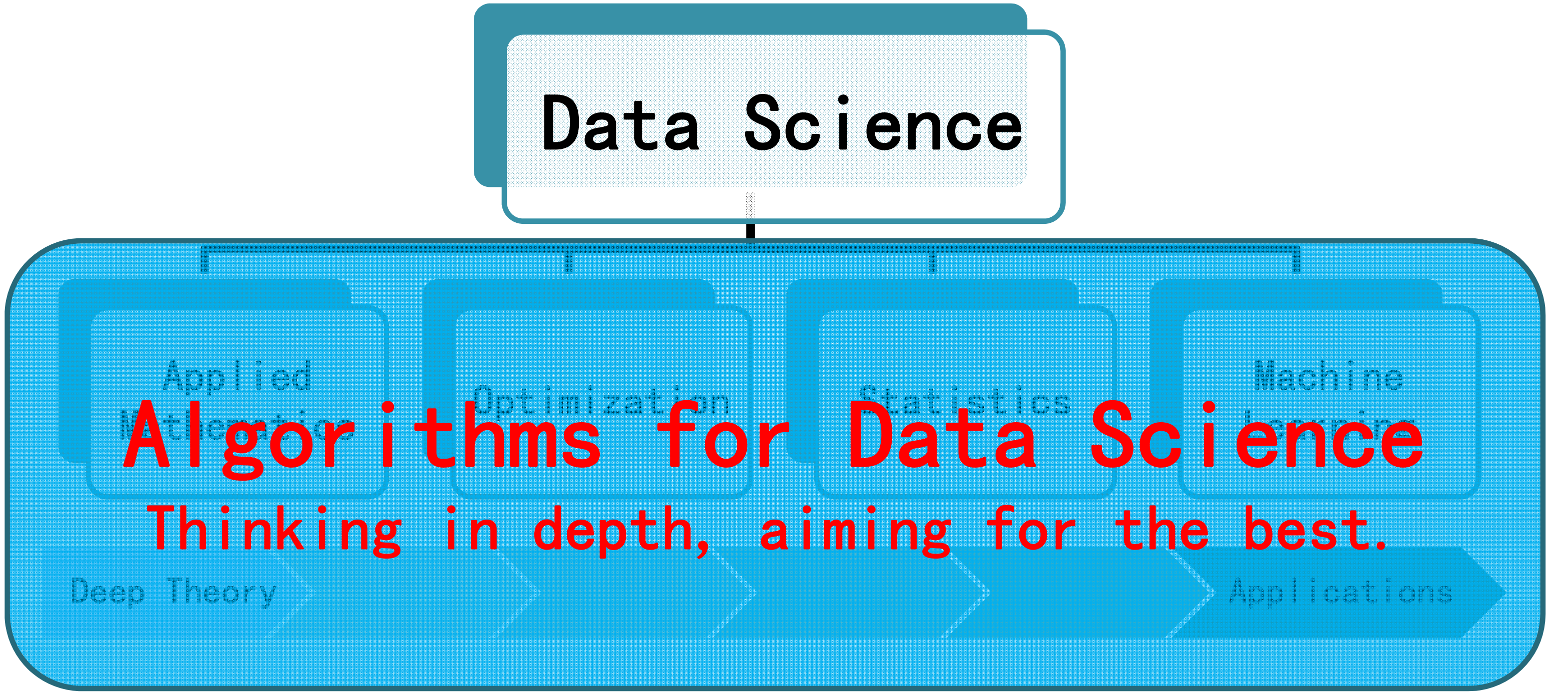
- Network data
- Webpage data
- Text
- Image
- Video
- Audio

Introduction



- Different area has different focus. Some has tight link with another.
- Broader links?

