Strengthened Sobolev inequalities for a random subspace of functions

Rachel Ward

University of Texas at Austin

April 2013

Discrete Sobolev inequalities

Proposition (Sobolev inequality for discrete images)

Let $x \in \mathbb{R}^{N \times N}$ be zero-mean. Then

$$\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N} x_{j,k}^2} \leq \sum_{j=1}^{N}\sum_{k=1}^{N} \Big[|x_{j+1,k} - x_{j,k}| + |x_{j,k+1} - x_{j,k}| \Big]$$

or

$$\|x\|_2 \leq \|x\|_{TV}$$

Theorem (New! Strengthened Sobolev inequality)

With probability $\geq 1 - e^{-cm}$, the following holds for all images $x \in \mathbb{R}^{N \times N}$ in the null space of an $m \times N^2$ random Gaussian matrix

$$\|x\|_2 \lesssim \frac{[\log(N)]^{3/2}}{\sqrt{m}} \|x\|_{TV}$$

Discrete Sobolev inequalities

Proposition (Sobolev inequality for discrete images)

Let $x \in \mathbb{R}^{N \times N}$ be zero-mean. Then

$$\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N} x_{j,k}^2} \leq \sum_{j=1}^{N}\sum_{k=1}^{N} \Big[|x_{j+1,k} - x_{j,k}| + |x_{j,k+1} - x_{j,k}| \Big]$$

or

$$\|x\|_2 \leq \|x\|_{TV}$$

Theorem (New! Strengthened Sobolev inequality)

With probability $\geq 1 - e^{-cm}$, the following holds for all images $x \in \mathbb{R}^{N \times N}$ in the null space of an $m \times N^2$ random Gaussian matrix

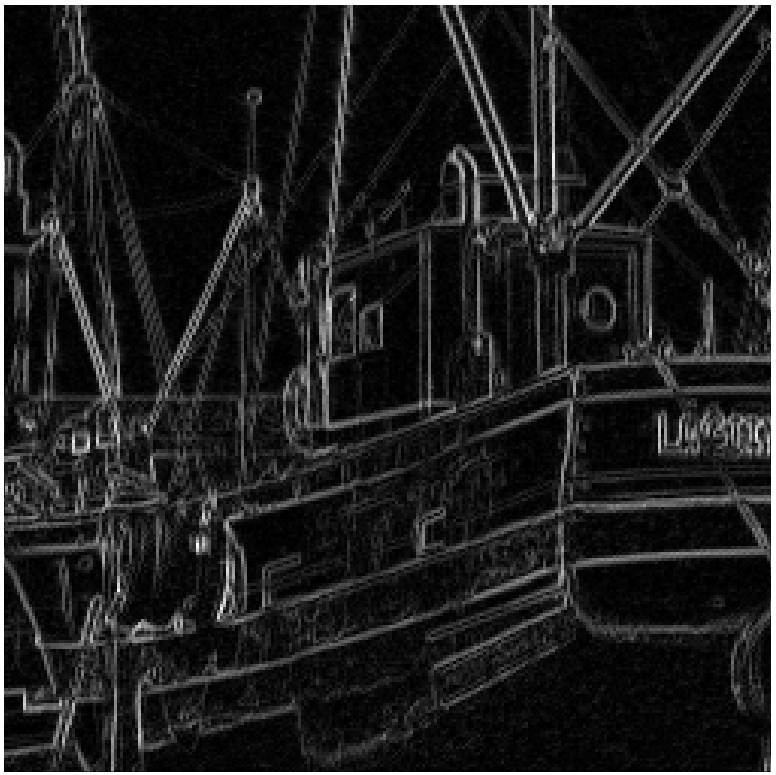
$$\|x\|_2 \lesssim \frac{[\log(N)]^{3/2}}{\sqrt{m}} \|x\|_{TV}$$

