On Continuous-Domain Inverse Problems with Sparse Superpositions of Decaying Sinusoids as Solutions

Rahul Parhi
Department of Electrical and Computer Engineering
University of Wisconsin–Madison

IEEE ICASSP 2022

(joint work with Robert D. Nowak)

Continuous-Domain Linear Inverse Problems

- A fundamental problem in science and engineering is to reconstruct a continuous-domain signal $f : \mathbb{R}^d \to \mathbb{R}$ from measurements.

$$y_n = \langle h_n, f \rangle + \varepsilon_n,\ n = 1, \ldots, N$$

- $\mathrm{H}\{f\} = (\langle h_1, f \rangle, \ldots, \langle h_N, f \rangle) \in \mathbb{R}^N$ symbolizes the linear measurement process.
- $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N) \in \mathbb{R}^N$ are perturbation or noise terms, typically zero-mean random variables.
- $\boldsymbol{y} = (y_1, \ldots, y_N) \in \mathbb{R}^N$ denotes the (possibly noisy) data.

Remark

This is an **ill-posed inverse problem**.

Many Real-World Problems are Inverse Problems

$$\boldsymbol{y} = \mathrm{H}\{f\} + \boldsymbol{\varepsilon}$$

- Medical imaging
  - $\Longrightarrow$ Image the interior of a body.
  - $\Longrightarrow$ e.g., in MRI, H corresponds to Fourier-domain measurements.
- Statistics/supervised machine learning
  - $\Longrightarrow$ Learn f from a dataset $\{(\boldsymbol{x}_n, y_n)\}_{n=1}^N$ such that $f(\boldsymbol{x}_n) \approx y_n$.
  - $\Longrightarrow$ $h_n = \delta(\cdot - \boldsymbol{x}_n)$ for some $\boldsymbol{x}_n \in \mathbb{R}^d$.
- Many other real-world problems...

Synthesis Formulation

- Assume *a priori* that f can be **synthesized** by a superposition of atoms from some dictionary, i.e.,

$$f = \sum_{k \in \mathbb{Z}} \alpha_k \varphi_k,$$

where $\{\varphi_k\}_{k \in \mathbb{Z}}$ is a dictionary of atoms, e.g.,
$\Longrightarrow$ wavelets (Donoho, 1998)

- The solution to the inverse problems is synthesized from coefficients that solve

$$\min_{\boldsymbol{\alpha} \in \ell^p(\mathbb{Z})} \left\| \boldsymbol{y} - \mathrm{H}\left\{ \sum_{k \in \mathbb{Z}} \alpha_k \varphi_k \right\} \right\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_p^p,$$

- The choice of $p = 2$ has classically been the common choice, the last few decades have shown that **sparsity** $(p = 1)$ plays a key role in signal reconstruction. (Bruckstein, 2009)
$\Longrightarrow$ Supported by the theory of compressed sensing. (Candès, 2006)

Analysis/Variational Formulation

- The solution to the inverse problem is a solution to the **variational problem**

$$\min_{f \in \mathcal{X}'} \|\boldsymbol{y} - \mathrm{H}\{f\}\|_2^2 + \lambda \|f\|_{\mathcal{X}'}^p,$$

- $\mathcal{X}'$ is an appropriate native space that models the signals to be reconstructed.
- $\|\cdot\|_{\mathcal{X}'}$ is the norm that defines $\mathcal{X}'$.

Sparse Fourier Reconstruction

- In this work we consider the analysis formulation for **sparse** signal reconstruction.
- Consider the **spectral Barron spaces** $\mathscr{B}^s(\mathbb{R}^d)$, $s \geq 0$, which are **Banach spaces** defined by the norm

$$\|f\|_{\mathscr{B}^s(\mathbb{R}^d)} = \|(1 + \|\cdot\|_2)^s F(\cdot)\|_{\mathcal{M}}.$$

$\Longrightarrow$ $F(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} f(\boldsymbol{x}) e^{-\mathrm{j}2\pi\boldsymbol{\xi}^\mathsf{T}\boldsymbol{x}} \, \mathrm{d}\boldsymbol{x}.$

$\Longrightarrow$ The $\mathcal{M}$-norm is a "generalization" of the L^1-norm that can also be applied to distributions such as the Dirac impulse. Morally,

$$\|f\|_{\mathscr{B}^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} (1 + \|\boldsymbol{\xi}\|_2)^s |F(\boldsymbol{\xi})| \, \mathrm{d}\boldsymbol{\xi}.$$

$\Longrightarrow$ The $\mathcal{M}$-norm is the continuous-domain analogue of the **sparsity-promoting** ℓ^1-norm.