Introduction



Importance of the merge:

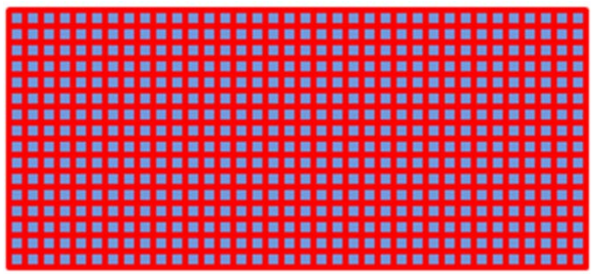
- Combining merits
- New insights on classical problems

Introduction

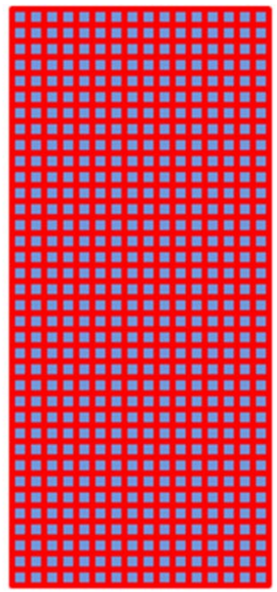
- Typical big data set: with $n \times p$ huge

$$X \in \mathbb{K}^{n \times p}, \quad \mathbb{K} = \mathbb{Z}, \mathbb{R}, \mathbb{C}, \text{ etc.}$$

- Classical v.s. modern



Classical: n<p



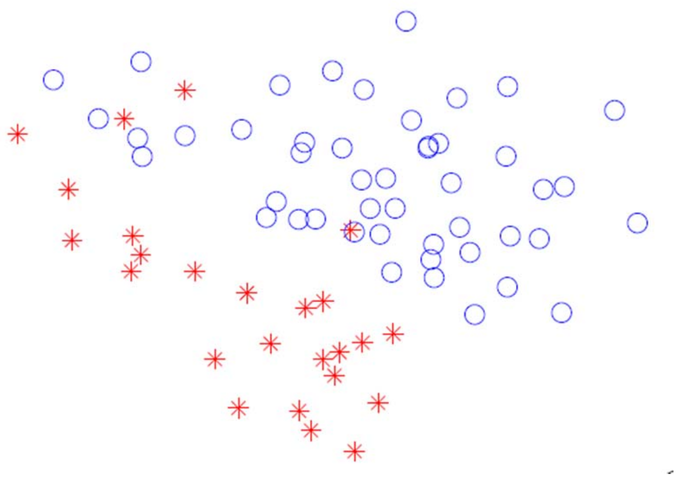
Modern: n>p

Introduction

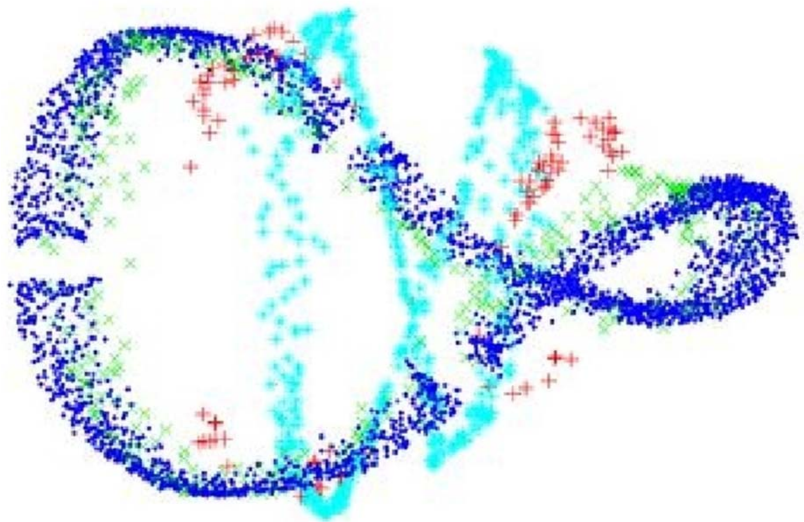
- Typical big data set: with $n \times p$ huge

$$X \in \mathbb{K}^{n \times p}, \quad \mathbb{K} = \mathbb{Z}, \mathbb{R}, \mathbb{C}, \text{ etc.}$$

- Classical v.s. modern



Classical: n<p



Modern: n>p

Introduction

- Sparsity is key
  - What is sparsity for general data sets?

