Sparsity Enforcing Edge Detection Method for blurred and noisy Fourier Data

Rosemary Renaut
This is joint work with
Wolfgang Stefan, Rice University,
Aditya Viswanathan, Cal Tech,
and Anne Gelb, Arizona State University

February Fourier Talks 2011

February 17, 2011

Outline

Brief Overview

Objective

Estimate the edges in a piecewise smooth function from blurred and noisy Fourier data.

Assume a finite number of Fourier Coefficients is available for a piecewise function.

Desire accurate and robust detection of jump discontinuities.

Aim to improve reconstructions, restorations and classifications of signals.

The Approach

Approximate the jump function using Concentration Function - l_1 minimization.
Remove Gibbs oscillations and aliasing introduced by concentration using

1. Matching waveform to estimate the jump function.
2. Impose sparseness on the jump function using regularization.

Overview of Test Problems

One dimensional examples considered

No noise, no blur, no under sampling: Best case scenario but still non trivial.

No noise, no blur, but under sampling: Only partial Fourier data available. Fourier coefficients are deleted from the middle of the spectrum (symmetrically), i.e. both low as well as high frequencies still present. In context of minimization problems, missing band of Fourier data corresponds to under sampling.

No noise, Gaussian blur, no under sampling: Fourier coefficients blurred by Gaussian filter

$$\hat{h}_k = e^{-\frac{k^2\tau^2}{2}}. \tag{1}$$

Smooths edges in signal, edge detection using classical methods is difficult.

Additive i.i.d. Gaussian noise, no blur, all samples: How does additive noise in Fourier coefficients impact edge detection?

Non-harmonic Fourier data: Examine edge detection for efficient data collection, eg in MRI.

Background

The function f is represented by a finite number of spectral coefficients

- f is 2π-periodic and piecewise-smooth in $[-\pi, \pi)$.
- It has Fourier series coefficients

$$\hat{f}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx \,, \quad k \in [-N, N]$$

- $\hat{f}$ is a global representation; i.e., $\hat{f}_k$ are obtained using values of f over the entire domain $[-\pi, \pi)$.

Assume f is piecewise smooth

- Its *jump function* is defined by

$$[f](x) := f(x^+) - f(x^-)$$

- A jump discontinuity is a local feature; i.e., the jump function at any point x only depends on the values of f at x^+ and x^-.

Concentration Factor Edge Detection Method (Gelb, Tadmor)

Concentrating the edges using the concentration factor

Approximate $[f](x)$ using generalized conjugate partial Fourier sum (convolution with $C_N^\sigma(x)$)

$$S_N^\sigma[f](x) = i \sum_{k=-N}^{N} \hat{f}_k \, \mathrm{sgn}(k) \, \sigma\left(\frac{|k|}{N}\right) e^{ikx} = (f * C_N^\sigma)(x) \tag{2}$$

$\sigma_{k,N}(\eta) = \sigma(\frac{|k|}{N})$ are known as *concentration factors*.
For convergence of (2) concentration factors have to satisfy admissibility properties:

1. $\displaystyle\sum_{k=1}^{N} \sigma\left(\frac{k}{N}\right) \sin(kx)$ isodd
2. $\displaystyle\frac{\sigma(\eta)}{\eta} \in C^2(0,1)$
3. $\displaystyle\int_\epsilon^1 \frac{\sigma(\eta)}{\eta} \to -\pi$, $\epsilon = \epsilon(N) > 0$ is small

Then the convergence

$$S_N^\sigma[f](x) = [f](x) + \mathcal{O}(\epsilon), \quad \epsilon = \epsilon(N) > 0 \text{ small}$$

depends on σ and the distance between x and a discontinuity of f.

Two relevant examples for Concentration Functions

Low or high order convergence away from a jump

- Polynomial: low order linear or quadratic

$$\sigma_{\text{poly}}(\xi) = p\xi^p, \; p \in \mathbb{N}^+$$

- Exponential: higher order.

$$\sigma_{\exp}(\xi) = \gamma\xi \exp\left(\frac{1}{\alpha\xi(\xi-1)}\right),$$

where

$$\gamma = \frac{\pi}{\int_{\epsilon}^{1-\epsilon} \exp\left(\frac{1}{\alpha\rho(\rho-1)}\right) d\rho}$$

and $\alpha > 0$.