Joint work with Deanna Needell

Research supported in part by the ONR and Alfred P. Sloan Foundation

Motivation: Image processing

Images are compressible in discrete gradient



Images are compressible in discrete gradient

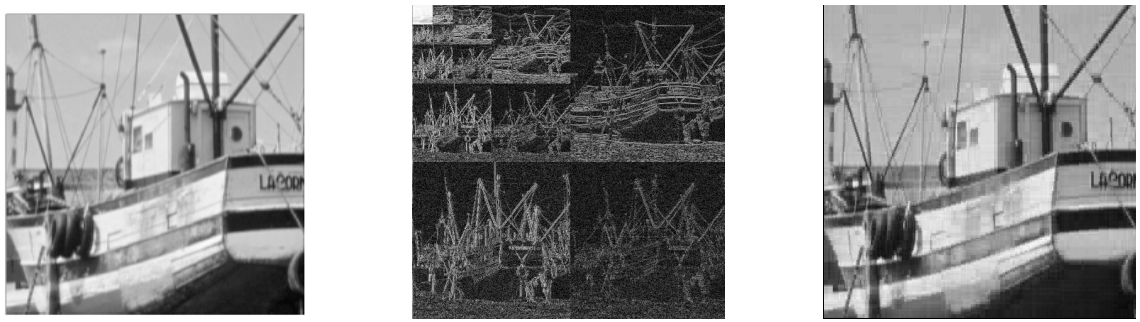
We define the discrete directional derivatives of an image $x \in \mathbb{R}^{N\times N}$ as

$$x_u : \mathbb{R}^{N\times N} \to \mathbb{R}^{(N-1)\times N}, \qquad (x_u)_{j,k} = x_{j,k} - x_{j-1,k},$$
$$x_v : \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times(N-1)}, \qquad (x_v)_{j,k} = x_{j,k} - x_{j,k-1},$$

and discrete gradient operator

$$\nabla x = (x_u, x_v) \in \mathbb{R}^{N\times N\times 2}$$

Images are also compressible in wavelet bases



Reconstruction of "Boats" from 10% of its Haar wavelet coefficients.

$\mathcal{H} : \mathbb{R}^{N \times N} \to \mathbb{R}^{N \times N}$ is bivariate Haar wavelet transform

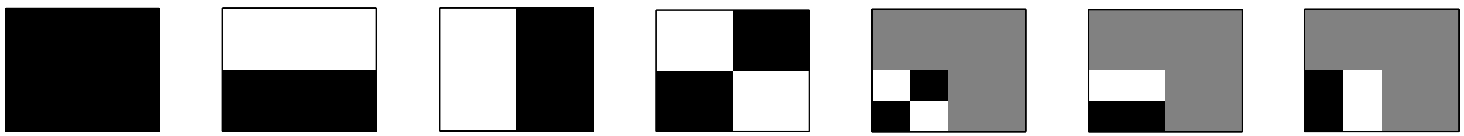


Figure: First few Haar basis functions

Notation

The image ℓ_p-norm is $\|u\|_p := \left(\sum_{j=1}^{N} \sum_{k=1}^{N} |u_{j,k}|^p \right)^{1/p}$

u is s-sparse if $\|u\|_0 := \{\#(j,k) : u_{j,k} \neq 0\} \leq s$

$u_s = \arg\min_{\{z:\|z\|_0=s\}} \|u - z\|$ is the best s-sparse approximation to u

$\sigma_s(u)_p = \|u - u_s\|_p$ is the s-sparse approximation error

Natural images are 'compressible':

$\sigma_s(\nabla x) = \|\nabla x - (\nabla x)_s\|_2$ decays quickly in s

$\sigma_s(\mathcal{H} x) = \|\mathcal{H} x - (\mathcal{H} x)_s\|_2$ decays quickly in s

Notation

The image ℓ_p-norm is $\|u\|_p := \left(\sum_{j=1}^{N} \sum_{k=1}^{N} |u_{j,k}|^p \right)^{1/p}$

u is s-sparse if $\|u\|_0 := \{\#(j,k) : u_{j,k} \neq 0\} \leq s$

$u_s = \arg\min_{\{z:\|z\|_0=s\}} \|u - z\|$ is the best s-sparse approximation to u

$\sigma_s(u)_p = \|u - u_s\|_p$ is the s-sparse approximation error

Natural images are 'compressible':