    Essential information is of much lower dimension than the dimension of the data itself.

  - How do we harvest sparsity?

    Sparse under certain (nonlinear) transformation.

  - Examples:
    - PCA and its siblings
    - Low rank approximation
    - Wavelet frame transform
    - Dictionary learning
    - Isomaps, LLE, diffusion maps
    - Autoencoder
    - … etc.

Nonlinear Classification

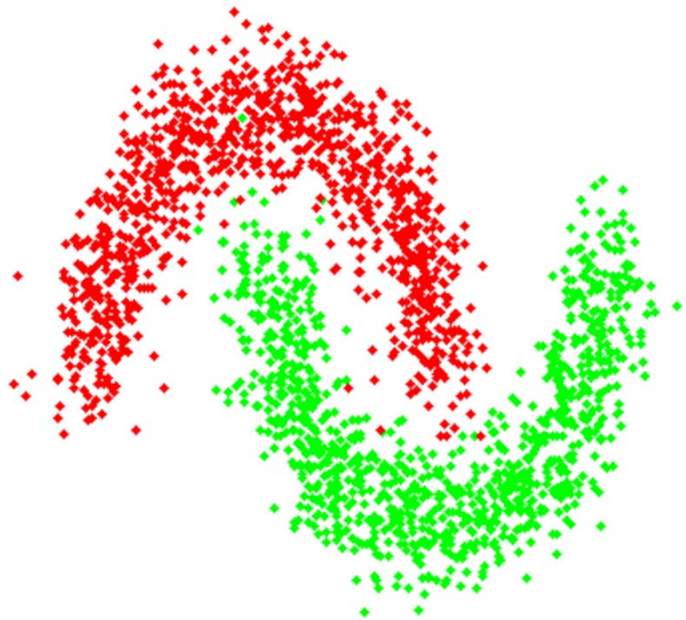
MODERN SCENARIO

❖ Bin Dong, *Sparse Representation on Graphs by Tight Wavelet Frames and Applications*, **Applied and Computational Harmonic Analysis**, 2015.

❖ Bin Dong and Ning Hao, *Semi-supervised high dimensional clustering by tight wavelet frames*, **Proceedings of SPIE**,2015.

Nonlinear Classifier

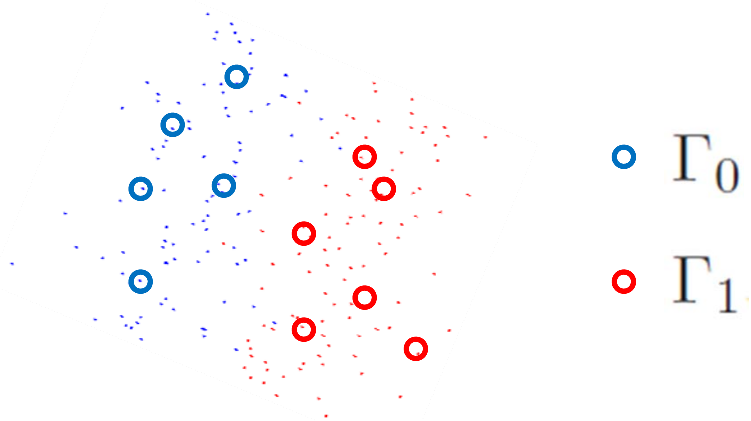
- When we have enough observations, nonlinear classifier leads to more accurate classification.