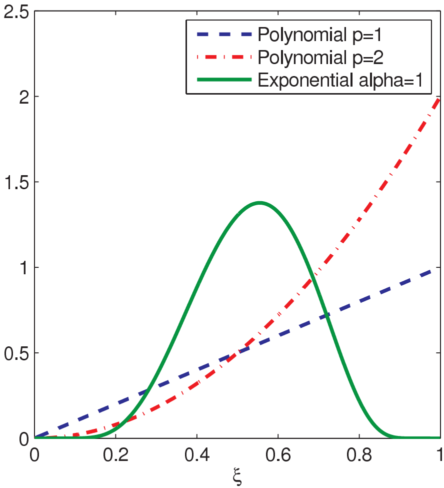
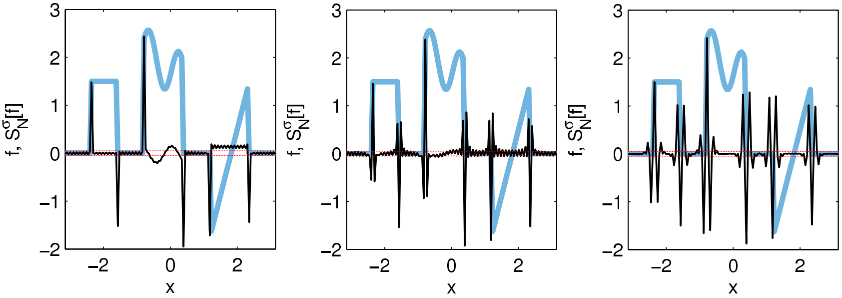


Figure: Illustration of the CFs σ_1 (dash), σ_2 (dash

Illustration of Edge Detection $N = 64$. Black line is the jump function

Example Case

$$f(x) = \begin{cases} 3/2 & \text{for} \quad -\frac{3\pi}{4} \le x < -\frac{\pi}{2} \\ 7/4 - x/2 + \sin(7x - 1/4) & \text{for} \quad -\frac{\pi}{4} \le x < \frac{\pi}{8} \\ x \cdot 11/4 - 5 & \text{for} \quad \frac{3\pi}{8} \le x < \frac{3\pi}{4} \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$



(a) Polynomial p=1 (σ_1) (b) Polynomial p=2 (σ_2) (c) Exponential ($\sigma_{\exp}$)

Observations

- Polynomial CFs only filter low frequencies
- Exponential also filters some high frequencies
- Fast convergence away from a jump leads to more oscillations around the jump
- Many false positive and false negatives with regard to classifying jumps.

The minmod to improve the approximation (Gelb and Tadmor (2006))

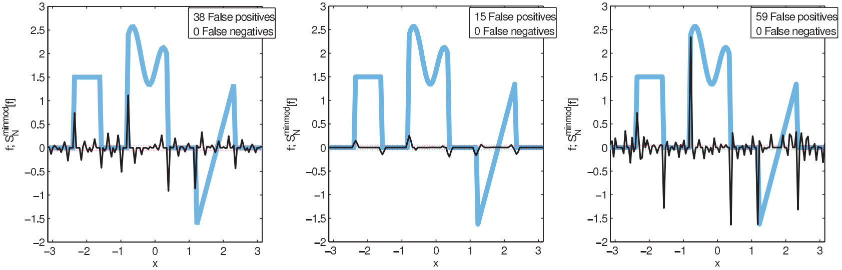
Use the **minmod** function over different concentration functions

$$\textbf{minmod}\{a_1, a_2, \dots, a_n\} := \begin{cases} s \cdot \min(|a_1|, |a_2|, \dots, |a_n|) & \text{if } \operatorname{sgn}(a_1) = \cdots = \operatorname{sgn}(a_n) := \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

yielding the approximation

$$S_N^{MM}[f](x) = \textbf{minmod}\{S_N^{\sigma_1}[f](x), S_N^{\sigma_2}[f](x), \dots, S_N^{\sigma_n}[f](x)\}. \tag{5}$$

Minmod CF approximation to the jump function for noisy and blurred functions



(d) Under sampling (e) Blurring by a Gaussian (f) Noisy Fourier coefficients

Figure: Noting false positives and false negatives for identifying edges using a 5% threshold. (a) 10% missing Fourier Coefficients. (b) Gaussian blur of variance $\tau = 0.05$, for point spread function coefficients $\hat{h}_k = e^{-\frac{k^2\tau^2}{2}}$. (c) Noise of variance .015 applied to Fourier Coefficients.

