Signal Recovery Method for Compressive Sensing Using Relaxation and Second-Order Cone Programming

Flávio C. A. Teixeira  Stuart W. A. Bergen  Andreas Antoniou

Department of Electrical and Computer Engineering
University of Victoria
Victoria, BC, Canada

2011 IEEE International Symposium on Circuits and Systems
Compressive sensing (CS) is a process of representing a large signal by a small number of measurements. The price that must be paid for compact signal representation is a nontrivial signal recovery process. The recovery process can be formulated as an undetermined least-squares problem where the solution is known to be sparse. The solution sparsity assumption is based on the fact that most practical signals can be represented concisely in a transform domain.
Compressive sensing (CS) is a process of representing a large signal by a small number of measurements.

The price that must be paid for compact signal representation is a nontrivial signal recovery process.

- The recovery process can be formulated as an undetermined least-squares problem where the solution is known to be sparse.

- The solution sparsity assumption is based on the fact that most practical signals can be represented concisely in a transform domain.
Compressive sensing (CS) is a process of representing a large signal by a small number of measurements.

The price that must be paid for compact signal representation is a nontrivial signal recovery process.

- The recovery process can be formulated as an undetermined least-squares problem where the solution is known to be sparse.

- The solution sparsity assumption is based on the fact that most practical signals can be represented concisely in a transform domain.
Compressive sensing (CS) is a process of representing a large signal by a small number of measurements.

The price that must be paid for compact signal representation is a nontrivial signal recovery process.

- The recovery process can be formulated as an undetermined least-squares problem where the solution is known to be sparse.

- The solution sparsity assumption is based on the fact that most practical signals can be represented concisely in a transform domain.
Motivation

- Widely known methods for signal recovery such as the $\ell_1$-Magic method promote sparsity by means of the $\ell_1$ norm:
  - Preferred sparsity promoting functions such as the $\ell_0$ norm are computationally intractable for large signals.
- We propose a new signal recovery method for CS using the smoothly clipped absolute deviation (SCAD) function as an alternative to the $\ell_0$ norm to promote sparsity.
- The resulting nonsmooth and nonconvex constrained optimization problem that must be solved to perform signal recovery is relaxed by:
  - Obtaining a series of local linear approximations of the SCAD, which results in a series of nonsmooth convex subproblems.
  - Reformulating each subproblem as a smooth second-order cone programming problem (SOCP).
Motivation

- Widely known methods for signal recovery such as the $\ell_1$-Magic method promote sparsity by means of the $\ell_1$ norm:
  - Preferred sparsity promoting functions such as the $\ell_0$ norm are computationally intractable for large signals.
- We propose a new signal recovery method for CS using the smoothly clipped absolute deviation (SCAD) function as an alternative to the $\ell_0$ norm to promote sparsity.
- The resulting nonsmooth and nonconvex constrained optimization problem that must be solved to perform signal recovery is relaxed by:
  - Obtaining a series of local linear approximations of the SCAD, which results in a series of nonsmooth convex subproblems.
  - Reformulating each subproblem as a smooth second-order cone programming problem (SOCP).
Motivation

- Widely known methods for signal recovery such as the $\ell_1$-Magic method promote sparsity by means of the $\ell_1$ norm:
  - Preferred sparsity promoting functions such as the $\ell_0$ norm are computationally intractable for large signals.

- We propose a new signal recovery method for CS using the smoothly clipped absolute deviation (SCAD) function as an alternative to the $\ell_0$ norm to promote sparsity.

- The resulting nonsmooth and nonconvex constrained optimization problem that must be solved to perform signal recovery is relaxed by:
  - Obtaining a series of local linear approximations of the SCAD, which results in a series of nonsmooth convex subproblems.
  - Reformulating each subproblem as a smooth second-order cone programming problem (SOCP).
Motivation

- Widely known methods for signal recovery such as the $\ell_1$-Magic method promote sparsity by means of the $\ell_1$ norm:
  - Preferred sparsity promoting functions such as the $\ell_0$ norm are computationally intractable for large signals.

- We propose a new signal recovery method for CS using the smoothly clipped absolute deviation (SCAD) function as an alternative to the $\ell_0$ norm to promote sparsity.

- The resulting nonsmooth and nonconvex constrained optimization problem that must be solved to perform signal recovery is relaxed by:
  1. Obtaining a series of local linear approximations of the SCAD, which results in a series of nonsmooth convex subproblems.
  2. Reformulating each subproblem as a smooth second-order cone programming problem (SOCP).
A method for sparse-signal recovery

Motivation

- Widely known methods for signal recovery such as the $\ell_1$-Magic method promote sparsity by means of the $\ell_1$ norm:
  - Preferred sparsity promoting functions such as the $\ell_0$ norm are computationally intractable for large signals.
- We propose a new signal recovery method for CS using the smoothly clipped absolute deviation (SCAD) function as an alternative to the $\ell_0$ norm to promote sparsity.
- The resulting nonsmooth and nonconvex constrained optimization problem that must be solved to perform signal recovery is relaxed by:
  1. Obtaining a series of local linear approximations of the SCAD, which results in a series of nonsmooth convex subproblems.
  2. Reformulating each subproblem as a smooth second-order cone programming problem (SOCP).
Motivation

- Widely known methods for signal recovery such as the $\ell_1$-Magic method promote sparsity by means of the $\ell_1$ norm:
  - Preferred sparsity promoting functions such as the $\ell_0$ norm are computationally intractable for large signals.
- We propose a new signal recovery method for CS using the smoothly clipped absolute deviation (SCAD) function as an alternative to the $\ell_0$ norm to promote sparsity.
- The resulting nonsmooth and nonconvex constrained optimization problem that must be solved to perform signal recovery is relaxed by:
  1. Obtaining a series of local linear approximations of the SCAD, which results in a series of nonsmooth convex subproblems.
  2. Reformulating each subproblem as a smooth second-order cone programming problem (SOCP).
Sparse Representation

- A vector $f$ of length $n$ represents the original signal.
- Vector $a$ of the same length represents a sparse or compressed version of the signal over an appropriate basis.
- This representation is obtained by using the linear operation $a = \Psi^T f$ where $\Psi \in \mathbb{R}^{n \times n}$ is orthonormal.
- The operation is reversible and the original signal $f$ can be exactly recovered from $a$ by using the relation $f = \Psi a$.
- Vector $a$ has only $s$ nonzero values with $s < n$. 
Sparse Representation

- A vector $\mathbf{f}$ of length $n$ represents the original signal.
- Vector $\mathbf{a}$ of the same length represents a sparse or compressed version of the signal over an appropriate basis.
- This representation is obtained by using the linear operation $\mathbf{a} = \Psi^T \mathbf{f}$ where $\Psi \in \mathbb{R}^{n \times n}$ is orthonormal.
- The operation is reversible and the original signal $\mathbf{f}$ can be exactly recovered from $\mathbf{a}$ by using the relation $\mathbf{f} = \Psi \mathbf{a}$.
- Vector $\mathbf{a}$ has only $s$ nonzero values with $s < n$. 
A vector $f$ of length $n$ represents the original signal.

