# Reconstruction of Sparse Signals by Minimizing a Re-Weighted Approximate $\ell_0$-Norm in the Null Space of the Measurement Matrix

J. K. Pant, W.-S. Lu, and A. Antoniou

University of Victoria

August 3, 2010

# Outline

- Comperssive Sensing

# Outline

- Comperssive Sensing
- Signal Recovery by $\ell_1$ Minimization

# Outline

- Comperssive Sensing
- Signal Recovery by $\ell_1$ Minimization
- Signal Recovery by $\ell_p$ Minimization with $p < 1$

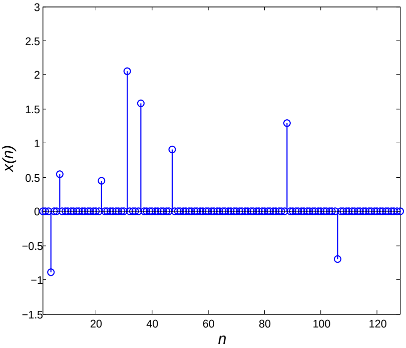
# Outline

- Comperssive Sensing
- Signal Recovery by $\ell_1$ Minimization
- Signal Recovery by $\ell_p$ Minimization with $p < 1$
- Performance Evaluation

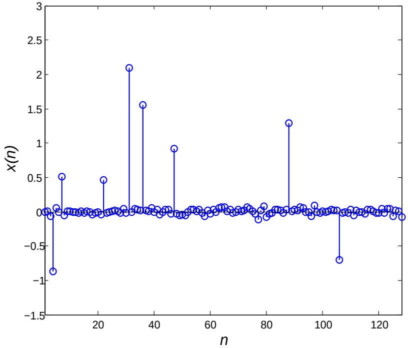
# Compressive Sensing

- A signal $\mathbf{x}(n)$ of length $N$ is $K$-sparse if it contains $K$ nonzero components with $K \ll N$.

- A signal $\mathbf{x}(n)$ of length $N$ is $K$-sparse if it contains $K$ nonzero components with $K \ll N$.
- A signal is near $K$-sparse if it contains $K$ significant components.




A sparse signal



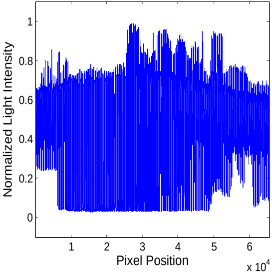

A near sparse signal

# Compressive Sensing, cont'd

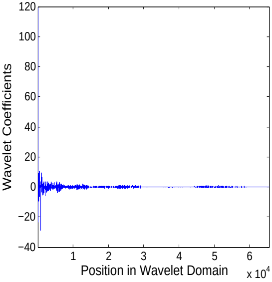
- Sparsity is a generic property of signals: A real-world signal always has a sparse or near-sparse representation with respect to an appropriate basis.

- Sparsity is a generic property of signals: A real-world signal always has a sparse or near-sparse representation with respect to an appropriate basis.

An Image




An equivalent 1-D signal



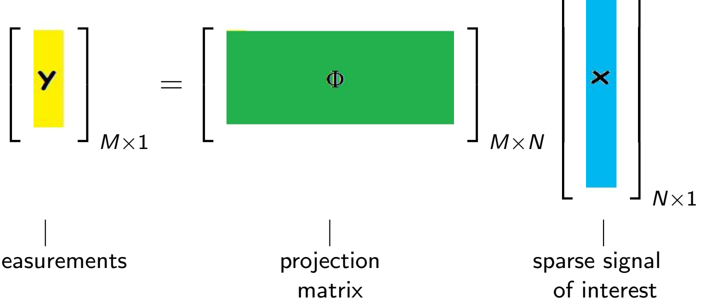

A wavelet representation of the image

# Compressive Sensing, cont'd

- Compressive sensing (CS) is a data acquisition process whereby a sparse signal $\mathbf{x}(n)$ represented by a vector $\mathbf{x}$ of length $N$ is determined using a small number of projections represented by a matrix $\mathbf{\Phi}$ of dimension $M \times N$.

- Compressive sensing (CS) is a data acquisition process whereby a sparse signal $\mathbf{x}(n)$ represented by a vector $\mathbf{x}$ of length $N$ is determined using a small number of projections represented by a matrix $\mathbf{\Phi}$ of dimension $M \times N$.
- In such a process, measurement vector $\mathbf{y}$ and signal vector $\mathbf{x}$ are interrelated by the equation

$$\mathbf{y} = \mathbf{\Phi} \cdot \mathbf{x}$$



## Compressive Sensing, cont'd

- CS theory shows that these random projections contain much, sometimes all, the information content of signal $\mathbf{x}$.