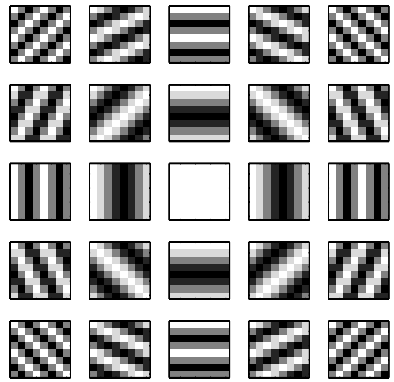
$\sigma_s(\nabla x) = \|\nabla x - (\nabla x)_s\|_2$ decays quickly in s

$\sigma_s(\mathcal{H} x) = \|\mathcal{H} x - (\mathcal{H} x)_s\|_2$ decays quickly in s

Imaging via compressed sensing

If images are so compressible (nearly low-dimensional), do we really need to know all of the pixel values for accurate reconstruction, or can we get away with taking a smaller number of linear measurements?

MRI imaging - speed up by exploiting sparsity?



In Magnetic Resonance Imaging (MRI), rotating magnetic field measurements are 2D Fourier transform measurements:

$$y_{k_1,k_2} = \left(\mathcal{F}(x)\right)_{k_1,k_2} = \frac{1}{N}\sum_{j_1,j_2} x_{j_1,j_2}\, e^{2\pi i(k_1 j_1 + k_2 j_2)/N},$$

$$-N/2+1 \leq k_1, k_2, \leq N/2$$

Each measurement takes a certain amount of time — reduced number of samples $m \ll N^2$ means reduced time of MRI scan.

Compressed sensing set-up

1. Signal (or image) of interest $f \in \mathbb{R}^d$ which is sparse: $\|f\|_0 \leq s$
2. Measurement operator $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^m$, $m \ll d$

$$\begin{bmatrix} \\ y \\ \\ \end{bmatrix} = \begin{bmatrix} & & \\ & \mathcal{A} & \\ & & \end{bmatrix} \begin{bmatrix} \\ \\ f \\ \\ \\ \end{bmatrix}$$

3. Problem: Using the sparsity of f, can we reconstruct f efficiently from y?

A sufficient condition: Restricted Isometry Property[1]

- $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^m$ satisfies the Restricted Isometry Property (RIP) of order s if

$$\frac{3}{4}\|f\|_2 \leq \|\mathcal{A}f\|_2 \leq \frac{5}{4}\|f\|_2 \quad \text{for all } s\text{-sparse } f.$$

- Gaussian or Bernoulli measurement matrices satisfy the RIP with high probability when

$$m \gtrsim s \log d.$$

- Random subsets of rows from Fourier and others with fast multiply have similar property: $m \gtrsim s \log^4 d$.

[1] Candès-Romberg-Tao 2006

A sufficient condition: Restricted Isometry Property[1]

- $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^m$ satisfies the Restricted Isometry Property (RIP) of order s if

$$\frac{3}{4}\|f\|_2 \leq \|\mathcal{A}f\|_2 \leq \frac{5}{4}\|f\|_2 \quad \text{for all } s\text{-sparse } f.$$

- Gaussian or Bernoulli measurement matrices satisfy the RIP with high probability when

$$m \gtrsim s \log d.$$

- Random subsets of rows from Fourier and others with fast multiply have similar property: $m \gtrsim s \log^4 d$.

[1] Candès-Romberg-Tao 2006

ℓ_1-minimization for sparse recovery[2]

Let $y = \mathcal{A}f$. Let $\mathcal{A}$ satisfy the Restricted Isometry Property of order s and set:

$$\hat{f} = \operatorname*{argmin}_g \|g\|_1 \quad \text{such that} \quad \mathcal{A}g = y$$

Then

1. **Exact recovery:** If f is s-sparse, then $\hat{f} = f$.
2. **Stable recovery:** If f is approximately s-sparse, then $\hat{f} \approx f$.

