Multimodal Data Processing

A Geometric Approach via Sparsity

João F. C. Mota

Heriot-Watt University, Edinburgh, UK

Side Information

Signal processing tasks

Denoising

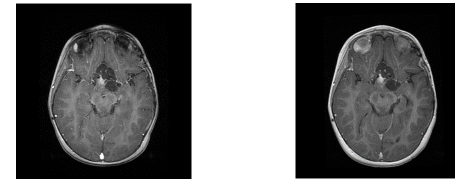
Reconstruction

Demixing (source separation)

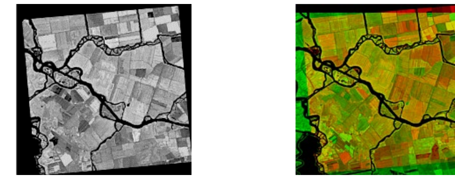
Compression

Inpainting, super-resolution, …

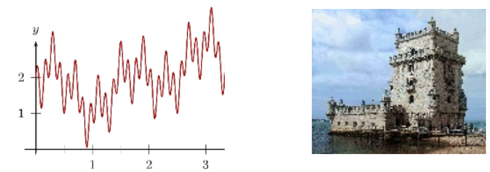
prior information



multi-modal

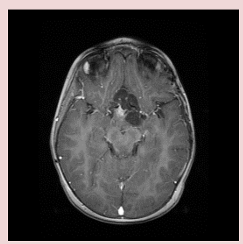


heterogeneous

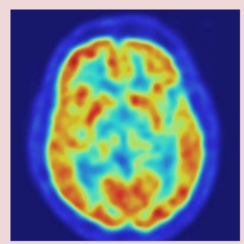


Recommender systems

Medical imaging



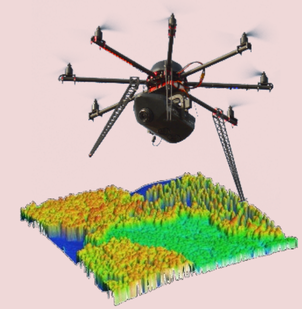
MRI



PET

Consumer electronics

Robotics



How to represent multi-modal or heterogeneous data ?

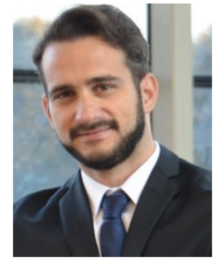
How to process it ?

Outline

- Compressed Sensing with Prior Information

- Towards Heterogeneous Data Processing

- Single-Image Super-Resolution



N Deligiannis
VUB-Belgium

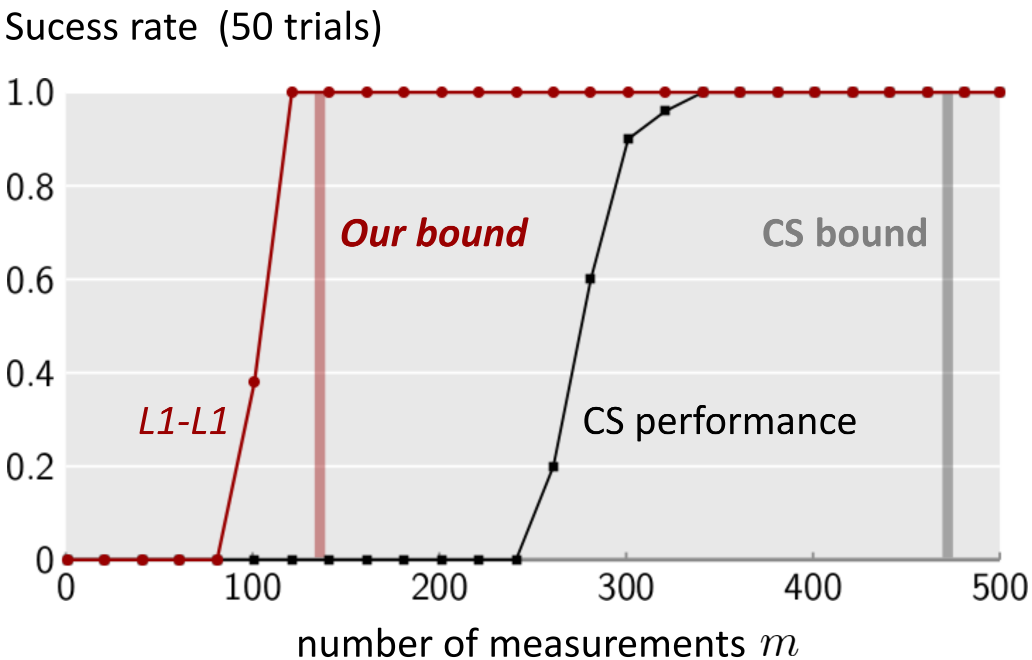
M Rodrigues
UCL

E Tsiligianni
VUB-Belgium

M Vella
Heriot-Watt U

Compressed Sensing

$$\|\overline{x} - x^\star\|_2 / \|x^\star\|_2 \simeq 0.45$$



Compressed Sensing (CS)

70-sparse

$$b = A : m \times n \; x^\star$$

iid Gaussian

1000

Basis pursuit

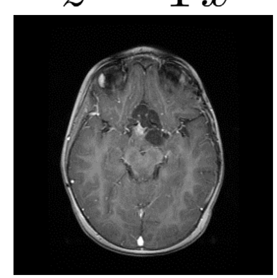
$$\widehat{x} = \arg\min_{x} \; \|x\|_1 \quad \text{s.t.} \quad b = Ax$$

What if we know $\overline{x} \sim x^\star$? ***prior information***

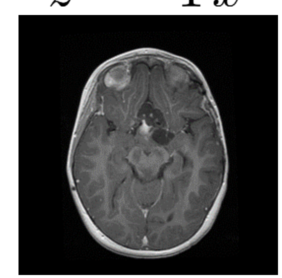
How to integrate $\overline{x}$ in the problem?

Reconstruction guarantees?