Explicit Classifier

PDE Method

- Ginzburg–Landau (GL) functional [Andrea and Flenner, 2011]



$$f[i] = \begin{cases} 0 & \text{for } i \in \Gamma_0, \\ 1 & \text{for } i \in \Gamma_1. \end{cases}$$

- GL model

$$E(u) = \frac{\epsilon}{2}\langle u, \mathcal{L}_s u\rangle + \frac{1}{4\epsilon}\|u^2 - 1\|_{2,G}^2 + \frac{\mu}{2}\|u_{|\Gamma} - f\|_{2,G}^2$$

where

$$\mathcal{L}_s = I - D^{-1/2} A D^{-1/2} \qquad \|f\|_{p,G} := \left(\sum_{k=1}^{K} |f[k]|^p d[k]\right)^{\frac{1}{p}}$$

PDE Method

- Splitting *E* to a sum of convex and concave parts

$$\boxed{E(u) = E_1(u) - E_2(u)}$$

$$E_1(u) = \frac{\epsilon}{2}\int |\nabla u(x)|^2 dx + \frac{c}{2}\int |u(x)|^2 dx,$$
$$E_2(u) = -\frac{1}{4\epsilon}\int (u(x)^2 - 1)^2 dx + \frac{c}{2}\int |u(x)|^2 dx - \int \frac{\lambda(x)}{2}(u(x) - u_0(x))^2 dx.$$

- Convex splitting scheme

$$\frac{u^{n+1} - u^n}{dt} = -\frac{\partial E_1}{\partial u}(u^{n+1}) + \frac{\partial E_2}{\partial u}(u^n)$$

❖ At each iteration, we need to solve a Laplace equation on graph.
❖ Fast graph Laplacian solver is needed, such as Nystrom's method.

Wavelet Frame Method

- Key idea: Eigenfunctions of Laplace-Beltrami operator (graph Laplacian in discrete setting) are understood as Fourier basis on manifolds (graphs in discrete setting) and the associated eigenvalues as frequency components.
- Spectrum of Laplace-Beltrami operator on $\{\mathcal{M}, g\}$

$$\Delta u + \lambda u = 0, \qquad u_{|S} = 0.$$

Eigenvalues and eigenfunctions: $0 < \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots$

$$\langle u_p, u_{p'} \rangle_{L_2(\mathcal{M})} = \int_{\mathcal{M}} u_p(x) u_{p'}^*(x) dx = \delta_{p,p'}$$

- Fourier transform $\widehat{f}[p] = \langle f, u_p \rangle_{L_2(\mathcal{M})}$
- Plancherel and Parseval's identities

$$\langle f, g \rangle_{L_2(\mathcal{M})} = \langle \widehat{f}, \widehat{g} \rangle_{\ell_2(\mathbb{Z}^+)} \quad \text{for } f, g \in L_2(\mathcal{M})$$
$$\|f\|^2_{L_2(\mathcal{M})} = \|\widehat{f}\|^2_{\ell_2(\mathbb{Z}^+)}.$$

Wavelet Frame Method

- Asymptotic properties of eigenfunctions and eigenvalues:
  - Weyl's asymptotic formula（1912）： $\lambda_p \asymp p^{\frac{2}{m}}$
  - Uniform bound (Grieser, 2002): $\|u_p\|_{L_\infty(\mathcal{M})} \leq C\lambda_p^{\frac{m-1}{4}}$

- Wavelet system (semi-continuous) on manifold $\mathcal{M}$:

$$X(\Psi) = \{\psi^{\mathcal{M}}_{j,n,y} \in L_2(\mathcal{M}) : \ 1 \leq j \leq r, n \in \mathbb{Z}, y \in \mathcal{M}\},$$

where $\psi^{\mathcal{M}}_{j,n,y} \in L_2(\mathcal{M})$ is generated by $\Psi = \{\psi_j : \ 1 \leq j \leq r\} \subset L_2(\mathbb{R})$ as

Eigenvalues Eigenfunctions

$$\psi^{\mathcal{M}}_{j,n,y}(x) = \sum_{p=0}^{\infty} \widehat{\psi}_j(2^{-n}\lambda_p)u_p^*(y)u_p(x), \quad \text{with } n \in \mathbb{Z}, \ x \in \mathcal{M}, \ y \in \mathcal{M},$$

Dilation Translation

where $\widehat{\psi}_j$ denotes that Fourier transform of $\psi_j \in L_2(\mathbb{R})$.