Precisely,

$$\|f - \hat{f}\|_1 \lesssim \sigma_s(f)_1$$

$$\|f - \hat{f}\|_2 \lesssim \frac{1}{\sqrt{s}} \sigma_s(f)_1$$

[2] Candès-Romberg-Tao, Donoho

ℓ_1-minimization for sparse recovery

Similarly, if $\mathcal{B}$ be an orthonormal basis in which f is assumed sparse, and $\mathcal{A}\mathcal{B}^* = \mathcal{A}\mathcal{B}^{-1}$ satisfies the Restricted Isometry Property of order s and

$$\hat{f} = \operatorname*{argmin}_g \|\mathcal{B}g\|_1 \quad \text{such that} \quad \mathcal{A}g = y$$

$$\Leftrightarrow \qquad \mathcal{B}\hat{f} = \operatorname*{argmin}_h \|h\|_1 \quad \text{such that} \quad \mathcal{A}\mathcal{B}^{-1}h = y$$

Then

$$\|\mathcal{B}(f - \hat{f})\|_1 \lesssim \sigma_s(\mathcal{B}f)_1$$

$$\|f - \hat{f}\|_2 \lesssim \frac{1}{\sqrt{s}}\sigma_s(\mathcal{B}f)_1$$

Total variation minimization for sparse recovery

For $x \in \mathbb{R}^{N \times N}$, the total variation seminorm is $\|x\|_{TV} = \|\nabla x\|_1$.

If x has s-sparse gradient, from observations $y = \mathcal{A}x$ one solves

Total Variation minimization

$$\hat{x} = \underset{z}{\operatorname{argmin}} \|z\|_{TV} \quad \text{subject to} \quad \mathcal{A}z = \mathcal{A}x$$

Immediate guarantees:

1. $\|\nabla \hat{x} - \nabla x\|_2 \lesssim \frac{1}{\sqrt{s}} \sigma_s(\nabla x)_1$,
2. $\|\nabla \hat{x} - \nabla x\|_1 \lesssim \sigma_s(\nabla x)_1$

Apply discrete Sobolev inequality $\Rightarrow \|x - \hat{x}\|_2 \lesssim \sigma_s(\nabla x)_1$

Can we do better?

Total variation minimization for sparse recovery

For $x \in \mathbb{R}^{N \times N}$, the total variation seminorm is $\|x\|_{TV} = \|\nabla x\|_1$.

If x has s-sparse gradient, from observations $y = \mathcal{A}x$ one solves

Total Variation minimization

$$\hat{x} = \underset{z}{\operatorname{argmin}} \|z\|_{TV} \quad \text{subject to} \quad \mathcal{A}z = \mathcal{A}x$$

Immediate guarantees:

1. $\|\nabla \hat{x} - \nabla x\|_2 \lesssim \frac{1}{\sqrt{s}} \sigma_s(\nabla x)_1$,
2. $\|\nabla \hat{x} - \nabla x\|_1 \lesssim \sigma_s(\nabla x)_1$

Apply discrete Sobolev inequality $\Rightarrow \|x - \hat{x}\|_2 \lesssim \sigma_s(\nabla x)_1$

Can we do better?

Discrete Sobolev inequalities

Proposition (Sobolev inequality)

Let $x \in \mathbb{R}^{N \times N}$ be zero-mean. Then

$$\|x\|_2 \leq \|x\|_{TV}$$

Theorem (Strengthened Sobolev inequality)

For all images $x \in \mathbb{R}^{N^2}$ in the null space of a matrix $\mathcal{A} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ which, when composed with the bivariate Haar wavelet transform, $\mathcal{A}\mathcal{H}^{-1} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the Restricted Isometry Property of order $2s$, the following holds:

$$\|x\|_2 \lesssim \frac{\log(N)}{\sqrt{s}} \|x\|_{TV}$$

Discrete Sobolev inequalities

Proposition (Sobolev inequality)

Let $x \in \mathbb{R}^{N \times N}$ be zero-mean. Then

$$\|x\|_2 \leq \|x\|_{TV}$$

Theorem (Strengthened Sobolev inequality)