Vector $a$ of the same length represents a sparse or compressed version of the signal over an appropriate basis.

This representation is obtained by using the linear operation $a = \Psi^T f$ where $\Psi \in \mathbb{R}^{n \times n}$ is orthonormal.

The operation is reversible and the original signal $f$ can be exactly recovered from $a$ by using the relation $f = \Psi a$.

Vector $a$ has only $s$ nonzero values with $s < n$. 
Sparse Representation

- A vector \( f \) of length \( n \) represents the original signal.
- Vector \( a \) of the same length represents a sparse or compressed version of the signal over an appropriate basis.
- This representation is obtained by using the linear operation \( a = \Psi^T f \) where \( \Psi \in \mathbb{R}^{n \times n} \) is orthonormal.
- The operation is reversible and the original signal \( f \) can be exactly recovered from \( a \) by using the relation \( f = \Psi a \).
- Vector \( a \) has only \( s \) nonzero values with \( s < n \).
Sparse Representation

- A vector \( f \) of length \( n \) represents the original signal.
- Vector \( a \) of the same length represents a sparse or compressed version of the signal over an appropriate basis.
- This representation is obtained by using the linear operation \( a = \Psi^T f \) where \( \Psi \in \mathbb{R}^{n \times n} \) is orthonormal.
- The operation is reversible and the original signal \( f \) can be exactly recovered from \( a \) by using the relation \( f = \Psi a \).
- Vector \( a \) has only \( s \) nonzero values with \( s < n \).
The measurement of the original signal is usually performed directly in the $\Psi$ domain in the presence of measurement noise $z$.

- $z$ has a known power bound $\varepsilon$ of the form $\|z\|_2 \leq \varepsilon$.
- The sensing operation in this context is given by $b = \Theta a + z$.
  - $\Theta \in \mathbb{R}^{q \times n}$ denotes a sensing matrix.
  - The entries of $\Theta$ are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $1/q$.
  - Vector $b$ of length $q$ represents the noisy measurements.

- The original signal $f$ must be recovered from a significantly reduced number of measurements $b$ such that $q \ll n$. 
The measurement of the original signal is usually performed directly in the $\Psi$ domain in the presence of measurement noise $z$.

$z$ has a known power bound $\varepsilon$ of the form $\|z\|_{\ell_2} \leq \varepsilon$.

The sensing operation in this context is given by $b = \Theta a + z$.

- $\Theta \in \mathbb{R}^{q \times n}$ denotes a sensing matrix.
- The entries of $\Theta$ are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $1/q$.
- Vector $b$ of length $q$ represents the noisy measurements.

The original signal $f$ must be recovered from a significantly reduced number of measurements $b$ such that $q \ll n$. 
The measurement of the original signal is usually performed directly in the $\Psi$ domain in the presence of measurement noise $z$.

$z$ has a known power bound $\varepsilon$ of the form $\|z\|_{\ell_2} \leq \varepsilon$.

The sensing operation in this context is given by $b = \Theta a + z$.

- $\Theta \in \mathbb{R}^{q \times n}$ denotes a sensing matrix.
- The entries of $\Theta$ are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $1/q$.
- Vector $b$ of length $q$ represents the noisy measurements.

The original signal $f$ must be recovered from a significantly reduced number of measurements $b$ such that $q \ll n$. 
Noisy Measurements

- The measurement of the original signal is usually performed directly in the $\Psi$ domain in the presence of measurement noise $z$.
- $z$ has a known power bound $\varepsilon$ of the form $\|z\|_{\ell_2} \leq \varepsilon$.
- The sensing operation in this context is given by $b = \Theta a + z$.
  - $\Theta \in \mathbb{R}^{q \times n}$ denotes a sensing matrix.
  - The entries of $\Theta$ are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $1/q$.
  - Vector $b$ of length $q$ represents the noisy measurements.
- The original signal $f$ must be recovered from a significantly reduced number of measurements $b$ such that $q \ll n$. 
Noisy Measurements

- The **measurement** of the original signal is usually performed directly in the $\Psi$ domain in the presence of measurement **noise** $z$.

- $z$ has a known **power bound** $\varepsilon$ of the form $\|z\|_{\ell_2} \leq \varepsilon$.

- The sensing operation in this context is given by $b = \Theta a + z$.

  - $\Theta \in \mathbb{R}^{q \times n}$ denotes a sensing matrix.
  - The entries of $\Theta$ are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $1/q$.
  - Vector $b$ of length $q$ represents the noisy measurements.

- The original signal $f$ must be recovered from a **significantly reduced** number of measurements $b$ such that $q \ll n$. 
The measurement of the original signal is usually performed directly in the \( \Psi \) domain in the presence of measurement noise \( z \).

\( z \) has a known power bound \( \varepsilon \) of the form \( \|z\|_{\ell_2} \leq \varepsilon \).

The sensing operation in this context is given by \( b = \Theta a + z \).

- \( \Theta \in \mathbb{R}^{q \times n} \) denotes a sensing matrix.
- The entries of \( \Theta \) are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance \( 1/q \).
- Vector \( b \) of length \( q \) represents the noisy measurements.

The original signal \( f \) must be recovered from a significantly reduced number of measurements \( b \) such that \( q \ll n \).
The measurement of the original signal is usually performed directly in the $\Psi$ domain in the presence of measurement noise $z$.