$\Longrightarrow$ These spaces were first studied in the context of approximation theory with neural networks. (Barron, 1993)

Sparse Fourier Reconstruction

Representer Theorem (P. & Nowak, 2022)

Consider the analysis formulation for signal reconstruction where the native space is $\mathscr{B}^s(\mathbb{R}^d)$. Furthermore, suppose that the measurement operator H is weak* continuous on $\mathscr{B}^s(\mathbb{R}^d)$. Then,

$$\mathcal{V} = \underset{f \in \mathscr{B}^s(\mathbb{R}^d)}{\arg\min} \ \|\boldsymbol{y} - \mathrm{H}\{f\}\|_2^2 + \lambda \|f\|_{\mathscr{B}^s(\mathbb{R}^d)},$$

is nonempty, convex, and weak* compact.
The extreme points of $\mathcal{V}$ are given by functions of the form

$$f_{\mathrm{sparse}}(\boldsymbol{x}) = \sum_{k=1}^{K} \alpha_k (1 + \|\boldsymbol{\xi}_k\|_2)^{-s} e^{\mathrm{j}2\pi \boldsymbol{\xi}_k^\mathsf{T} \boldsymbol{x}},$$

where $\boldsymbol{\xi}_k \in \mathbb{R}^d$, $k = 1, \ldots, K$, and $K \leq N$. The weak* closure of the convex hull of these extreme points is the full solution set.

- Why is this kind of result called a representer theorem?

  $\Longrightarrow$ The solution set to the optimization problem is completely characterized by functions with a **finite-dimensional representation** in terms of certain atoms.

- The term representer theorem is commonly used when talking about kernel methods in machine learning.

  $\Longrightarrow$ The notion of a representer theorem is much more general and can be applied to many problems about convex optimization in Banach spaces. (Unser 2021)

Sparse Fourier Reconstruction

$$f_{\text{sparse}}(\boldsymbol{x}) = \sum_{k=1}^{K} \alpha_k (1 + \|\boldsymbol{\xi}_k\|_2)^{-s} e^{\mathrm{j}2\pi \boldsymbol{\xi}_k^\mathsf{T} \boldsymbol{x}},$$

- Solution set of is completely characterized by **sparse** superpositions of **decaying sinusoids**.
  - $\Longrightarrow$ Atoms of solutions: $(1 + \|\boldsymbol{\xi}_k\|_2)^{-s} e^{\mathrm{j}2\pi \boldsymbol{\xi}_k^\mathsf{T} \boldsymbol{x}}$.
  - $\Longrightarrow$ As a function of the frequency variable $\boldsymbol{\xi}_k$, the atoms are decaying sinusoids.
  - $\Longrightarrow$ $\|f_{\text{sparse}}\|_{\mathscr{B}^s(\mathbb{R}^d)} = \|\boldsymbol{\alpha}\|_1$.
  - $\Longrightarrow$ Larger s penalizes high-frequencies due to decay factor $(1 + \|\boldsymbol{\xi}\|_2)^{-s}$.
- Solutions have sparse Fourier transforms.
  - $\Longrightarrow$ Fourier transform is superposition of weighted impulses.
- Condition of weak* continuity of the measurement operator is satisfied by convolution operators whose Fourier transforms decay as $\|\boldsymbol{\xi}\|_2^t$, where $t > -s$, as $\|\boldsymbol{\xi}\|_2 \to \infty$,
  - $\Longrightarrow$ Relatively mild condition.

Numerical Examples

$$\|f\|_{\mathscr{B}^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} (1 + \|\boldsymbol{\xi}\|_2)^s |F(\boldsymbol{\xi})| \, \mathrm{d}\boldsymbol{\xi}.$$

- Consider digital images ($d = 2$).

  $\implies$ $f \rightsquigarrow \boldsymbol{f}$, where $\boldsymbol{f}$ is an $N \times N$ digital image.
- Replace continuous Fourier transforms with discrete Fourier transforms.

  $\implies$ $\|f\|_{\mathscr{B}^s(\mathbb{R}^2)} \rightsquigarrow \|\boldsymbol{f}\|_{\ell^s}$

  $\implies$ $$\|\boldsymbol{f}\|_{\ell^s} := \sum_{k_1, k_2} \left(1 + \sqrt{\left|\frac{k_1}{N}\right|^2 + \left|\frac{k_2}{N}\right|^2} \right)^s |\boldsymbol{F}[k_1, k_2]|.$$

  $\implies$ $\boldsymbol{F}[k_1, k_2]$ denotes the DFT of $\boldsymbol{f}$.