$\overline{z} = \Psi \overline{x}$



$z^\star = \Psi x^\star$



L1-L1 minimization

s-sparse

$$b = \underbrace{A : m \times n}_{\text{i.i.d. } \mathcal{N}(0, 1/m)} \; x^\star$$

Theorem (CS) **[Chandrasekaran et al, 12']**

$$m \geq 2s \log\left(\frac{n}{s}\right) + \frac{7}{5}s + 1 \implies x^\star = \begin{array}{ll} \underset{x}{\text{argmin}} & \|x\|_1 \\ \text{s.t.} & b = Ax \end{array} \quad \text{w.h.p.}$$

Theorem (L1-L1 minimization) **[Mota et al, 17']**

parameter-free

$\beta = 1$

$$\overline{h} := \left|\{i : x_i^\star > 0,\, \overline{x}_i < x_i^\star\} \cup \{i : x_i^\star < 0,\, \overline{x}_i > x_i^\star\}\right| > 0$$

$$m \geq 2\overline{h} \log\left(\frac{n}{s + \xi/2}\right) + \frac{7}{5}\left(s + \frac{\xi}{2}\right) + 1 \implies x^\star = \begin{array}{ll} \underset{x}{\text{argmin}} & \|x\|_1 + \|x - \overline{x}\|_1 \\ \text{s.t.} & b = Ax \end{array} \quad \text{w.h.p.}$$

$$0 \leq \overline{h} \leq s$$

$$\xi = \underbrace{\left|\{i : \overline{x}_i \neq x_i^\star = 0\}\right|}_{\text{support overestimation}} - \left|\{i : \overline{x}_i = x_i^\star \neq 0\}\right|$$

Heterogeneous Data Processing

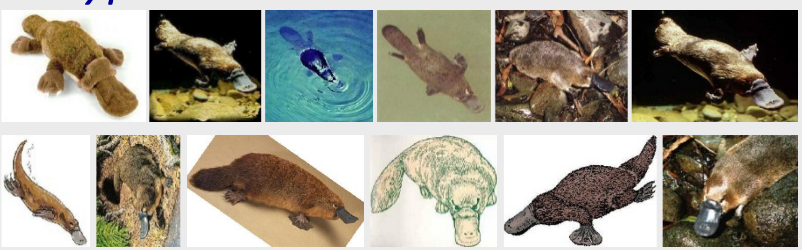
How to perform reconstruction, inpainting, classification w/ *heterogeneous* data?

different representations

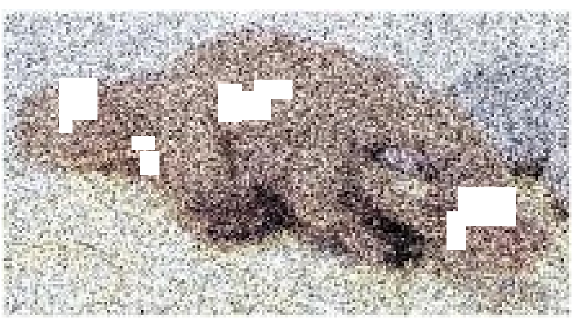
Conceptual example

Known database: *images* + *annotation (or description)*

Platypus



Airplane



+

Platypus **(side information)**

Heterogeneous Data Processing

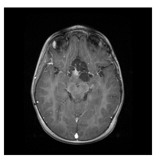
Our Approach

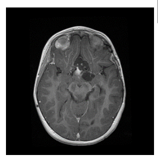
Recall L1-L1 minimization

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x\|_1 + \|x - \overline{x}\|_1 \\ \text{subject to} & Ax = b \end{array}$$

$$m \geq 2\,\overline{h}(x^\star, \overline{x}) \, \log\left(\frac{n}{s + \xi/2}\right) + \frac{7}{5}\left(s + \frac{\xi}{2}\right) + 1$$

$$\left|\{i : x_i^\star > 0,\, \overline{x}_i < x_i^\star\} \cup \{i : x_i^\star < 0,\, \overline{x}_i > x_i^\star\}\right|$$





homogeneous side info

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x\|_1 + \|x - \mathbf{g}(\overline{x})\|_1 \\ \text{subject to} & Ax = b \end{array}$$

learn $\mathbf{g}$

training data

$$\underset{\mathbf{g} \in \mathcal{G}}{\text{minimize}} \quad \frac{1}{T} \sum_{t=1}^{T} \overline{h}\left(x^{(t)},\, \mathbf{g}\left(\overline{x}^{(t)}\right)\right)$$

platypus

allowable maps $\mathbb{R}^q \to \mathbb{R}^n$

heterogeneous side info

Warning: we minimize $\overline{h}$, not full bound

Proof of concept

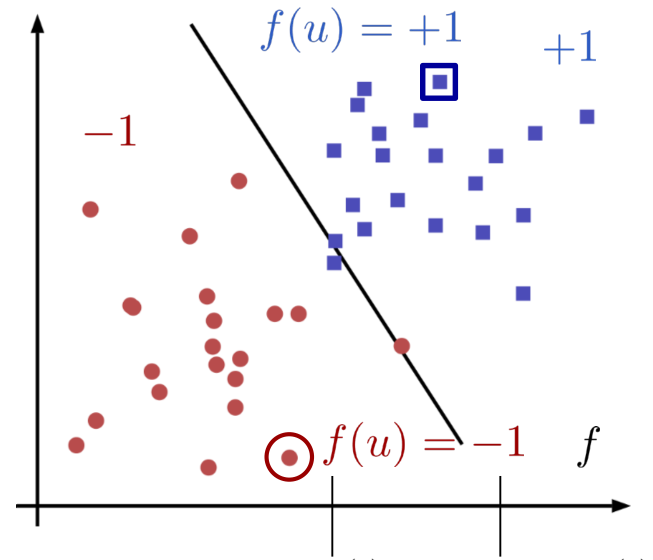
A Classification Problem ?

$$\underset{\mathbf{g} \in \mathcal{G}}{\text{minimize}} \ \frac{1}{T} \sum_{t=1}^{T} \overline{h}\left(x^{(t)}, \mathbf{g}(\overline{x}^{(t)})\right)$$

$$\overline{h}(x, \mathbf{g}(\overline{x})) = \left|\{i \,:\, x_i > 0,\ g_i(\overline{x}) < x_i\} \cup \{i \,:\, x_i < 0,\ g_i(\overline{x}) > x_i\}\right| = \sum_{i=1}^{n} 1_{\{\text{sign}(x_i)\cdot(x_i - g_i(y)) > 0\}}$$

$$\underbrace{\underset{\{g_i \in \mathcal{G}_i\}_{i=1}^{n}}{\text{minimize}}}_{\text{n independent problems}} \ \frac{1}{T} \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{1_{\left\{\text{sign}\left(x_i^{(t)}\right)\cdot\left(x_i^{(t)} - g_i\left(\overline{x}^{(t)}\right)\right) > 0\right\}}}_{\text{0-1 loss in a classification problem}}$$

$$L^{0\text{-}1} = \begin{cases} 1 & , \ f(u) \neq v \\ 0 & , \ f(u) = v \end{cases}$$