$z$ has a known power bound $\varepsilon$ of the form $\|z\|_{\ell^2} \leq \varepsilon$.

The sensing operation in this context is given by $b = \Theta a + z$.

- $\Theta \in \mathbb{R}^{q \times n}$ denotes a sensing matrix.
- The entries of $\Theta$ are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $1/q$.
- Vector $b$ of length $q$ represents the noisy measurements.

The original signal $f$ must be recovered from a significantly reduced number of measurements $b$ such that $q \ll n$. 
Recovery Process: Goals

The goal of the recovery process is twofold:
1. To find the **sparsest** signal.
2. To ensure that the signal found is **consistent** with the measurements.

The **sparsity** of $f$ can be measured in terms of its transform coefficients $a$ and a function of the form:

$$P_{\tau}(a) = \sum_{i=1}^{n} p_{\tau}(|a_i|)$$

- $p_{\tau}(|a_i|)$ quantifies the magnitude of each individual coefficient of $a$.
- The minimization of $P_{\tau}(a)$ has a sparse solution.
- For this reason, we call $p_{\tau}(|a_i|)$ a sparsity promoting function.
Recovery Process: Goals

- The goal of the recovery process is twofold:
  1. To find the **sparsest** signal.
  2. To ensure that the signal found is consistent with the measurements.

- The sparsity of $f$ can be measured in terms of its transform coefficients $a$ and a function of the form:

$$P_\tau(a) = \sum_{i=1}^{n} p_\tau(|a_i|)$$

  - $p_\tau(|a_i|)$ quantifies the magnitude of each individual coefficient of $a$.

- The minimization of $P_\tau(a)$ has a sparse solution.
  - For this reason, we call $p_\tau(|a_i|)$ a sparsity promoting function.
Recovery Process: Goals

- The goal of the recovery process is twofold:
  1. To find the \textit{sparsest} signal.
  2. To ensure that the signal found is \textit{consistent} with the measurements.

- The sparsity of $f$ can be measured in terms of its transform coefficients $a$ and a function of the form:

$$P_\tau(a) = \sum_{i=1}^{n} p_\tau(|a_i|)$$

- $p_\tau(|a_i|)$ quantifies the magnitude of each individual coefficient of $a$.
- The minimization of $P_\tau(a)$ has a sparse solution.
- For this reason, we call $p_\tau(|a_i|)$ a sparsity promoting function.
The goal of the recovery process is twofold:

1. To find the \textit{sparsest} signal.
2. To ensure that the signal found is \textit{consistent} with the measurements.

The \textit{sparsity} of $f$ can be measured in terms of its transform coefficients $a$ and a function of the form:

$$P_{\tau}(a) = \sum_{i=1}^{n} p_{\tau}(|a_i|)$$

- $p_{\tau}(|a_i|)$ quantifies the magnitude of each individual coefficient of $a$.
- The minimization of $P_{\tau}(a)$ has a sparse solution.
- For this reason, we call $p_{\tau}(|a_i|)$ a \textit{sparsity promoting function}. 
Recovery Process: Goals

- The goal of the recovery process is twofold:
  1. To find the **sparsest** signal.
  2. To ensure that the signal found is **consistent** with the measurements.

- The **sparsity** of $f$ can be measured in terms of its transform coefficients $a$ and a function of the form:

$$P_\tau(a) = \sum_{i=1}^{n} p_\tau(|a_i|)$$

- $p_\tau(|a_i|)$ quantifies the magnitude of each individual coefficient of $a$.
  - The minimization of $P_\tau(a)$ has a sparse solution.
  - For this reason, we call $p_\tau(|a_i|)$ a sparsity promoting function.
Recovery Process: Goals

- The goal of the recovery process is twofold:
  1. To find the **sparsest** signal.
  2. To ensure that the signal found is **consistent** with the measurements.

- The **sparsity** of $f$ can be measured in terms of its transform coefficients $a$ and a function of the form:

$$P_\tau(a) = \sum_{i=1}^{n} p_\tau(|a_i|)$$

- $p_\tau(|a_i|)$ quantifies the magnitude of each individual coefficient of $a$.
- The minimization of $P_\tau(a)$ has a sparse solution.
  - For this reason, we call $p_\tau(|a_i|)$ a sparsity promoting function.
Recovery Process: Goals

- The goal of the recovery process is twofold:
  - To find the sparsest signal.
  - To ensure that the signal found is consistent with the measurements.
- The sparsity of \( \mathbf{f} \) can be measured in terms of its transform coefficients \( \mathbf{a} \) and a function of the form:

\[
P_\tau(\mathbf{a}) = \sum_{i=1}^{n} p_\tau(|a_i|)
\]

- \( p_\tau(|a_i|) \) quantifies the magnitude of each individual coefficient of \( \mathbf{a} \).
- The minimization of \( P_\tau(\mathbf{a}) \) has a sparse solution.
  - For this reason, we call \( p_\tau(|a_i|) \) a sparsity promoting function.
Sparse-Signal Recovery: Problem Definition

**Sparse-Signal Recovery Problem**

- The problem can be approached via two different formulations.
  - The *unconstrained* formulation (or Lagrangian Form) defined by
    \[
    \text{minimize}_{\mathbf{a}} \quad \|\Theta \mathbf{a} - \mathbf{b}\|_{\ell_2} + \frac{1}{\lambda} P_{\tau}(\mathbf{a})
    \]
  - The *constrained* formulation defined by
    \[
    \text{minimize}_{\mathbf{a}} \quad P_{\tau}(\mathbf{a}) \quad \text{subject to:} \quad \|\Theta \mathbf{a} - \mathbf{b}\|_{\ell_2} \leq \varepsilon
    \]

- Optimization theory asserts that the two problems are equivalent.
  - The constrained formulation is harder to solve.
  - The relationship between \(\varepsilon\) and \(1/\lambda\) is nontrivial.
  - It is easier to determine an appropriate \(\varepsilon\) rather than a \(\lambda\).
Sparse-Signal Recovery Problem

- The problem can be approached via two different formulations.
- The unconstrained formulation (or Lagrangian Form) defined by
  \[
  \min_a \| \Theta a - b \|_{\ell_2} + \frac{1}{\lambda} P_\tau(a)
  \]
- The constrained formulation defined by
  \[
  \min_a P_\tau(a) \quad \text{subject to:} \quad \| \Theta a - b \|_{\ell_2} \leq \varepsilon
  \]

- Optimization theory asserts that the two problems are equivalent.
- The constrained formulation is harder to solve.
- The relationship between \( \varepsilon \) and \( 1/\lambda \) is nontrivial.
- It is easier to determine an appropriate \( \varepsilon \) rather than a \( \lambda \).
The problem can be approached via two different formulations.