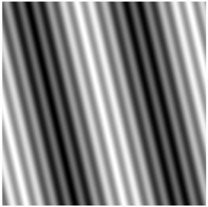
Numerical Examples

We will compare the proposed regularization approach with classical Tikhonov regularization.

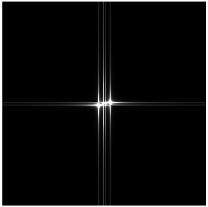
- Proposed approach: $\min_{\boldsymbol{f}} \|\boldsymbol{y} - \mathbf{H}\boldsymbol{f}\|_2^2 + \lambda \|\boldsymbol{f}\|_{\mathscr{b}^s}$
- Tikhonov regularization: $\min_{\boldsymbol{f}} \|\boldsymbol{y} - \mathbf{H}\boldsymbol{f}\|_2^2 + \lambda \|\boldsymbol{f}\|_2^2$
- $\mathbf{H}$ will correspond to convolution with a Gaussian followed by downsampling.
- $\boldsymbol{\varepsilon}$ will be i.i.d. $\mathcal{N}(0,1)$ noise.
- $s = 1$.

Toy Problem

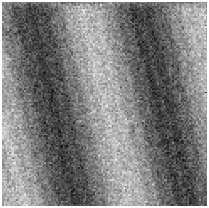
$$\boldsymbol{y} = \mathbf{H}\boldsymbol{f} + \boldsymbol{\varepsilon}$$



512×512 image

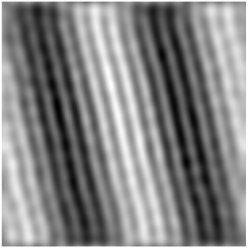


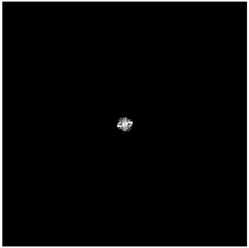
DFT



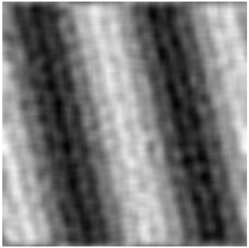
128×128 measurement

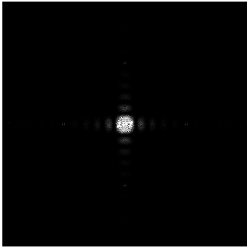
Sparse Fourier Reconstruction



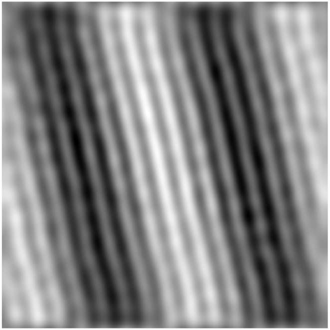


Tikhonov Reconstruction

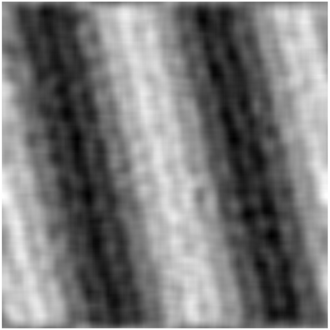




Comparison



Proposed
#coeffs $= 67$



Tikhonov
#Fourier coeffs $= 16384$

Cameraman

512×512 image

128×128 measurement

Cameraman

Proposed
$\#$coeffs $= 180$

Tikhonov
$\#$Fourier coeffs $= 16384$

Concluding Remarks

- Proposed new regularization procedure for sparse signal reconstruction in the **continuous-domain**, adding to a long line of work in this area (Vetterli, 2002), (Eldar, 2005), (Adcock, 2016), (Candés, 2014)

$\Longrightarrow$ Proved a **representer theorem** for the spectral Barron spaces $\mathscr{B}^s(\mathbb{R}^d)$.

- Future work:

$\Longrightarrow$ Comparison of this new approach with existing approaches for sparse signal reconstruction, e.g., wavelet shrinkage and total variation methods.

$\Longrightarrow$ Quantify conditions for **exact recovery** for signals in $f \in \mathscr{B}^s(\mathbb{R}^d)$ from measurements.
When $s = 0$ the problem has been solved under the name "off-the-grid" compressed sensing. (Tang, 2013), (Candés, 2014)