Problems: 0-1 loss minimization

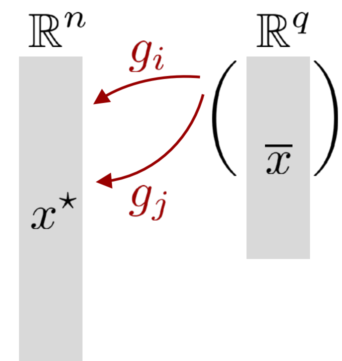
- NP-Hard
- Does not generalize well

Our solution: $\mathcal{G}_i =$ set of affine functions & 0-1 loss ➡ hinge loss

support vector machine

Performance Guarantees

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x\|_1 + \|x - \mathbf{g}(\overline{x})\|_1 \\ \text{subject to} & Ax = b \end{array}$$



$$m > 2\,\overline{h}\big(x^\star, \mathbf{g}(\overline{x})\big)\,\log\left(\frac{n}{s+\xi/2}\right) + \frac{7}{5}\left(s + \frac{\xi}{2}\right)$$

Learn each g_i : $\displaystyle \min_{w,r}\ \|w\|_2^2 + \frac{\lambda}{T}\sum_{t=1}^{T} \max\Big\{0,\ 1 - v^{(t)}\cdot\big(w^\top u^{(t)} + r\big)\Big\}$

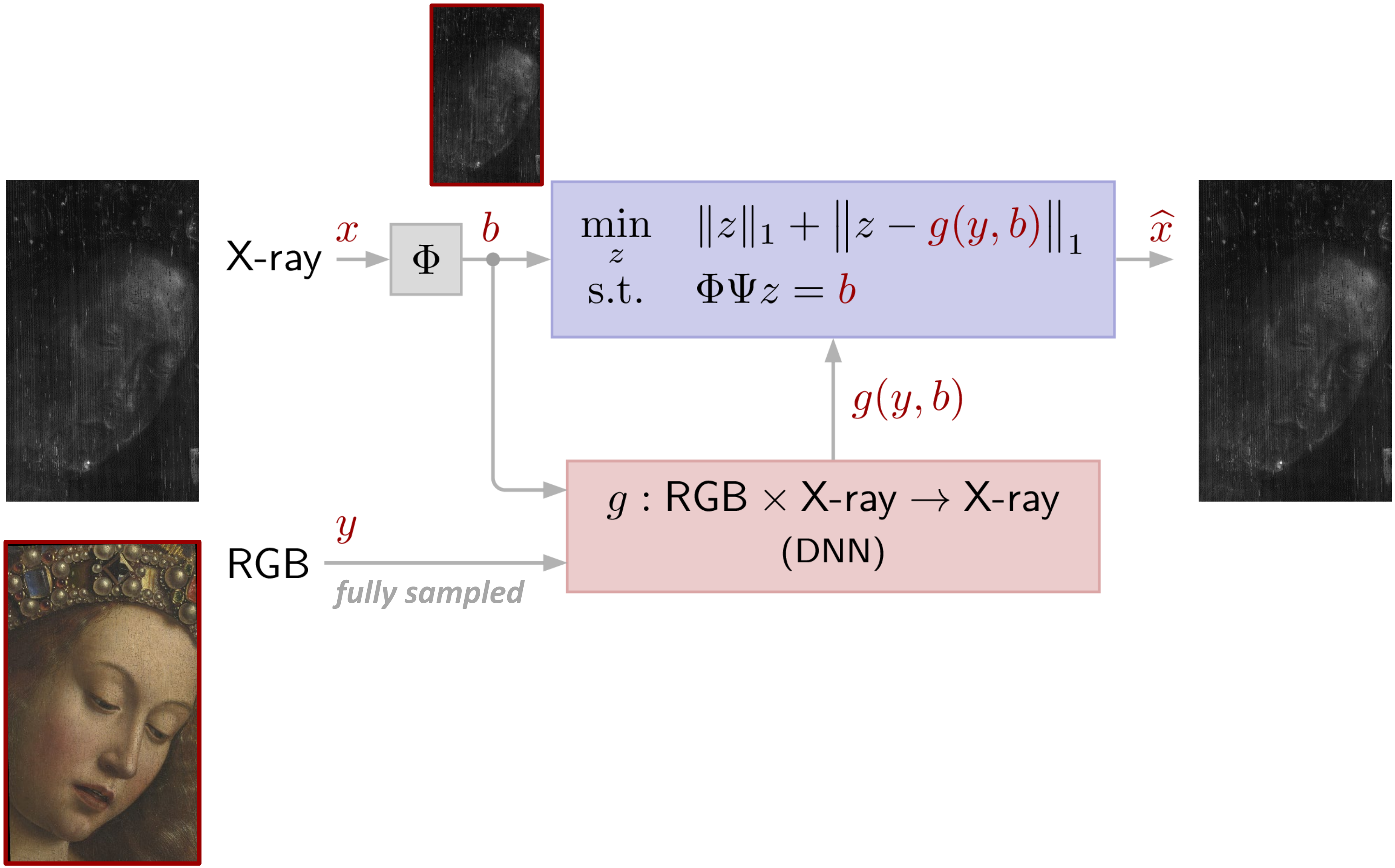
Theorem **[Mota et al., 17']**

Assume $\{x^{(t)},\, \overline{x}^{(t)}\}_{t=1}^{T}$ are iid realizations of $(X,\, \overline{X})$ w/ unknown joint pdf

Each g_i is learned as above (+ technical assumptions)