- The **unconstrained** formulation (or Lagrangian Form) defined by

  \[
  \min_{a} \| \Theta a - b \|_{2} + \frac{1}{\lambda} P_{\tau}(a)
  \]

- The **constrained** formulation defined by

  \[
  \min_{a} P_{\tau}(a) \quad \text{subject to: } \| \Theta a - b \|_{2} \leq \varepsilon
  \]

Optimization theory asserts that the two problems are equivalent.

- The constrained formulation is harder to solve.
- The relationship between \( \varepsilon \) and \( 1/\lambda \) is nontrivial.
- It is easier to determine an appropriate \( \varepsilon \) rather than a \( \lambda \).
Sparse-Signal Recovery Problem

- The problem can be approached via two different formulations.
  - The *unconstrained* formulation (or Lagrangian Form) defined by
    \[
    \min_{a} \| \Theta a - b \|_{\ell_2} + \frac{1}{\lambda} P_\tau(a)
    \]
  - The *constrained* formulation defined by
    \[
    \min_{a} P_\tau(a) \quad \text{subject to: } \| \Theta a - b \|_{\ell_2} \leq \varepsilon
    \]

- Optimization theory asserts that the two problems are equivalent.
  - The constrained formulation is harder to solve.
  - The relationship between $\varepsilon$ and $1/\lambda$ is nontrivial.
  - It is easier to determine an appropriate $\varepsilon$ rather than $\lambda$. 

F. Teixeira, S. Bergen, A. Antoniou (UVic)
Sparse-Signal Recovery Problem

- The problem can be approached via two different formulations.
  - The **unconstrained** formulation (or Lagrangian Form) defined by
    \[
    \min_{a} \| \Theta a - b \|_{\ell_2} + \frac{1}{\lambda} P_{\tau}(a)
    \]
  - The **constrained** formulation defined by
    \[
    \min_{a} P_{\tau}(a) \quad \text{subject to:} \quad \| \Theta a - b \|_{\ell_2} \leq \varepsilon
    \]

- Optimization theory asserts that the two problems are **equivalent**.
  1. The constrained formulation is **harder** to solve.
  2. The relationship between \( \varepsilon \) and \( 1/\lambda \) is **nontrivial**.
  3. It is **easier** to determine an appropriate \( \varepsilon \) rather than a \( \lambda \).
Sparse-Signal Recovery Problem

- The problem can be approached via two different formulations.
  - The *unconstrained* formulation (or Lagrangian Form) defined by
    \[
    \min_{\mathbf{a}} \quad \| \Theta \mathbf{a} - \mathbf{b} \|_2 + \frac{1}{\lambda} P_\tau(\mathbf{a})
    \]
  - The *constrained* formulation defined by
    \[
    \min_{\mathbf{a}} \quad P_\tau(\mathbf{a}) \quad \text{subject to:} \quad \| \Theta \mathbf{a} - \mathbf{b} \|_2 \leq \varepsilon
    \]

- Optimization theory asserts that the two problems are equivalent.
  1. The constrained formulation is **harder** to solve.
  2. The relationship between \( \varepsilon \) and \( 1/\lambda \) is **nontrivial**.
  3. It is easier to determine an appropriate \( \varepsilon \) rather than a \( \lambda \).
Sparse-Signal Recovery Problem

The problem can be approached via two different formulations.

The unconstrained formulation (or Lagrangian Form) defined by

$$\min_{a} \| \Theta a - b \|_{2} + \frac{1}{\lambda} P_{\tau}(a)$$

The constrained formulation defined by

$$\min_{a} P_{\tau}(a) \quad \text{subject to: } \| \Theta a - b \|_{2} \leq \varepsilon$$

Optimization theory asserts that the two problems are equivalent.

1. The constrained formulation is harder to solve.
2. The relationship between $\varepsilon$ and $1/\lambda$ is nontrivial.
3. It is easier to determine an appropriate $\varepsilon$ rather than a $\lambda$. 

F. Teixeira, S. Bergen, A. Antoniou (UVic)
On the sparsest solution of the recovery problem

On the sparsest solution of the recovery problem

Obtaining the Sparsest Solution

- The **sparsest solution** for the two problems can be obtained when $p_\tau(|a_i|) = \tau|a_i|^p$ and $p = 0$, i.e., by computing the $\ell_0$ norm of $a$.
  - Unfortunately, the use of the $\ell_0$ norm in the two problems requires an intractable combinatorial search for large signals.
- Past work in CS has shown that when certain conditions on the transform matrix $\Psi$ and measurement matrix $\Theta$ are met:
  - We are able to recover $f$ from $b$ by using $p_\tau(|a_i|) = \tau|a_i|$ as the sparsity promoting function, i.e., by computing the $\ell_1$ norm of $a$.
  - The price that must be paid for this approximation is that more measurements $q$ are required to recover $f$ than when using the $\ell_0$ norm.
Obtaining the Sparsest Solution

- The **sparsest solution** for the two problems can be obtained when $p_r(|a_i|) = \tau |a_i|^p$ and $p = 0$, i.e., by computing the $\ell_0$ norm of $a$.
- Unfortunately, the use of the $\ell_0$ norm in the two problems requires an intractable combinatorial search for large signals.

Past work in CS has shown that when certain conditions on the transform matrix $\Psi$ and measurement matrix $\Theta$ are met:

- We are able to recover $f$ from $b$ by using $p_r(|a_i|) = \tau |a_i|$ as the sparsity promoting function, i.e., by computing the $\ell_1$ norm of $a$.
- The price that must be paid for this approximation is that more measurements $q$ are required to recover $f$ than when using the $\ell_0$ norm.
On the sparsest solution of the recovery problem

Obtaining the Sparsest Solution

- The **sparsest solution** for the two problems can be obtained when $p_\tau(|a_i|) = \tau |a_i|^p$ and $p = 0$, i.e., by computing the $\ell_0$ norm of $a$.