- CS theory shows that these random projections contain much, sometimes all, the information content of signal $\mathbf{x}$.
- If a sufficient number of such measurements are collected, recovering signal $\mathbf{x}$ from measurements $\mathbf{y}$ is possible.

- CS theory shows that these random projections contain much, sometimes all, the information content of signal $\mathbf{x}$.
- If a sufficient number of such measurements are collected, recovering signal $\mathbf{x}$ from measurements $\mathbf{y}$ is possible.
- A condition for this to be possible is

$$M \geq c \cdot K \cdot \log(N/K)$$

where $c$ is a small constant.

- CS theory shows that these random projections contain much, sometimes all, the information content of signal $\mathbf{x}$.
- If a sufficient number of such measurements are collected, recovering signal $\mathbf{x}$ from measurements $\mathbf{y}$ is possible.
- A condition for this to be possible is

$$M \geq c \cdot K \cdot \log(N/K)$$

where $c$ is a small constant.

- Typically,

$$K < M < N$$

- Recovering signal vector **x** from measurement vector **y** such that

$$\underset{M\times N}{\boldsymbol{\Phi}} \cdot \underset{N\times 1}{\mathbf{x}} = \underset{M\times 1}{\mathbf{y}}$$

is an ill-posed problem.

# Signal Recovery by $\ell_1$ Minimization

- Recovering signal vector $\mathbf{x}$ from measurement vector $\mathbf{y}$ such that

$$\underset{M \times N}{\boldsymbol{\Phi}} \cdot \underset{N \times 1}{\mathbf{x}} = \underset{M \times 1}{\mathbf{y}}$$

  is an ill-posed problem.

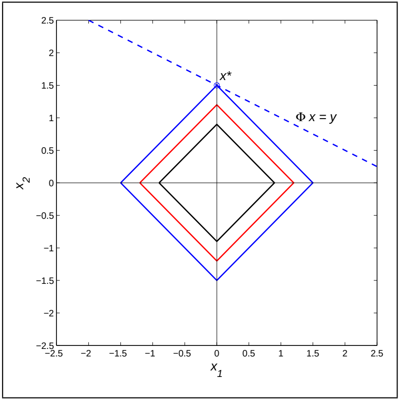
- Given that $\mathbf{x}$ is sparse, $\mathbf{x}$ can be reconstructed by solving the $\ell_1$-minimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & ||\mathbf{x}||_1 \\ \text{subject to} \quad & \boldsymbol{\Phi}\mathbf{x} = \mathbf{y} \end{aligned}$$

  where $||\mathbf{x}||_1 = \sum_{i=1}^{N} |x_i|$.

- Why $\ell_1$-norm minimization?

- Why $\ell_1$-norm minimization?




Contours for $||\mathbf{x}||_1 = c$

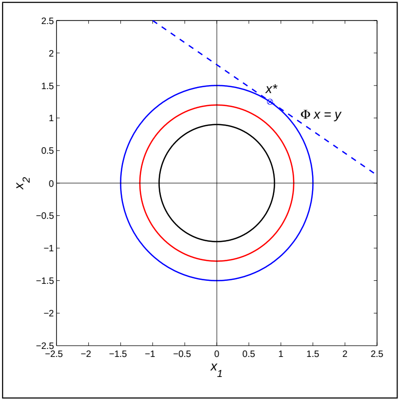
As $c$ increases, the contour of $||\mathbf{x}||_1 = c$ grows and touches the hyperplane $\mathbf{\Phi x} = \mathbf{y}$, yielding a sparse solution

$$\mathbf{x}^* = \begin{bmatrix} 0 \\ c \end{bmatrix}$$

- Why $\ell_2$-norm minimization fails to work?

# Signal Recovery by $\ell_1$ Minimization, cont'd

- Why $\ell_2$-norm minimization fails to work?




Contours of $||\mathbf{x}||_2 = r$

As $r$ increases, the contour of $||\mathbf{x}||_2 = r$ grows and touches the hyperplane $\mathbf{\Phi x} = \mathbf{y}$.

The solution $\mathbf{x}^*$ obtained is not sparse.