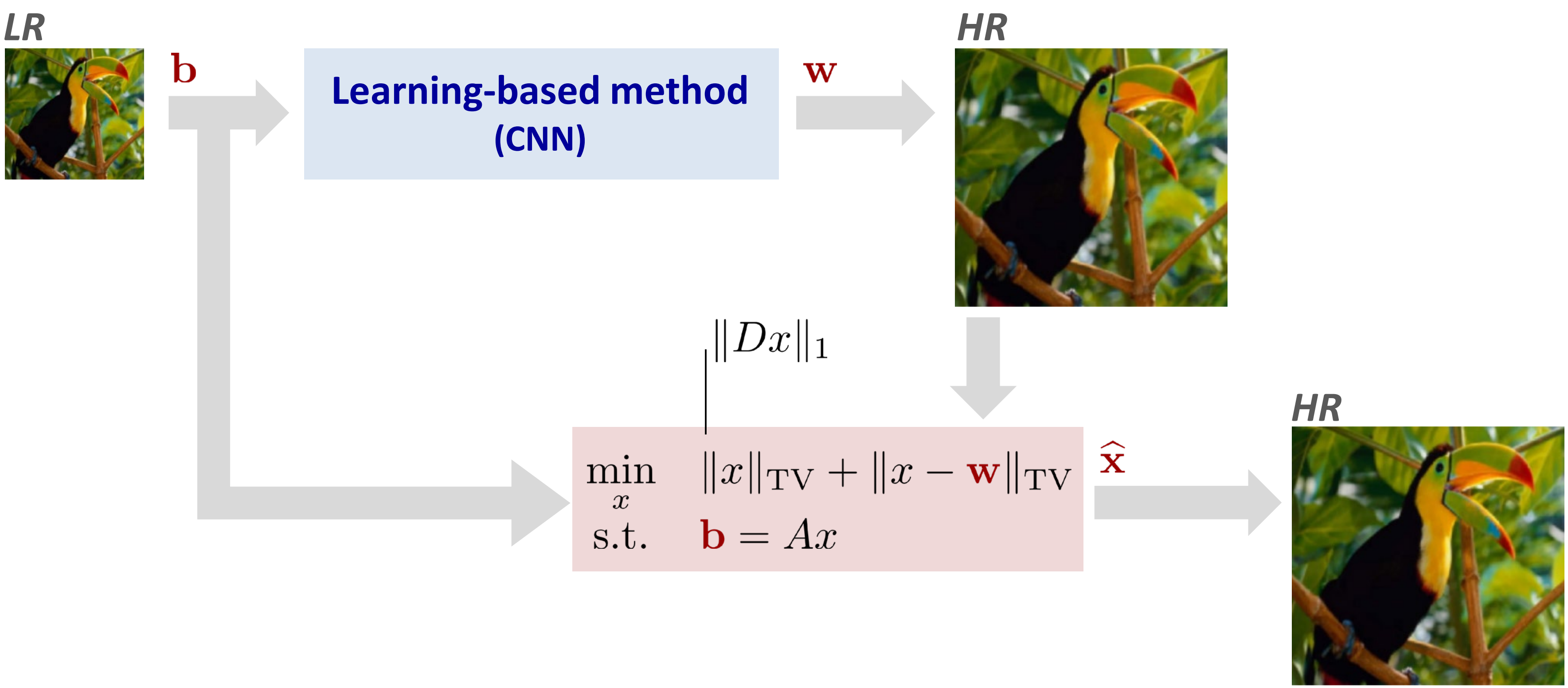
Then, with high probability, $\mathbb{E}\Big[\overline{h}\big(X,\, \mathbf{g}(\overline{X})\big)\Big] \leq O\Big(\frac{1}{\sqrt{T}}\Big)$ *# training data points*

X-Ray Super-Resolution aided by RGB



Preliminary Results

Single modality super-resolution



[Vella & Mota 19’, BMVC]

Preliminary Results

$$\min_x \ \|x\|_{\mathrm{TV}} \quad \text{s.t.} \quad b = Ax$$

Dataset	Scale	TVAL3	SRCNN*	*Ours*
Set5	×2	34.0315 (0.9354)	36.2772 (0.9509)	**36.5288 (0.9536)**
	×4	29.1708 (0.8349)	30.0765 (0.8525)	**30.2669 (0.8590)**
Set14	×2	31.0033 (0.8871)	31.9954 (0.9012)	**32.2949 (0.9057)**
	×4	26.6742 (0.7278)	27.1254 (0.7395)	**27.3040 (0.7480)**
BSD100	×2	30.1373 (0.8671)	31.1087 (0.8835)	**31.2241 (0.8866)**
	×4	26.3402 (0.6900)	26.7027 (0.7018)	**26.7838 (0.7085)**
Urban100	×2	27.5143 (0.8728)	28.6505 (0.8909)	**28.8415 (0.8939)**
	×4	23.7529 (0.6977)	24.1443 (0.7047)	**24.2368 (0.7114)**

PSNR (SSIM)

*C. Dong, C. C. Loy, K. He, X. Tang
Learning a deep convolutional neural network for image super-resolution
ECCV, 2014

Dataset	Scale	FSRCNN*	*Ours*	LapSRN**	*Ours*
Set5	×2	36.9912 (0.9556)	**37.0394** (**0.9559**)	37.7008 (0.9590)	**37.7219** (**0.9592**)
	×4	30.7122 (0.8658)	**30.8005** (**0.8691**)	31.7181 (0.8891)	**31.7428** (**0.8894**)
	×8			26.3314 (**0.7548**)	**26.3881** (0.7545)
Set14	×2	32.6515 (0.9089)	**32.6935** (**0.9092**)	33.2518 (0.9138)	**33.2709** (**0.9142**)
	×4	27.6179 (0.7550)	**27.6890** (**0.7574**)	28.2533 (0.7730)	**28.2722** (**0.7734**)
	×8			24.5643 (**0.6266**)	**24.5993** (0.6264)
BSD100	×2	31.5075 (0.8905)	**31.5250** (**0.8907**)	32.0214 (0.8970)	**32.0274** (**0.8975**)
	×4	26.9675 (0.7130)	**27.0011** (**0.7149**)	27.4164 (0.7296)	**27.4317** (**0.7300**)
	×8			24.6495 (**0.5887**)	**24.6769** (0.5886)
Urban100	×2	29.8734 (0.9010)	**29.8926** (**0.9013**)	31.1319 (0.9180)	**31.1462** (**0.9183**)
	×4	24.6196 (0.7270)	**24.6619** (**0.7297**)	25.5026 (0.7661)	**25.5167** (**0.7662**)
	×8			22.0547 (**0.5956**)	**22.0675** (0.5944)