- Unfortunately, the use of the $\ell_0$ norm in the two problems requires an **intractable** combinatorial search for large signals.

- Past work in CS has shown that when certain **conditions** on the transform matrix $\Psi$ and measurement matrix $\Theta$ are met:
  - We are able to recover $f$ from $b$ by using $p_\tau(|a_i|) = \tau |a_i|$ as the sparsity promoting function, i.e., by computing the $\ell_1$ norm of $a$.
  - The **price** that must be paid for this approximation is that more measurements $q$ are required to recover $f$ than when using the $\ell_0$ norm.
Obtaining the Sparsest Solution

The **sparsest solution** for the two problems can be obtained when $p_\tau(|a_i|) = \tau |a_i|^p$ and $p = 0$, i.e., by computing the $\ell_0$ norm of $a$.

Unfortunately, the use of the $\ell_0$ norm in the two problems requires an intractable combinatorial search for large signals.

Past work in CS has shown that when certain **conditions** on the transform matrix $\Psi$ and measurement matrix $\Theta$ are met:

- We are able to recover $f$ from $b$ by using $p_\tau(|a_i|) = \tau |a_i|$ as the sparsity promoting function, i.e., by computing the $\ell_1$ norm of $a$.
- The price that must be paid for this approximation is that more measurements $q$ are required to recover $f$ than when using the $\ell_0$ norm.
On the sparsest solution of the recovery problem

Obtaining the Sparsest Solution

- The **sparsest solution** for the two problems can be obtained when
  \( p_\tau(|a_i|) = \tau |a_i|^p \) and \( p = 0 \), i.e., by computing the \( \ell_0 \) norm of \( a \).

- Unfortunately, the use of the \( \ell_0 \) norm in the two problems requires an intractable combinatorial search for large signals.

- Past work in CS has shown that when certain **conditions** on the transform matrix \( \Psi \) and measurement matrix \( \Theta \) are met:
  - We are able to recover \( f \) from \( b \) by using \( p_\tau(|a_i|) = \tau |a_i| \) as the sparsity promoting function, i.e., by computing the \( \ell_1 \) norm of \( a \).
  - The **price** that must be paid for this approximation is that more measurements \( q \) are required to recover \( f \) than when using the \( \ell_0 \) norm.
An interesting alternative to the $\ell_0$ norm as a sparsity-promoting function is the smoothly clipped absolute deviation (SCAD) function. We are interested in using the SCAD because it performs as well as the oracle estimator for a problem similar to the unconstrained formulation for sparse-signal recovery. This means that the SCAD is asymptotically as efficient as an ideal estimator, namely, it performs as well as if the coefficients that are zero were known.
The SCAD as sparsity promoting function

SCAD Function

- An interesting alternative to the $\ell_0$ norm as a sparsity-promoting function is the smoothly clipped absolute deviation (SCAD) function.

- We are interested in using the SCAD because it performs as well as the oracle estimator for a problem similar to the unconstrained formulation for sparse-signal recovery.

- This means that the SCAD is asymptotically as efficient as an ideal estimator, namely, it performs as well as if the coefficients that are zero were known.
An interesting alternative to the $\ell_0$ norm as a sparsity-promoting function is the smoothly clipped absolute deviation (SCAD) function. We are interested in using the SCAD because it performs as well as the oracle estimator for a problem similar to the unconstrained formulation for sparse-signal recovery. This means that the SCAD is asymptotically as efficient as an ideal estimator, namely, it performs as well as if the coefficients that are zero were known.
Under the assumption that the noise level $\varepsilon$ is known in advance, it is usually more \textbf{natural} and \textbf{efficient} to solve the \textbf{constrained} version of the recovery problem instead of the unconstrained one.

Unfortunately, use of the SCAD function on the constrained version of the recovery problem has the following drawbacks:

- The objective function $P_r(a)$ is now \textbf{concave} and \textbf{nonsmooth}.
- The recovery problem becomes a \textbf{nonconvex} and \textbf{nonsmooth} constrained optimization problem.
- This means that the recovery problem is computationally \textbf{intractable} in its current form.
Using the SCAD in the Recovery Problem

- Under the assumption that the noise level $\varepsilon$ is known in advance, it is usually more natural and efficient to solve the constrained version of the recovery problem instead of the unconstrained one.

- Unfortunately, use of the SCAD function on the constrained version of the recovery problem has the following drawbacks:
  - The objective function $P_\tau(a)$ is now concave and nonsmooth.
  - The recovery problem becomes a nonconvex and nonsmooth constrained optimization problem.
  - This means that the recovery problem is computationally intractable in its current form.
Using the SCAD in the Recovery Problem

- Under the assumption that the noise level $\varepsilon$ is known in advance, it is usually more natural and efficient to solve the constrained version of the recovery problem instead of the unconstrained one.

- Unfortunately, use of the SCAD function on the constrained version of the recovery problem has the following drawbacks:
  - The objective function $P_\tau(a)$ is now concave and nonsmooth.
  - The recovery problem becomes a nonconvex and nonsmooth constrained optimization problem.
  - This means that the recovery problem is computationally intractable in its current form.
Using the SCAD in the Recovery Problem

- Under the assumption that the noise level $\varepsilon$ is known in advance, it is usually more natural and efficient to solve the constrained version of the recovery problem instead of the unconstrained one.

- Unfortunately, use of the SCAD function on the constrained version of the recovery problem has the following drawbacks:
  - The objective function $P_T(a)$ is now concave and nonsmooth.
  - The recovery problem becomes a nonconvex and nonsmooth constrained optimization problem.
  - This means that the recovery problem is computationally intractable in its current form.
Using the SCAD in the Recovery Problem

- Under the assumption that the noise level $\varepsilon$ is known in advance, it is usually more natural and efficient to solve the constrained version of the recovery problem instead of the unconstrained one.