- Question: how to construct ψ_j so that $X(\Psi)$ is a tight frame on $\mathcal{M}$?

Wavelet Frame Method

- Further restriction on ψ_j :

Given $\widehat{\phi}(2\xi) = \widehat{a}(\xi)\widehat{\phi}(\xi)$ and $a_j \in \ell_0(\mathbb{Z})$

let $\widehat{\psi}_j(2\xi) := \widehat{a}_j(\xi)\widehat{\phi}(\xi), \quad 1 \leq j \leq r.$

- Question: how to construct a_j so that $X(\Psi)$ is a tight frame on $\mathcal{M}$?
- Benefits of such restriction
  - Grants a natural transition from continuum to discrete setting
  - Makes construction of tight frames on manifolds/graphs painless
  - Grants fast decomposition and reconstruction algorithms (Chebyshev polynomial approximation)

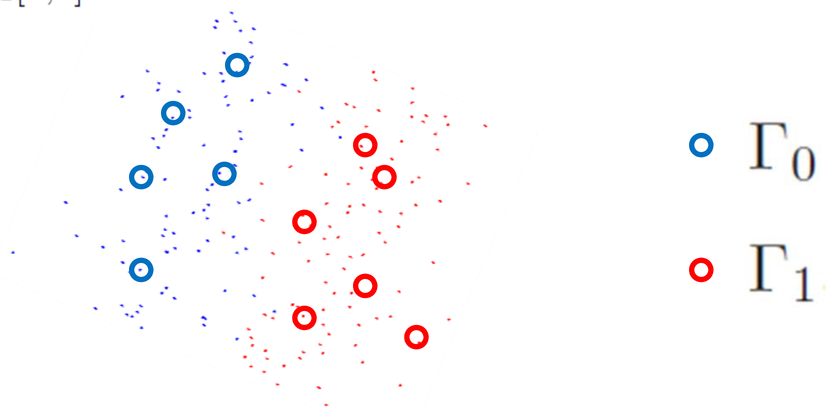
Wavelet Frame Method

- Sparsity based semi-supervised learning models

Model L2: $$\min_{u\in[0,1]} \|\boldsymbol{\nu}\cdot\boldsymbol{W}u\|_{1,G}+\frac{1}{2}\|u_{|\Gamma}-f\|_{2,G}^2,$$

Exact Model: $$\begin{array}{ll}\min_{u\in[0,1]} & \|\boldsymbol{\nu}\cdot\boldsymbol{W}u\|_{1,G} \\ \text{s.t.} & u_{|\Gamma}=f\end{array} \qquad f[i]=\begin{cases}0 & \text{for } i\in\Gamma_0,\\ 1 & \text{for } i\in\Gamma_1.\end{cases}$$

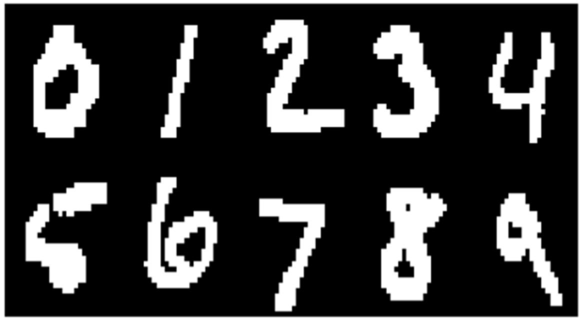
Robust Model: $$\min_{u\in[0,1]} \|\boldsymbol{\nu}\cdot\boldsymbol{W}u\|_{1,G}+\|u_{|\Gamma}-f\|_{1,G}$$



- Transform ***W*** is the fast tight wavelet frame transform on graphs **[B. Dong, ACHA, 2015]**.