*C. Dong, C. C. Loy, X. Tang
Accelerating the super-resolution convolutional neural network
ECCV, 2016

**W.-S. Lai, J.-B. Huang, N. Ahuja, M.-H. Yang
Deep Laplacian pyramid networks for fast and accurate super-resolution
CVPR, 2017

Dataset	Scale	SRMD*	*Ours*	ESRGAN**	*Ours*
Set5	$\times 2$	37.4496 (0.9579)	**37.5817** (**0.9585**)		
	$\times 4$	31.5750 (0.8853)	**31.6531** (**0.8863**)	32.7072 (0.9001)	**32.7170** (**0.9002**)
Set14	$\times 2$	33.1035 (0.9127)	**33.1868** (**0.9137**)		
	$\times 4$	28.1593 (0.7716)	**28.2174** (**0.7728**)	28.8342 (0.7877)	**28.9253** (**0.7891**)
BSD100	$\times 2$	31.8722 (0.8953)	**31.9009** (**0.8959**)		
	$\times 4$	27.3350 (0.7273)	**27.3579** (**0.7280**)	27.8332 (0.7447)	**27.8489** (**0.7447**)
Urban100	$\times 2$	30.8799 (0.9146)	**30.9253** (**0.9151**)		
	$\times 4$	25.3494 (0.7605)	**25.3834** (**0.7609**)	27.0270 (**0.8146**)	**27.0404** (**0.8146**)

*K. Zhang, W. Zuo, L. Zhang
Learning a single convolutional super-resolution network for multiple degradations
CVPR, 2018

**X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, C. C. Loy,
ESRGAN: Enhanced super-resolution generative adversarial networks
ECCV workshops, 2018

Conclusions

- Multimodal data processing: from *prior information* to *heterogeneous data*

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x\|_1 + \|x - \mathbf{g}(\overline{x})\|_1 \\ \text{subject to} & Ax = b \end{array}$$

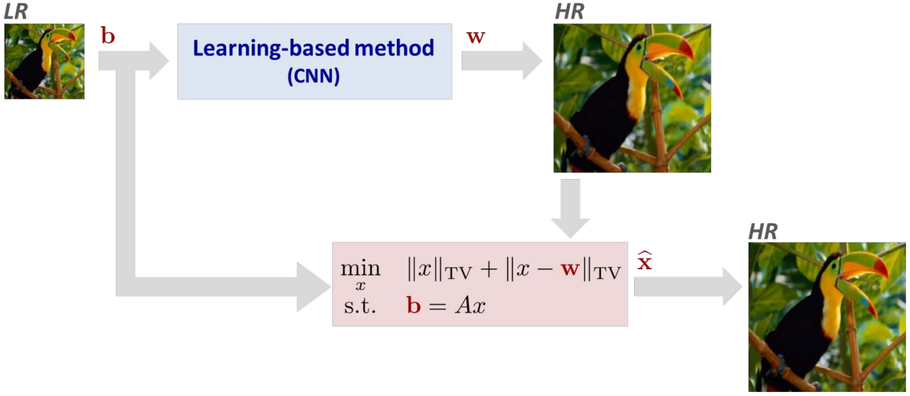
platypus

heterogeneous side info

- Theory informs practice

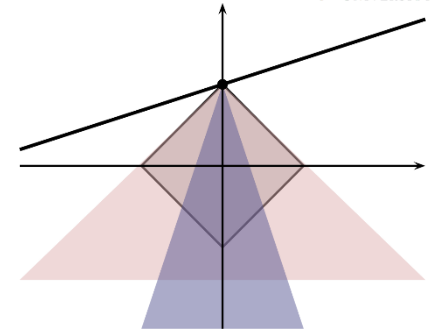
- *Models* still have a role and can complement data-driven methods

- More *theory needed*

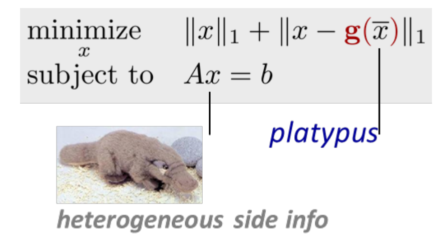


References

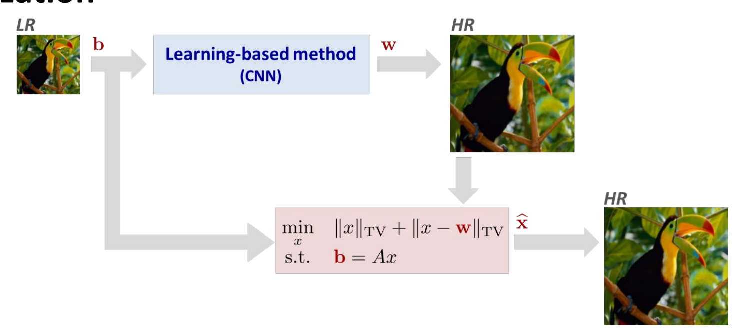
J. F. C. Mota, N. Deligiannis, M. R. D. Rodrigues
Compressed Sensing with Prior Information: Optimal Strategies, Geometries, and Bounds
IEEE Transactions on Information Theory, Vol 63, No 7, 2017



J. F. C. Mota, E. Tsiligianni, N. Deligiannis
A framework of learning affine transformations for multimodal sparse reconstruction
Wavelets and Sparsity XVII, SPIE Optical Engineering + Applications, 2017



M. Vella, J. F. C. Mota
Single Image Super-Resolution via CNN Architectures and TV-TV Minimization
British Machine Vision Conference (BMVC), 2019