For all images $x \in \mathbb{R}^{N^2}$ in the null space of a matrix $\mathcal{A} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ which, when composed with the bivariate Haar wavelet transform, $\mathcal{A}\mathcal{H}^{-1} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the Restricted Isometry Property of order $2s$, the following holds:

$$\|x\|_2 \lesssim \frac{\log(N)}{\sqrt{s}} \|x\|_{TV}$$

Consequence for total variation minimization

Suppose that the matrix $\mathcal{A} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ is such that, composed with the bivariate Haar wavelet transform, $\mathcal{A}\mathcal{H}^* : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the Restricted Isometry Property of order $2s$. From observations $y = \mathcal{A}x$, consider

Total Variation minimization

$$\hat{x} = \underset{z}{\operatorname{argmin}} \|z\|_{TV} \quad \text{subject to} \quad \mathcal{A}z = y$$

Then

1. $\|\nabla\hat{x} - \nabla x\|_2 \lesssim \frac{1}{\sqrt{s}}\sigma_s(\nabla x)_1$
2. $\|\hat{x} - x\|_{TV} \lesssim \sigma_s(\nabla x)_1$
3. $\|x - \hat{x}\|_2 \lesssim \frac{\log(N^2/s)}{\sqrt{s}}\sigma_s(\nabla x)_1$

Consequence for total variation minimization

Suppose that the matrix $\mathcal{A} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ is such that, composed with the bivariate Haar wavelet transform, $\mathcal{A}\mathcal{H}^* : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the Restricted Isometry Property of order $2s$. From observations $y = \mathcal{A}x$, consider

Total Variation minimization

$$\hat{x} = \underset{z}{\operatorname{argmin}} \|z\|_{TV} \quad \text{subject to} \quad \mathcal{A}z = y$$

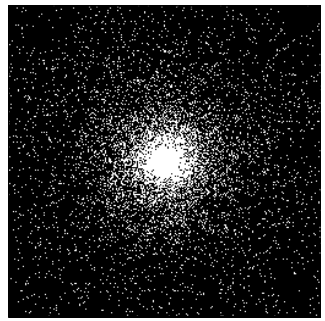
Then

1. $\|\nabla\hat{x} - \nabla x\|_2 \lesssim \frac{1}{\sqrt{s}}\sigma_s(\nabla x)_1$
2. $\|\hat{x} - x\|_{TV} \lesssim \sigma_s(\nabla x)_1$
3. $\|x - \hat{x}\|_2 \lesssim \frac{\log(N^2/s)}{\sqrt{s}}\sigma_s(\nabla x)_1$

Corollary (Sobolev inequality for random matrices)

With probability $\geq 1 - e^{-cm}$, *the following holds for all images* $x \in \mathbb{R}^{N \times N}$ *in the null space of an* $m \times N^2$ *random Gaussian matrix:*

$$\|x\|_2 \lesssim \frac{[\log(N)]^{3/2}}{\sqrt{m}} \|x\|_{TV}$$



Corollary (Sobolev inequality for Fourier constraints)

Select m frequency pairs (k_1, k_2) with replacement from the distribution

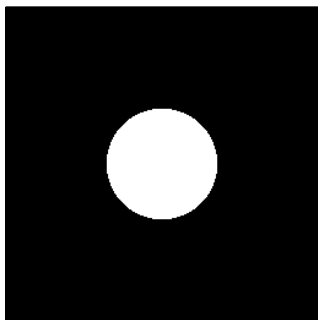
$$p_{k_1,k_2} \propto \frac{1}{(|k_1|+1)^2 + (|k_2|+1)^2}.$$

With probability $\geq 1 - e^{-cm}$, the following holds for all images $x \in \mathbb{R}^{N \times N}$ with vanishing Fourier transform $\big(\mathcal{F}(x)\big)_{k_1,k_2} = 0$:

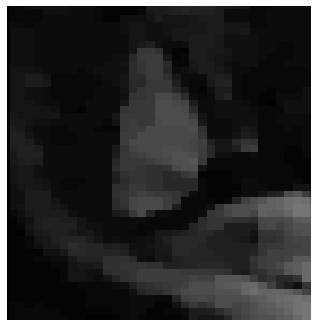
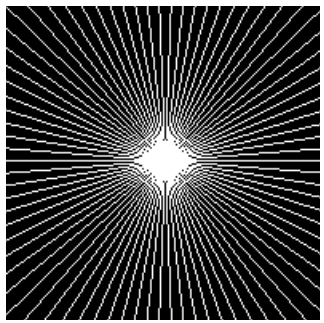
$$\|x\|_2 \lesssim \frac{[\log(N)]^5}{\sqrt{m}} \|x\|_{TV}$$

Sobolev inequalities in action!