# Signal Recovery by $\ell_1$ Minimization, cont'd

**Theorem**

If $\mathbf{\Phi} = \{\phi_{ij}\}$ where $\phi_{ij}$ are independent and identically distributed random variables with zero-mean and variance $1/N$ and $M \geq cK \log(N/K)$, the solution of the $\ell_1$-minimization problem would recover exactly a $K$-sparse signal with high probability.

# Signal Recovery by $\ell_1$ Minimization, cont'd

**Theorem**

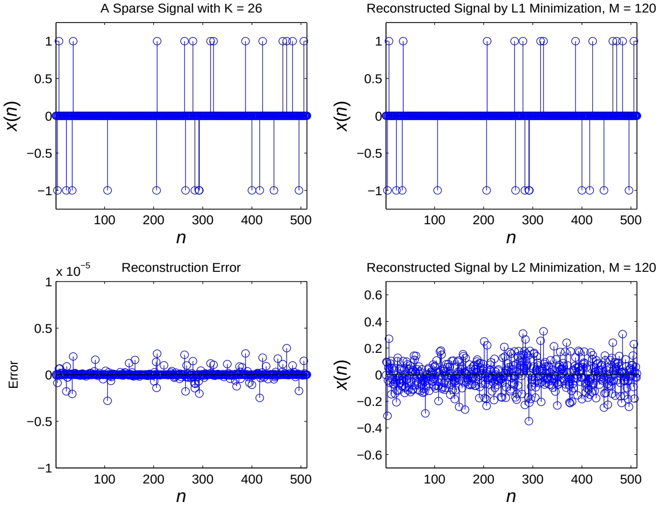
If $\mathbf{\Phi} = \{\phi_{ij}\}$ where $\phi_{ij}$ are independent and identically distributed random variables with zero-mean and variance $1/N$ and $M \geq cK \log(N/K)$, the solution of the $\ell_1$-minimization problem would recover exactly a $K$-sparse signal with high probability.

- For real-valued data $\{\mathbf{\Phi}, \mathbf{y}\}$, the $\ell_1$-minimization problem is a linear programming problem.

# Signal Recovery by $\ell_1$ Minimization, cont'd

- Example: $N = 512$, $M = 120$, $K = 26$

- Example: $N = 512$, $M = 120$, $K = 26$



# Signal Recovery by $\ell_p$ Minimization

- The sparsity of a signal can be measured by using its $\ell_0$ pseudonorm

$$||\mathbf{x}||_0 = \sum_{i=1}^{N} |x_i|^0$$

# Signal Recovery by $\ell_p$ Minimization

- The sparsity of a signal can be measured by using its $\ell_0$ pseudonorm

$$||\mathbf{x}||_0 = \sum_{i=1}^{N} |x_i|^0$$

- Hence the sparsest solution of $\mathbf{\Phi x} = \mathbf{y}$ can be obtained by solving the $\ell_0$-norm minimization problem

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & ||\mathbf{x}||_0 \\ \text{subject to} & \mathbf{\Phi x} = \mathbf{y} \end{array}$$

# Signal Recovery by $\ell_p$ Minimization

- The sparsity of a signal can be measured by using its $\ell_0$ pseudonorm

$$||\mathbf{x}||_0 = \sum_{i=1}^{N} |x_i|^0$$

- Hence the sparsest solution of $\mathbf{\Phi x} = \mathbf{y}$ can be obtained by solving the $\ell_0$-norm minimization problem

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & ||\mathbf{x}||_0 \\ \text{subject to} & \mathbf{\Phi x} = \mathbf{y} \end{array}$$

- Unfortunately, the $\ell_0$-norm minimization problem is nonconvex with combinatorial complexity.

# Signal Recovery by $\ell_p$ Minimization, cont'd

- An effective signal recovery strategy is to solve the $\ell_p$-minimization problem

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & ||\mathbf{x}||_p^p \quad \text{with} \quad 0 < p < 1 \\ \text{subject to} & \mathbf{\Phi x} = \mathbf{y} \end{array}$$

where $||\mathbf{x}||_p^p = \sum_{i=1}^{N} |x_i|^p$.