Experimental Results

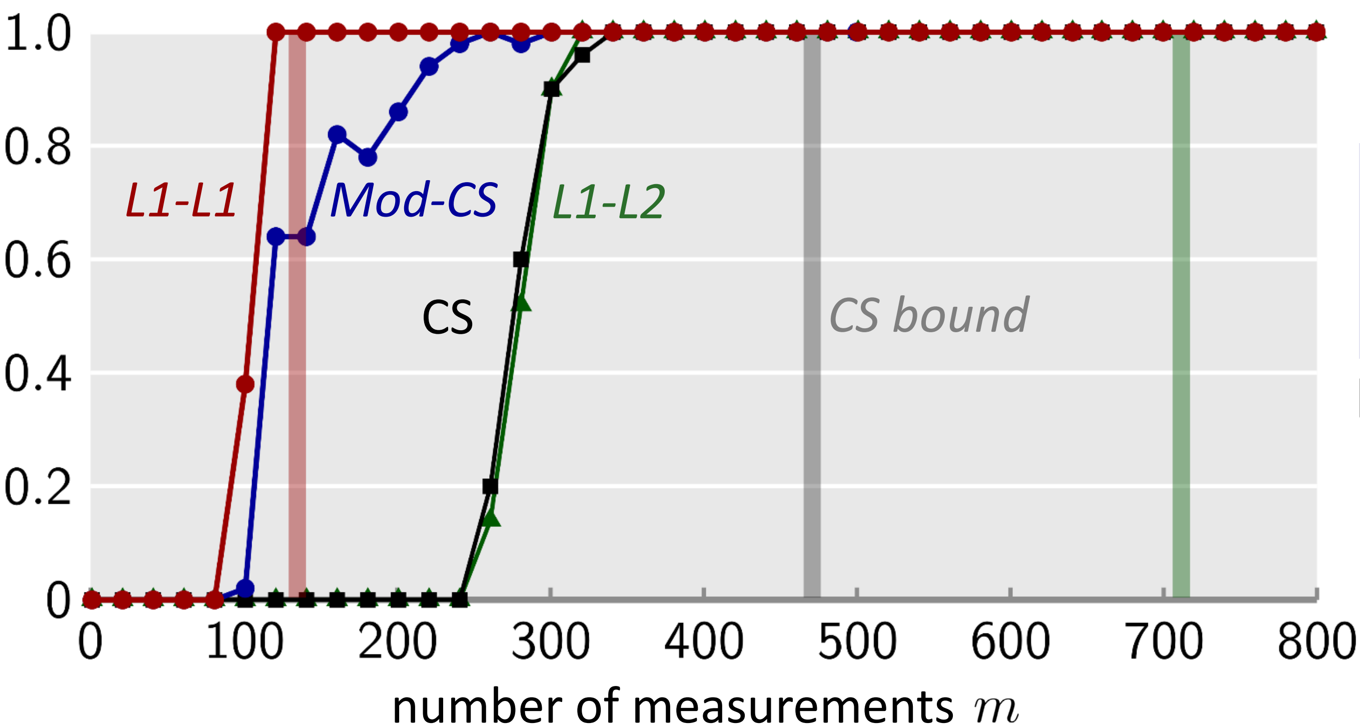
A_{ij} : Gaussian

$x^\star$: $n = 1000$ $\quad s = 70$ $\quad x_i^\star \sim \mathcal{N}(0,1)$ $\quad \|\overline{x} - x^\star\|_2 / \|x^\star\|_2 \simeq 0.45$

$\overline{x}$: $\overline{x} = x^\star + z$ $\quad \mathrm{card}(z) = 28$ $\quad z_i \sim \mathcal{N}(0, 0.8)$ $\quad \overline{h} = 11 \quad \xi = -42$

$\|\hat{x} - x^\star\|_\infty / \|x^\star\|_\infty \leq 10^{-2}$

Sucess rate (50 trials)



Mod-CS

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x_{T^c}\|_1 \\ \text{subject to} & Ax = b \end{array}$$

[Vaswani and Lu, 2010]

Intuition

$Ax^\star = b$ measurements

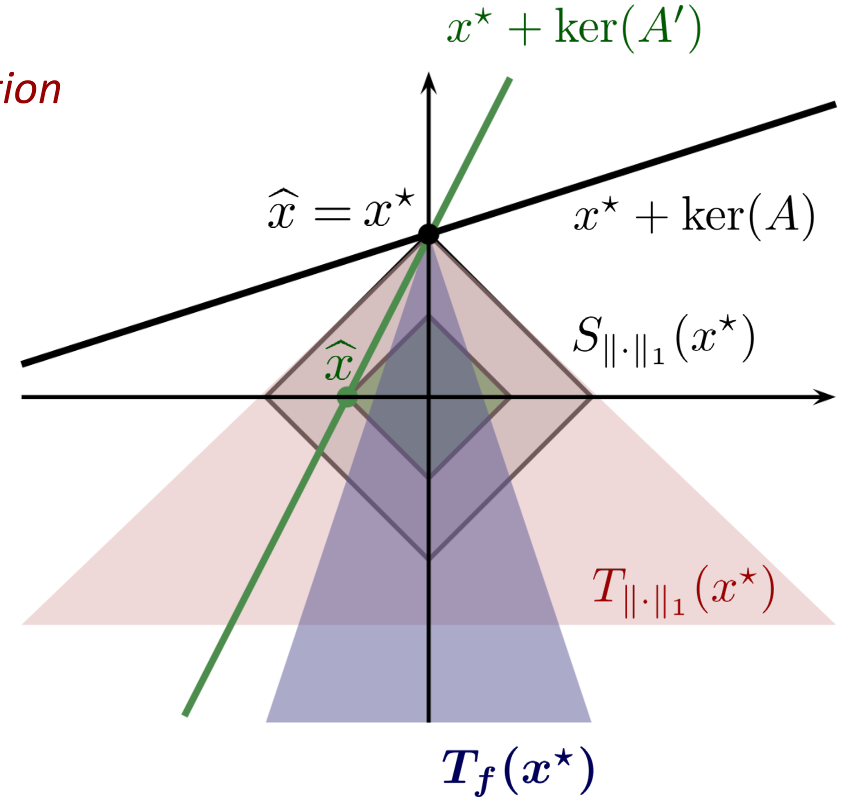
random orientation

solutions of $Ax = b$: $x^\star + \ker(A)$

$$\widehat{x} = \operatorname*{argmin}_{x} \; \|x\|_1 \quad \text{s.t.} \quad Ax = b$$

Tangent cone of f at $x^\star$

$$T_f(x^\star) = \text{cone}\big(\underbrace{S_f(x^\star)}_{\{x \,:\, f(x) \leq f(x^\star)\}} - x^\star\big)$$