- Unfortunately, use of the SCAD function on the constrained version of the recovery problem has the following drawbacks:
  - The objective function $P_\tau(a)$ is now concave and nonsmooth.
  - The recovery problem becomes a nonconvex and nonsmooth constrained optimization problem.
  - This means that the recovery problem is computationally intractable in its current form.
Relaxing the Objective Function of the Recovery Problem

- An effective convex approximation of $P_\tau(a)$ is based on a local linear approximation (LLA) to $p_\tau(|a_i|)$ near a point $a^{(k)}$ given by

\[
\mathcal{L}_{a^{(k)}}(a) = \sum_{i=1}^{n} \left[ p_\tau(|a_i^{(k)}|) + \frac{d}{da_i} p_\tau(|a_i^{(k)}|) (|a_i| - |a_i^{(k)}|) \right]
\]

- When $a^{(k)} \approx a$, then $\mathcal{L}_{a^{(k)}}(a) \approx P_\tau(a)$.
- Past work in statistical estimation proposed utilizing the LLA in the context of penalized likelihood models:
  - In this context, a problem similar to the unconstrained version of the recovery problem is addressed.
  - The least angle regression (LARS) algorithm is usually employed in these problems for finding the sparsest solution.
  - The LARS algorithm is known to have limited applicability when $q \ll n$. 

Relaxing the Objective Function of the Recovery Problem

- An effective convex approximation of $P_\tau(a)$ is based on a local linear approximation (LLA) to $p_\tau(|a_i|)$ near a point $a^{(k)}$ given by

$$L_{a^{(k)}}(a) = \sum_{i=1}^{n} \left[ p_\tau(|a_i^{(k)}|) + \frac{d}{da_i} p_\tau(|a_i^{(k)}|) (|a_i| - |a_i^{(k)}|) \right]$$

- When $a^{(k)} \approx a$, then $L_{a^{(k)}}(a) \approx P_\tau(a)$.

- Past work in statistical estimation proposed utilizing the LLA in the context of penalized likelihood models:
  - In this context, a problem similar to the unconstrained version of the recovery problem is addressed.
  - The least angle regression (LARS) algorithm is usually employed in these problems for finding the sparsest solution.
  - The LARS algorithm is known to have limited applicability when $q \ll n$. 
Relaxing the Objective Function of the Recovery Problem

- An effective convex approximation of $P_\tau(a)$ is based on a local linear approximation (LLA) to $p_\tau(|a_i|)$ near a point $a^{(k)}$ given by

$$L_{a^{(k)}}(a) = \sum_{i=1}^{n} \left[ p_\tau(|a_i^{(k)}|) + \frac{d}{d a_i} p_\tau(|a_i^{(k)}|) (|a_i| - |a_i^{(k)}|) \right]$$

- When $a^{(k)} \approx a$, then $L_{a^{(k)}}(a) \approx P_\tau(a)$.

- Past work in statistical estimation proposed utilizing the LLA in the context of penalized likelihood models:
  - In this context, a problem similar to the unconstrained version of the recovery problem is addressed.
  - The least angle regression (LARS) algorithm is usually employed in these problems for finding the sparsest solution.
  - The LARS algorithm is known to have limited applicability when $q \ll n$. 
Relaxing the Objective Function of the Recovery Problem

- An effective convex approximation of $P_\tau(a)$ is based on a local linear approximation (LLA) to $p_\tau(|a_i|)$ near a point $a^{(k)}$ given by

$$
\mathcal{L}_{a^{(k)}}(a) = \sum_{i=1}^{n} \left[ p_\tau(|a_i^{(k)}|) + \frac{d}{da_i} p_\tau(|a_i^{(k)}|) \left( |a_i| - |a_i^{(k)}| \right) \right]
$$

- When $a^{(k)} \approx a$, then $\mathcal{L}_{a^{(k)}}(a) \approx P_\tau(a)$.

- Past work in statistical estimation proposed utilizing the LLA in the context of penalized likelihood models:
  - In this context, a problem similar to the unconstrained version of the recovery problem is addressed.
  - The least angle regression (LARS) algorithm is usually employed in these problems for finding the sparsest solution.
  - The LARS algorithm is known to have limited applicability when $q \ll n$. 

Relaxing the Objective Function of the Recovery Problem

- An effective convex approximation of $P_{\tau}(a)$ is based on a local linear approximation (LLA) to $p_{\tau}(|a_i|)$ near a point $a^{(k)}$ given by

$$\mathcal{L}_{a^{(k)}}(a) = \sum_{i=1}^{n} \left[ p_{\tau}(|a_i^{(k)}|) + \frac{d}{da_i} p_{\tau}(|a_i^{(k)}|) (|a_i| - |a_i^{(k)}|) \right]$$

- When $a^{(k)} \approx a$, then $\mathcal{L}_{a^{(k)}}(a) \approx P_{\tau}(a)$.
- Past work in statistical estimation proposed utilizing the LLA in the context of penalized likelihood models:
  - In this context, a problem similar to the unconstrained version of the recovery problem is addressed.
  - The least angle regression (LARS) algorithm is usually employed in these problems for finding the sparsest solution.
  - The LARS algorithm is known to have limited applicability when $q \ll n$. 
Relaxing the Objective Function of the Recovery Problem

- An effective convex approximation of $P_\tau(a)$ is based on a local linear approximation (LLA) to $p_\tau(|a_i|)$ near a point $a^{(k)}$ given by

$$L_{a^{(k)}}(a) = \sum_{i=1}^{n} \left[ p_\tau(|a_i^{(k)}|) + \frac{d}{da_i} p_\tau(|a_i^{(k)}|) \left(|a_i| - |a_i^{(k)}|\right) \right]$$

- When $a^{(k)} \approx a$, then $L_{a^{(k)}}(a) \approx P_\tau(a)$.

- Past work in statistical estimation proposed utilizing the LLA in the context of penalized likelihood models:
  - In this context, a problem similar to the unconstrained version of the recovery problem is addressed.
  - The least angle regression (LARS) algorithm is usually employed in these problems for finding the sparsest solution.
  - The LARS algorithm is known to have limited applicability when $q \ll n$. 