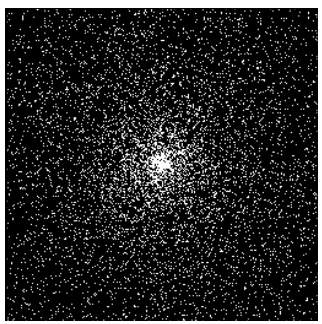
Original MRI image



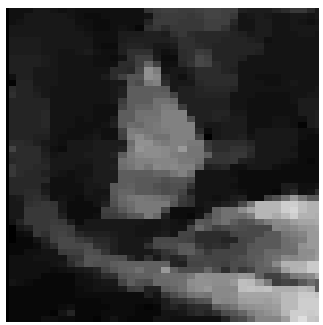
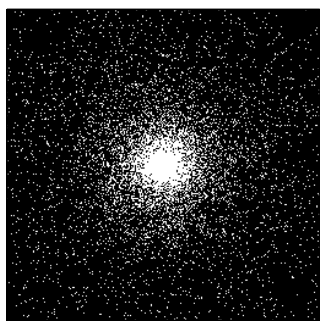
Low frequencies only

Equispaced radial lines



$(k_1, k_2) \sim (k_1^2 + k_2^2)^{-1/2}$

$(k_1, k_2) \sim (k_1^2 + k_2^2)^{-1}$

Sketch of the proof

Theorem (Strengthened Sobolev inequality)

For all images $x \in \mathbb{R}^{N^2}$ in the null space of a matrix $\mathcal{A} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ which, when composed with the bivariate Haar wavelet transform, $\mathcal{A}\mathcal{H}^{-1} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the Restricted Isometry Property of order $2s$, the following holds:

$$\|x\|_2 \lesssim \frac{\log(N)}{\sqrt{s}} \|x\|_{TV}$$

Sketch of the proof

Theorem (Strengthened Sobolev inequality)

For all images $x \in \mathbb{R}^{N^2}$ in the null space of a matrix $\mathcal{A} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ which, when composed with the bivariate Haar wavelet transform, $\mathcal{A}\mathcal{H}^{-1} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the Restricted Isometry Property of order $2s$, the following holds:

$$\|x\|_2 \lesssim \frac{\log(N)}{\sqrt{s}} \|x\|_{TV}$$

Lemmas we use:

1. Denote the ordered bivariate Haar wavelet coefficients of mean-zero $x \in \mathbb{R}^{N^2}$ by $c_{(1)} \geq c_{(2)} \geq \cdots \geq c_{(N^2)}$. Then[3]

$$|c_{(k)}| \lesssim \frac{\|x\|_{TV}}{k}$$

   That is, the sequence is in weak-ℓ_1.

2. If $c = c_{S_0} + c_{S_1} + c_{S_2} + \ldots$ where $c_{S_0} = (c_{(1)}, c_{(2)}, \ldots, c_{(s)}, 0, \ldots, 0)$, and $c_{S_1} = (0, \ldots, 0, c_{(s+1)}, \ldots, c_{(2s)}, 0, \ldots, 0)$, and so on are s-sparse, then $\|c_{S_{j+1}}\|_2 \leq \frac{1}{\sqrt{s}} \|c_{S_j}\|_1$.

3. Recall that $\mathcal{A}\mathcal{H}^{-1} : \mathbb{R}^{N^2} \to \mathbb{R}^m$ having the RIP of order s means that:

$$\frac{3}{4}\|x\|_2 \leq \|\mathcal{A}\mathcal{H}^{-1}x\|_2 \leq \frac{5}{4}\|x\|_2 \quad \forall s\text{-sparse } x.$$

[3] Cohen, Devore, Petrushev, Xu, 1999

Suppose that $\Psi = \mathcal{A}\mathcal{H}^* : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the RIP of order $2s$.