Our approach

$\overline{x}$: prior information (PI)

g : model for PI $\quad g(x^\star - \overline{x}) \simeq$ small

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f(x) = \|x\|_1 + \beta\, g(x - \overline{x}) \\ \text{subject to} & Ax = b \end{array} \quad \Rightarrow \quad \boldsymbol{T_f(x^\star)}$$

$$g_1(x - \overline{x}) = \|x - \overline{x}\|_1$$

$$g_2(x - \overline{x}) = \tfrac{1}{2}\|x - \overline{x}\|_2^2$$

$\beta = 1$

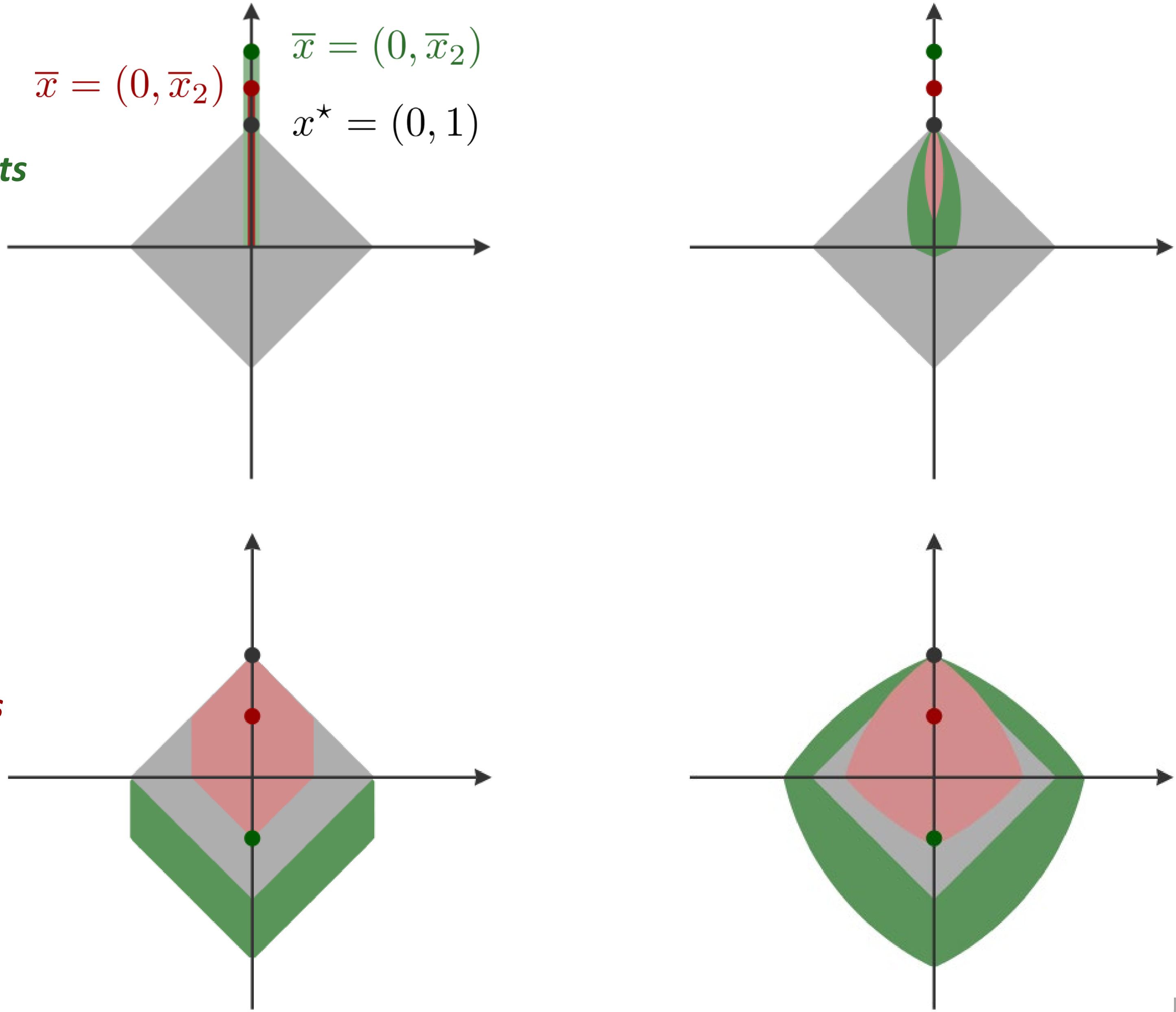
$\ell_1 - \ell_1$ $\quad\quad\quad\quad$ $\ell_1 - \ell_2$