Wavelet Frame Method

- Classification – Real Datasets
  - MNIST data set (http://yann.lecun.com/exdb/mnist/)
  - Banknote authentication dataset (UCI machine learning repository)



Machine Learning Repository

- Results: Model L2 **[Dong, ACHA, 2015]**

~3-5% label

Errors (%)	Our Method	Max-Flow	PAL	Binary MBO	GL
MNIST	2.76 (8.5 sec.)	1.52	1.56	1.64	1.75
Banknote	1.64 (2.9 sec.)	1.17	1.71	6.52	3.90

- Max-Flow & PAL : [Merkurjev, Bae, Bertozzi, and Tai, preprint, 2014]
- Binary MBO: [Merkurjev, Kostic, and Bertozzi, 2013]
- GL: [Bertozzi and Flenner, 2012]

Further Studies of Wavelet Frame Transform on Graphs

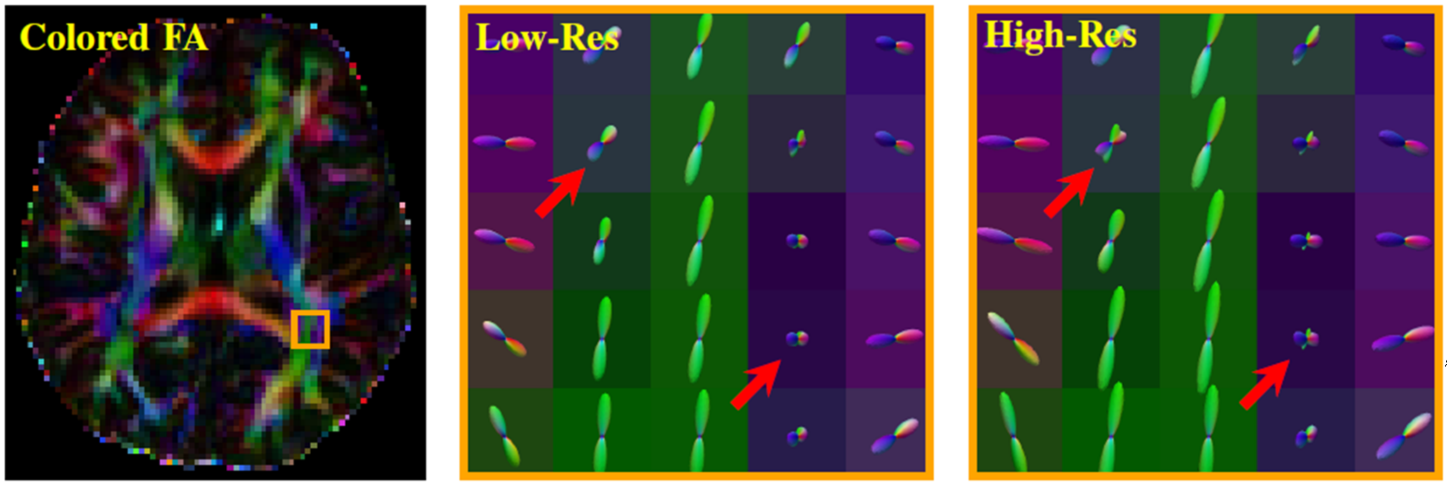
- High dimensional classification **[Dong and Hao, SPIE 2015]**
  - V.S. LDA methods: Leukemia (n=72, p=7129) and Lung (n=181, p=12533)

Error % (Std %)	NSC	IR	ROAD	RS-ROAD	Exact Model
Leukemia	8.51 (3.0)	4.27 (8.4)	6.35 (6.0)	4.46 (3.1)	5.57 (4.2)
Lung Cancer	10.44 (1.4)	3.47 (7.3)	1.37 (1.1)	0.93 (0.9)	0.59 (0.6)

- Application in super-resolution diffusion MRI **[Yap, Dong, Zhang,**

Leukemia

Lung



40

CONCLUDING REMARKS

Conclusions and Future Work

- **Conclusions**
  - Bridging wavelet frame transforms and differential operators
  - New insights, models/algorithms and applications
  - Sparse approximation for general data analysis
- **Future work**
  - Idea of "end-to-end" in classical problems such as imaging
  - Learning PDEs from data

Thanks for Your Attention
and
Questions?

http://bicmr.pku.edu.cn/~dongbin