Suppose $\mathcal{A}x = 0$

Decompose $c = \mathcal{H}x$ into ordered s-sparse blocks
$c = c_{S_0} + c_{S_1} + c_{S_2} + \dots$

Then $\Psi c = \mathcal{A}\mathcal{H}^*\mathcal{H}x = \mathcal{A}x = 0$ and

$$
\begin{aligned}
0 &\geq \|\Psi(c_{S_0} + c_{S_1})\|_2 - \sum_{j\geq 2} \|\Psi c_{S_j}\|_2 \\
\text{(RIP of } \Psi\text{)} \qquad &\geq \frac{3}{4}\|c_{S_0} + c_{S_1}\|_2 - \frac{5}{4}\sum_{j\geq 2} \|c_{S_j}\|_2 \\
\text{(block trick)} \qquad &\geq \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\sum_{j\geq 1} \|c_{S_j}\|_1 \\
\text{(c in weak } \ell_1\text{)} \qquad &\geq \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\|x\|_{TV}\log(N^2/s)
\end{aligned}
$$

Suppose that $\Psi = \mathcal{A}\mathcal{H}^* : \mathbb{R}^{N^2} \to \mathbb{R}^m$ has the RIP of order $2s$.

Suppose $\mathcal{A}x = 0$

Decompose $c = \mathcal{H}x$ into ordered s-sparse blocks
$c = c_{S_0} + c_{S_1} + c_{S_2} + \dots$

Then $\Psi c = \mathcal{A}\mathcal{H}^*\mathcal{H}x = \mathcal{A}x = 0$ and

$$
\begin{aligned}
0 &\geq \|\Psi(c_{S_0} + c_{S_1})\|_2 - \sum_{j\geq 2} \|\Psi c_{S_j}\|_2 \\
\text{(RIP of } \Psi) \qquad &\geq \frac{3}{4}\|c_{S_0} + c_{S_1}\|_2 - \frac{5}{4}\sum_{j\geq 2} \|c_{S_j}\|_2 \\
\text{(block trick)} \qquad &\geq \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\sum_{j\geq 1} \|c_{S_j}\|_1 \\
\text{(c in weak } \ell_1) \qquad &\geq \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\|x\|_{TV}\log(N^2/s)
\end{aligned}
$$

So

1. $\|c_{S_0}\|_2 \leq \frac{1}{\sqrt{s}}\|x\|_{TV} \log(N/s)$,
2. $\|c - c_{S_0}\|_2 \leq \frac{1}{\sqrt{s}}\|x\|_{TV}$ (c is in weak ℓ_1)

Then

$$\|x\|_2 = \|c\|_2 \leq \|c_{S_0}\|_2 + \|c - c_{S_0}\|_2 \leq \frac{\log(N^2/s)}{\sqrt{s}}\|x\|_{TV}$$

Extensions

1. Strengthened Sobolev inequalities and total variation minimization guarantees can be extended to multidimensional signals.
2. Strengthened Sobolev inequalities and total variation minimization guarantees are robust to noisy measurements: $y = \mathcal{A}x + \xi$, with $\|\xi\|_2 \leq \varepsilon$.
3. Sobolev inequalities for random subspaces in dimension $d = 1$? Error guarantees for total variation minimization in dimension $d = 1$?
4. Can we remove the extra factor of $\log(N)$ in the derived Sobolev inequalities?
5. Deterministic constructions of subspaces satisfying strengthened Sobolev inequalities (warning: hard!!!)
6. ...

References

- Needell, Ward. "Stable image reconstruction using total variation minimization." *SIAM Journal on Imaging Sciences* 6.2 (2013): 1035-1058.
- Needell, Ward. "Near-optimal compressed sensing guarantees for total variation minimization." *IEEE transactions on image processing* 22.10 (2013): 3941.
- Krahmer, Ward, "Beyond incoherence: stable and robust sampling strategies for compressive imaging". arXiv preprint arXiv:1210.2380 (2012).