Good components

$\overline{x}_2 > x_2^\star$

Bad components

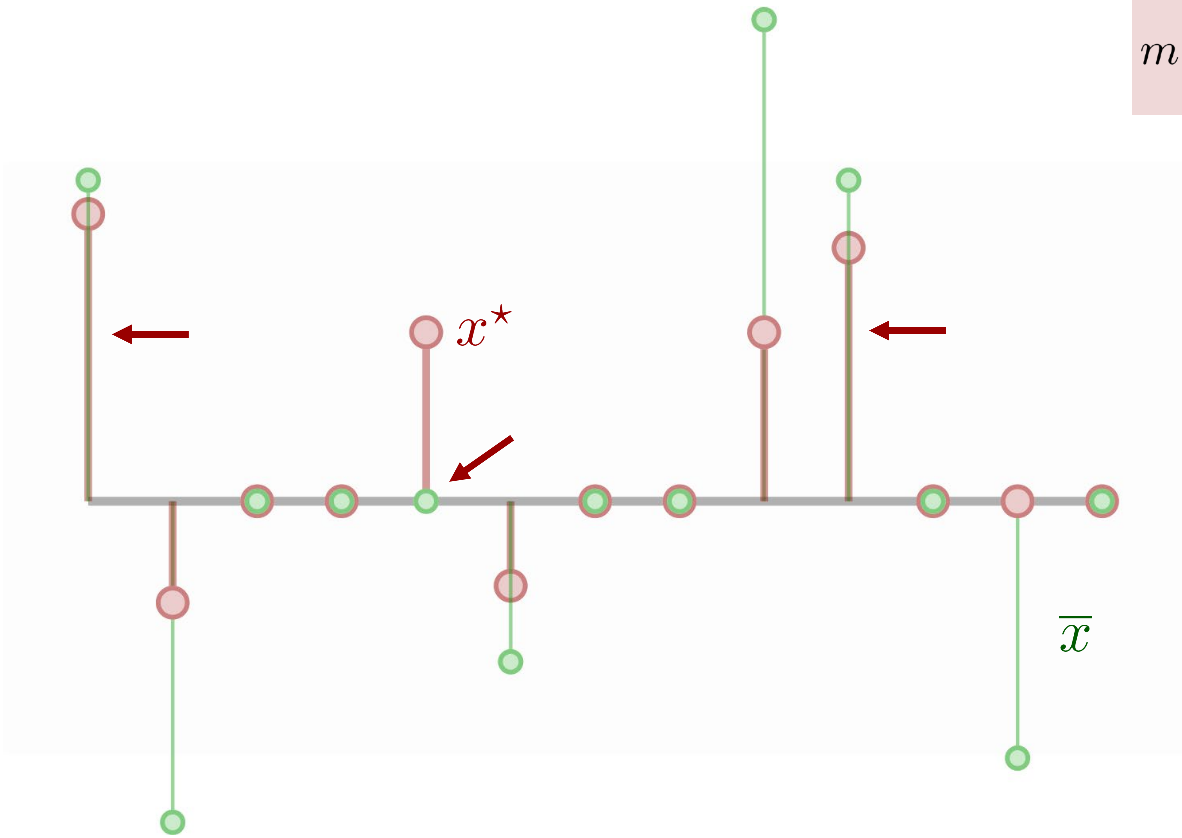
$\overline{x}_2 < x_2^\star$



Improving Prior Information?

$$m \geq 2\overline{\boldsymbol{h}} \log\left(\frac{n}{s+\xi/2}\right) + \frac{7}{5}(s+\frac{\xi}{2}) + 1$$

 # bad components



$\boldsymbol{c = 1.0}$

$\overline{h} = 3 \Rightarrow \overline{h} = 1$

Improve $\overline{x}$ by $\overline{x} \leftarrow c \cdot \overline{x}$ $\quad c > 1$

Warning: $\overline{h} > 0$

CS Geometry & Known Results

Theorem *(Chandrasekaran et al "The Convex Geometry of Linear Inverse Problems" 2012)*

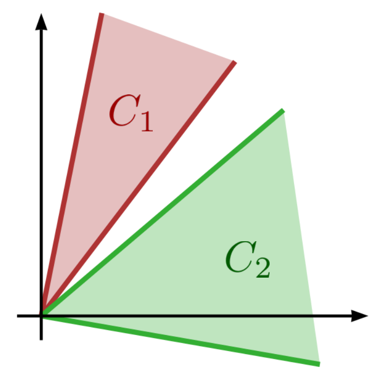
$$A : m \times n\,, \quad A_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/m)$$

$f : \mathbb{R}^n \to \mathbb{R}$ convex

$$\begin{cases} b = Ax^\star \\ m \geq w(T_f(x^\star))^2 + 1 \end{cases} \qquad \widehat{x} = \begin{array}{ll} \underset{x}{\text{argmin}} & f(x) \\ \text{s.t.} & Ax = b \end{array} \qquad \Longrightarrow \qquad \widehat{x} = x^\star \quad \text{w.h.p.}$$

Gaussian width

$$w(C) = \mathbb{E}_g\Big[\sup_{z \in C\, \cap\, \mathbb{B}^n(0,1)} g^\top z\Big], \quad g \sim \mathcal{N}(0,1)$$



$w(C_2) > w(C_1)$

s-sparse

$$f(x) = \|x\|_1 \;\Rightarrow\; w\big(T_f(x^\star)\big)^2 \leq 2s\log\Big(\frac{n}{s}\Big) + \frac{7}{5}s$$

$$f_1(x) = \|x\|_1 + \beta\|x - \overline{x}\|_1 \;\Rightarrow\; \boldsymbol{w\big(T_{f_1}(x^\star)\big)^2 \leq \;?}$$

ℓ_1-ℓ_1 *minimization*

$$f_2(x) = \|x\|_1 + \frac{\beta}{2}\|x - \overline{x}\|_2^2 \;\Rightarrow\; \boldsymbol{w\big(T_{f_2}(x^\star)\big)^2 \leq \;?}$$

ℓ_1-ℓ_2 *minimization*

Theorem: L1-L1 $(\beta = 1)$

$A : m \times n\,,$ $A_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/m)$

$Ax^\star = b$ measurements $\overline{h} > 0$

$$\widehat{x} = \begin{array}{cl} \underset{x}{\operatorname{argmin}} & \|x\|_1 + \|x - \overline{x}\|_1 \\ \text{s.t.} & Ax = b \end{array}$$

$$m \geq 2\overline{h} \log\left(\frac{2n}{q + h + \overline{h}}\right) + \frac{7}{10}(q + h + \overline{h}) + 1 \quad \Longrightarrow \quad \widehat{x} = x^\star \ \text{ w.h.p.}$$

$$m \geq 2\overline{h} \log\left(\frac{n}{s + \xi/2}\right) + \frac{7}{5}(s + \xi/2) + 1$$

standard CS: $m \geq 2s \log\left(\frac{n}{s}\right) + \frac{7}{5}s + 1$

Theorem: L1-L2 $(\beta = 1)$

$A : m \times n\,,$ $A_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/m)$

$Ax^\star = b$ measurements *(+ some realistic assumptions)*

$$\widehat{x} = \begin{array}{cl} \underset{x}{\operatorname{argmin}} & \|x\|_1 + \frac{1}{2}\|x - \overline{x}\|_2^2 \\ \text{s.t.} & Ax = b \end{array}$$

$\#i \,:\, x_i^\star = 0,\ |\overline{x}_i| \geq 1$

$$m \geq 2v \log\left(\frac{n}{q}\right) + s + 2R + \frac{4}{5}q + 1 \quad \Longrightarrow \quad \widehat{x} = x^\star \ \text{ w.h.p.}$$

$$v := \sum_{i : x_i^\star > 0} (1 + x_i^\star - \overline{x}_i)^2 + \sum_{i : x_i^\star < 0} (1 + \overline{x}_i - x_i^\star)^2 + \sum_{i : x_i^\star = 0, \overline{x}_i \neq 0} (|\overline{x}_i| - 1)^2$$