F. Teixeira, S. Bergen, A. Antoniou (UVic)
Proposed Method for Signal Recovery

- We propose a new signal recovery method that uses the SCAD as sparsity promoting function in the constrained version of the recovery problem.
- In order to overcome nonconvexity, we relax the concave objective function $P_\tau(a)$ to its convex linear approximation:
  - This problem setting results in a sequence of convex nonsmooth constrained subproblems.
  - The sequence of solutions of these subproblems generates a monotonically decreasing sequence of values of the original concave objective function $P_\tau(a)$.
- We show that the resulting nonsmooth constrained subproblems can be formulated as smooth second-order cone programming (SOCP) subproblems.
  - This formulation is handy since each SOCP subproblem can be solved efficiently using standard state-of-the-art solvers such as SeDuMi.
Proposed Method for Signal Recovery

- We propose a new signal recovery method that uses the SCAD as sparsity promoting function in the constrained version of the recovery problem.

- In order to overcome nonconvexity, we relax the concave objective function $P_\tau(a)$ to its convex linear approximation:
  - This problem setting results in a sequence of convex nonsmooth constrained subproblems.
  - The sequence of solutions of these subproblems generates a monotonically decreasing sequence of values of the original concave objective function $P_\tau(a)$.

- We show that the resulting nonsmooth constrained subproblems can be formulated as smooth second-order cone programming (SOCP) subproblems.
  - This formulation is handy since each SOCP subproblem can be solved efficiently using standard state-of-the-art solvers such as SeDuMi.
A signal recovery method: LLA and SOCP subproblems

A Signal Recovery Method for CS

**Proposed Method for Signal Recovery**

- We propose a new signal recovery method that uses the SCAD as sparsity promoting function in the constrained version of the recovery problem.
- **In order to overcome nonconvexity**, we relax the concave objective function $P_\tau(a)$ to its convex linear approximation:
  - This problem setting results in a sequence of convex **nonsmooth** constrained subproblems.
  - The sequence of solutions of these subproblems generates a monotonically decreasing sequence of values of the original concave objective function $P_\tau(a)$.
- We show that the resulting **nonsmooth** constrained subproblems can be formulated as **smooth** second-order cone programming (SOCP) subproblems.
  - This formulation is handy since each SOCP subproblem can be solved efficiently using standard state-of-the-art solvers such as SeDuMi.
Proposed Method for Signal Recovery

- We propose a new signal recovery method that uses the SCAD as sparsity promoting function in the constrained version of the recovery problem.
- In order to overcome nonconvexity, we relax the concave objective function $P_\tau(a)$ to its convex linear approximation:
  - This problem setting results in a sequence of convex nonsmooth constrained subproblems.
  - The sequence of solutions of these subproblems generates a monotonically decreasing sequence of values of the original concave objective function $P_\tau(a)$.
- We show that the resulting nonsmooth constrained subproblems can be formulated as smooth second-order cone programming (SOCP) subproblems.
  - This formulation is handy since each SOCP subproblem can be solved efficiently using standard state-of-the-art solvers such as SeDuMi.

F. Teixeira, S. Bergen, A. Antoniou (UVic)
Proposed Method for Signal Recovery

- We propose a new signal recovery method that uses the SCAD as sparsity promoting function in the constrained version of the recovery problem.
- In order to overcome nonconvexity, we relax the concave objective function $P_{\tau}(a)$ to its convex linear approximation:
  - This problem setting results in a sequence of convex nonsmooth constrained subproblems.
  - The sequence of solutions of these subproblems generates a monotonically decreasing sequence of values of the original concave objective function $P_{\tau}(a)$.
- We show that the resulting nonsmooth constrained subproblems can be formulated as smooth second-order cone programming (SOCP) subproblems.
  - This formulation is handy since each SOCP subproblem can be solved efficiently using standard state-of-the-art solvers such as SeDuMi.
A signal recovery method: LLA and SOCP subproblems

Proposed Method for Signal Recovery

- We propose a new signal recovery method that uses the SCAD as sparsity promoting function in the constrained version of the recovery problem.
- In order to overcome nonconvexity, we relax the concave objective function $P_\tau(a)$ to its convex linear approximation:
  - This problem setting results in a sequence of convex nonsmooth constrained subproblems.
  - The sequence of solutions of these subproblems generates a monotonically decreasing sequence of values of the original concave objective function $P_\tau(a)$.
- We show that the resulting nonsmooth constrained subproblems can be formulated as smooth second-order cone programming (SOCP) subproblems.
  - This formulation is handy since each SOCP subproblem can be solved efficiently using standard state-of-the-art solvers such as SeDuMi.
Reconstruction Performance of the Proposed Method

- Reconstruction performance is usually compared in terms of the probability of perfect signal recovery (PPSR).
  - Perfect signal recovery is declared when the solution obtained for the recovery problem $a'$ is close to the true known solution $a^*$. Closeness is measured in the $l_\infty$ sense, i.e., $||a' - a^*||_{l_\infty} \leq 10^{-3}$.
  - The PPSR is estimated by performing $r$ recovery trials for a range of $s$.
- The performance of the proposed method was compared to:
  - The $\ell_1$-Magic suite of algorithms which uses the $\ell_1$ norm as the sparsity promoting function.
  - Our previous method which solves the unconstrained version of the recovery problem with a local quadratic approximation (LQA) of the SCAD.
Reconstruction Performance of the Proposed Method

- Reconstruction performance is usually compared in terms of the probability of perfect signal recovery (PPSR).
  - Perfect signal recovery is declared when the solution obtained for the recovery problem $a'$ is close to the true known solution $a^*$.
  - Closeness is measured in the $\ell_\infty$ sense, i.e., $||a' - a^*||_{\ell_\infty} \leq 10^{-3}$.
  - The PPSR is estimated by performing $r$ recovery trials for a range of $s$.
- The performance of the proposed method was compared to:
  - The $\ell_1$-Magic suite of algorithms which uses the $\ell_1$ norm as the sparsity promoting function.
  - Our previous method which solves the unconstrained version of the recovery problem with a local quadratic approximation (LQA) of the SCAD.
Reconstruction Performance of the Proposed Method

- Reconstruction performance is usually compared in terms of the probability of perfect signal recovery (PPSR).
  - Perfect signal recovery is declared when the solution obtained for the recovery problem $a'$ is close to the true known solution $a^*$.
  - Closeness is measured in the $\ell_\infty$ sense, i.e., $||a' - a^*||_{\ell_\infty} \leq 10^{-3}$.
  - The PPSR is estimated by performing $r$ recovery trials for a range of $s$.