- An effective signal recovery strategy is to solve the $\ell_p$-minimization problem

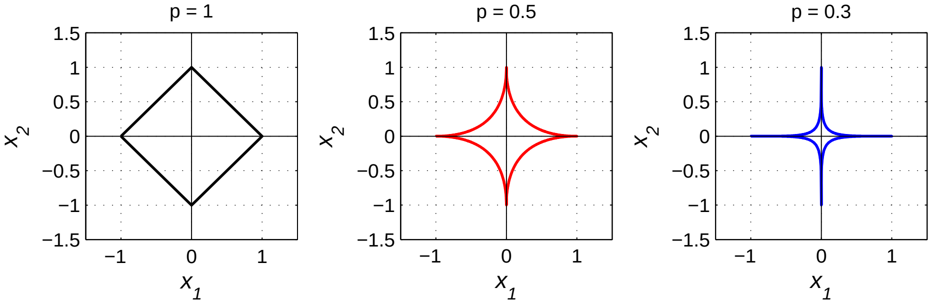
$$\begin{array}{llll} \underset{\mathbf{x}}{\text{minimize}} & ||\mathbf{x}||_p^p & \text{with} & 0 < p < 1 \\ \text{subject to} & & \mathbf{\Phi x} = \mathbf{y} & \end{array}$$

where $||\mathbf{x}||_p^p = \sum_{i=1}^{N} |x_i|^p$.

- The $\ell_p$-norm minimization problem is nonconvex.

# Signal Recovery by $\ell_p$ Minimization, cont'd

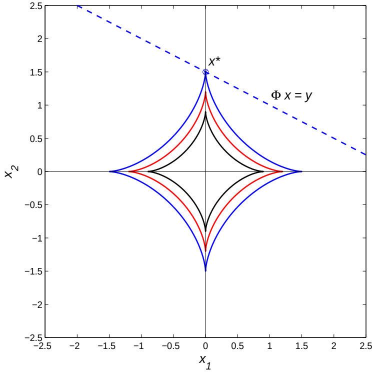
- Contours of $||\mathbf{x}||_p = 1$ with $p < 1$




$||\mathbf{x}||_1 = 1$ $||\mathbf{x}||_{0.5} = 1$ $||\mathbf{x}||_{0.3} = 1$

- Why $\ell_p$ minimization with $p < 1$?

- Why $\ell_p$ minimization with $p < 1$?




Contours of $||\mathbf{x}||_p^p = c$ with $p < 1$

As $c$ increases, the contour $||\mathbf{x}||_p^p = c$ grows and touches the hyperplane $\mathbf{\Phi x} = \mathbf{y}$, yielding a sparse solution

$$\mathbf{x}^* = \begin{bmatrix} 0 \\ c \end{bmatrix}.$$

The possibility that the contour will touch the hyperplane at another point is eliminated.

# Signal Recovery by $\ell_p$ Minimization, cont'd

- We propose to minimize an approximate $\ell_0$-norm

$$||\mathbf{x}||_{0,\sigma} = \sum_{i=1}^{N} \left(1 - e^{-x_i^2/2\sigma^2}\right)$$

where $\mathbf{x}$ lies in the solution space of $\mathbf{\Phi x} = \mathbf{y}$, namely,

$$\mathbf{x} = \mathbf{x}_s + \mathbf{V}_r \xi$$

where $\mathbf{x}_s$ is a solution of $\mathbf{\Phi x} = \mathbf{y}$ and $\mathbf{V}_r$ is an orthonormal basis of the null space of $\mathbf{\Phi}$.

- Why norm $||\mathbf{x}||_{0,\sigma}$ works?

# Signal Recovery by $\ell_p$ Minimization, cont'd

- Why norm $||\mathbf{x}||_{0,\sigma}$ works?

  With $\sigma$ small,

$$\left(1 - e^{-x_i^2/2\sigma^2}\right)\Big|_{x_i=0} = 0$$

  and

$$\left(1 - e^{-x_i^2/2\sigma^2}\right)\Big|_{x_i\neq 0} \approx 1$$

- Why norm $||\mathbf{x}||_{0,\sigma}$ works?

  With $\sigma$ small,

$$\left(1 - e^{-x_i^2/2\sigma^2}\right)\Big|_{x_i=0} = 0$$

  and

$$\left(1 - e^{-x_i^2/2\sigma^2}\right)\Big|_{x_i\neq 0} \approx 1$$

  Therefore, for a $K$-sparse signal,

$$||\mathbf{x}||_{0,\sigma} = \sum_{i=1}^{N} \left(1 - e^{-x_i^2/2\sigma^2}\right) \approx K = ||\mathbf{x}||_0$$

# Signal Recovery by $\ell_p$ Minimization, cont'd

- Improved recovery rate can be achieved by using a re-weighting technique.