- The performance of the proposed method was compared to:
  - The $\ell_1$-Magic suite of algorithms which uses the $\ell_1$ norm as the sparsity promoting function.
  - Our previous method which solves the unconstrained version of the recovery problem with a local quadratic approximation (LQA) of the SCAD.
Reconstruction Performance of the Proposed Method

- Reconstruction **performance** is usually compared in terms of the probability of perfect signal recovery (PPSR).
  - Perfect signal recovery is declared when the solution obtained for the recovery problem $a'$ is close to the true known solution $a^*$.
  - Closeness is measured in the $\ell_\infty$ sense, i.e., $||a' - a^*||_{\ell_\infty} \leq 10^{-3}$.
  - The PPSR is estimated by performing $r$ recovery trials for a range of $s$.

- The performance of the proposed method was compared to:
  - The $\ell_1$-Magic suite of algorithms which uses the $\ell_1$ norm as the sparsity promoting function.
  - Our previous method which solves the unconstrained version of the recovery problem with a local quadratic approximation (LQA) of the SCAD.
Reconstruction Performance of the Proposed Method

- Reconstruction **performance** is usually compared in terms of the probability of perfect signal recovery (PPSR).
  - Perfect signal recovery is declared when the solution obtained for the recovery problem $a'$ is close to the true known solution $a^*$.
  - Closeness is measured in the $\ell_\infty$ sense, i.e., $\|a' - a^*\|_{\ell_\infty} \leq 10^{-3}$.
  - The PPSR is estimated by performing $r$ recovery trials for a range of $s$.
- The **performance** of the proposed method was compared to:
  - The $\ell_1$-Magic suite of algorithms which uses the $\ell_1$ norm as the sparsity promoting function.
  - Our previous method which solves the unconstrained version of the recovery problem with a local quadratic approximation (LQA) of the SCAD.
Reconstruction Performance of the Proposed Method

- Reconstruction **performance** is usually compared in terms of the probability of perfect signal recovery (PPSR).
  - Perfect signal recovery is declared when the solution obtained for the recovery problem $a'$ is close to the true known solution $a^*$.
  - Closeness is measured in the $\ell_\infty$ sense, i.e., $||a' - a^*||_{\ell_\infty} \leq 10^{-3}$.
  - The PPSR is estimated by performing $r$ recovery trials for a range of $s$.
- The **performance** of the proposed method was compared to:
  - The $\ell_1$-Magic suite of algorithms which uses the $\ell_1$ norm as the sparsity promoting function.
  - Our previous method which solves the unconstrained version of the recovery problem with a local quadratic approximation (LQA) of the SCAD.
Reconstruction Performance of the Proposed Method

- Reconstruction performance is usually compared in terms of the probability of perfect signal recovery (PPSR).
  - Perfect signal recovery is declared when the solution obtained for the recovery problem \( \mathbf{a}' \) is close to the true known solution \( \mathbf{a}^* \).
  - Closeness is measured in the \( \ell_\infty \) sense, i.e., \( ||\mathbf{a}' - \mathbf{a}^*||_{\ell_\infty} \leq 10^{-3} \).
  - The PPSR is estimated by performing \( r \) recovery trials for a range of \( s \).
- The performance of the proposed method was compared to:
  - The \( \ell_1 \)-Magic suite of algorithms which uses the \( \ell_1 \) norm as the sparsity promoting function.
  - Our previous method which solves the unconstrained version of the recovery problem with a local quadratic approximation (LQA) of the SCAD.
Numerical Simulations

For a typical PPSR setup such as \( n = 512, q = 100, \) and \( r = 250 \):

A marked improvement in signal recovery is achieved over the two competing methods.
For a typical PPSR setup such as $n = 512$, $q = 100$, and $r = 250$:

A marked improvement in signal recovery is achieved over the two competing methods.
Numerical Simulations, cont.

- For a typical PPSR setup such as $n = 512$, $q = 100$, and $r = 250$:

![Graph showing CPU time vs. sparsity](image)

- The average CPU time is roughly the same as those for the two competing methods for $s \leq 20$, i.e., when the event of a sparse signal being perfectly recovered occurs with probability one.
Numerical Simulations, cont.

- For a typical PPSR setup such as $n = 512$, $q = 100$, and $r = 250$:

  ![Graph showing average CPU time vs. sparsity $s$.]

  - The average CPU time is roughly the same as those for the two competing methods for $s \leq 20$, i.e., when the event of a sparse signal being perfectly recovered occurs with probability one.
Conclusions

- In this presentation we have:
  - Addressed a central problem in CS, which involves the recovery of the original signal from its compressed samples.
  - Proposed a new method for sparse-signal recovery that when compared with two competing methods:
    - Exhibits superior reconstruction performance.
    - Offers approximately the same computational cost when the signal is always perfectly recovered.
Conclusions

- In this presentation we have:
  - Addressed a central problem in CS, which involves the recovery of the original signal from its compressed samples.
  - Proposed a new method for sparse-signal recovery that when compared with two competing methods:
    - Exhibits superior reconstruction performance.
    - Offers approximately the same computational cost when the signal is always perfectly recovered.
In this presentation we have:

- Addressed a **central problem** in CS, which involves the **recovery** of the original signal from its compressed samples.
- Proposed a **new method** for sparse-signal recovery that when compared with two competing methods:
  - Exhibits **superior reconstruction** performance.
  - Offers approximately the **same** computational cost when the signal is always perfectly recovered.
In this presentation we have:

- Addressed a central problem in CS, which involves the recovery of the original signal from its compressed samples.
- Proposed a new method for sparse-signal recovery that when compared with two competing methods:
  - Exhibits superior reconstruction performance.
  - Offers approximately the same computational cost when the signal is always perfectly recovered.
Conclusions

In this presentation we have:

- Addressed a **central problem** in CS, which involves the **recovery** of the original signal from its compressed samples.
- Proposed a **new method** for sparse-signal recovery that when compared with two competing methods:
  - Exhibits **superior reconstruction** performance.
  - Offers approximately the **same** computational cost when the signal is always perfectly recovered.
Thank you for your attention.