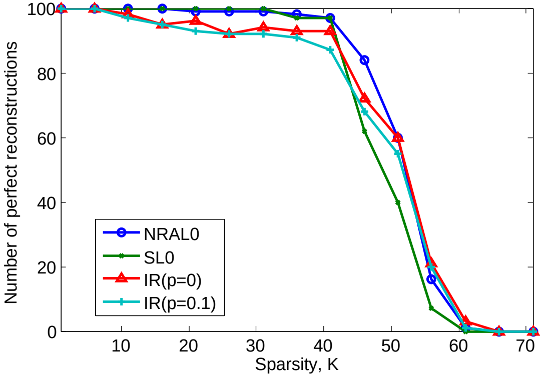
- Improved recovery rate can be achieved by using a re-weighting technique.
- This involves solving the optimization problem

$$\underset{\boldsymbol{\xi}}{\text{minimize}} \sum_{i=1}^{n} w_i \left\{ 1 - e^{-[x_s(i) + \mathbf{v}_i^T \boldsymbol{\xi}]^2 / 2\sigma^2} \right\}$$

where

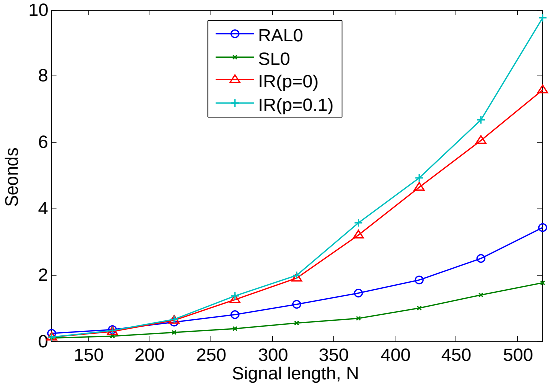
$$w_i^{(k+1)} = \frac{1}{|x_i^{(k)}| + \epsilon}$$

Number of perfectly recovered instances versus sparsity $K$ by various algorithms with $N = 256$ and $M = 100$ over 100 runs.



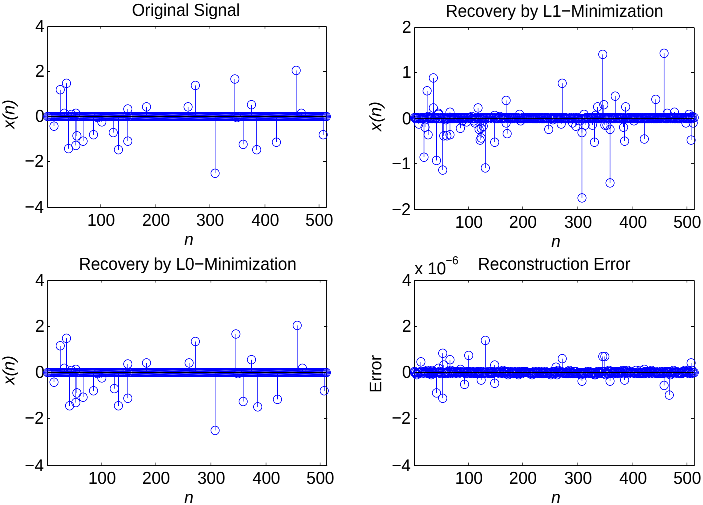

IR: Iterative re-weighting (Chartrand and Yin, 2008)
SL0: Smoothed $\ell_0$-norm minimization (Mohimani et. al., 2009)
NRAL0: Proposed

Average CPU time versus signal length for various algorithms with $M = N/2$ and $K = M/2.5$.




IR: Iterative re-weighting (Chartrand and Yin, 2008)
SL0: Smoothed $\ell_0$-norm minimization (Mohimani et. al., 2009)
NRAL0: Proposed

Performance comparison of $\ell_1$ minimization with approximate $\ell_0$ minimization for $N = 512$, $M = 80$, $K = 30$.



# Conclusions

- Compressive sensing is an effective technique for signal sampling.

# Conclusions

- Compressive sensing is an effective technique for signal sampling.
- $\ell_1$ minimization works in general for the reconstruction of sparse signals.

# Conclusions

- Compressive sensing is an effective technique for signal sampling.
- $\ell_1$ minimization works in general for the reconstruction of sparse signals.
- $\ell_p$ minimization with $p < 1$ can improve the recovery performance for signals that are less sparse.

# Conclusions

- Compressive sensing is an effective technique for signal sampling.
- $\ell_1$ minimization works in general for the reconstruction of sparse signals.
- $\ell_p$ minimization with $p < 1$ can improve the recovery performance for signals that are less sparse.
- Approximate $\ell_0$-norm minimization offers good performance with improved complexity.

*Thank you for your attention.*