For blurred functions the edges may be missed, for noisy functions or with missing data too many edges are determined.

Improving Concentration using the Matching Waveform (A. Gelb and D. Cates, 2008)

- Jump function approximation at $x = \xi$ depends on size and location of jump but not on f:

$$S_N^\sigma[f](x) = \frac{[f](\xi)}{\pi} \sum_{k=1}^{N} \sigma\left(\frac{k}{N}\right) \frac{\cos k(x-\xi)}{k} + \mathcal{O}\left(\frac{\log N}{N}\right).$$

- Introduce the waveform

$$W_N^\sigma(x) = \sum_{k=1}^{N} \sigma\left(\frac{k}{N}\right) \frac{\cos kx}{k}. \tag{6}$$

- Apply the CF and correlate the obtained waveform with CF applied to an indicator function

$$S_N^{\sigma_{mw}}[f](x) = \frac{1}{\gamma_{mw}} (S_N^\sigma[f] * W_N^\sigma)(x), \quad \text{normalization} \quad \gamma_{mw} = \frac{1}{\pi} \sum_{k=1}^{N} \left(\frac{\sigma(\frac{k}{N})}{k}\right)^2 \tag{7}$$

- This leads to the admissible *matching waveform concentration factor* (MWCF)

$$\sigma_{mw}\left(\frac{|k|}{N}\right) := \frac{1}{\gamma_{mw}} \sigma\left(\frac{|k|}{N}\right) \int_{-\pi}^{\pi} W_N^\sigma(\rho) \exp\left(-ik\rho\right) d\rho. \tag{8}$$

MWCF performs better in the presence of noise, but does not remove oscillations.
Performance deteriorates for nearby jumps.

Appealing to Sparsity (Tadmor and Zou 2008)

- A minimization formulation (iterative) provides an alternative to the matched filter edge detector.
- We take inspiration from sparsity enforcing regularization routines and their iterative solutions (Tadmor and Zou).
- Using a *fit to data* functional with a constraint condition on the sparsity of total variation in the signal is one approach.
- We consider l_1 sparsity as an alternative and combine with the matching waveform technique for correlating edges.

Find Approximate jump function for f using coefficients of Noisy Blurred Function f

Appealing to sparsity

- Given are $\hat{g}_k$ for blur function h and noise n, $\hat{g}_k = \hat{h}_k \cdot \hat{f}_k + \hat{n}_k$
- Approximating jump function for f is jump function for g with Fourier coefficients

$$(\hat{S_N^\sigma[g]})_k = \left(i \cdot \sigma\left(\frac{|k|}{N}\right) \cdot \operatorname{sgn}(k)\right) \cdot \hat{g}_k$$

- Using $\hat{g}_k \approx \hat{h}_k \cdot \hat{f}_k$ yields

$$\left(i \cdot \sigma\left(\frac{|k|}{N}\right) \cdot \operatorname{sgn}(k)\right) \cdot \hat{g}_k = (\hat{S_N^\sigma[g]})_k \approx \hat{h}_k (\hat{S_N^\sigma[f]})_k$$

- We seek a sparse y which also approximates the jump function of f
- Convolving y with $W_N^\sigma(x)$ should also approximate jump function $S_N^\sigma[f](x)$

$$(\hat{S_N^\sigma[f]})_k \approx (\hat{W_N^\sigma * y})_k = (\hat{W_N^\sigma})_k \cdot \hat{y}_k \tag{9}$$

- We obtain

$$\hat{h}_k \cdot (\hat{W_N^\sigma})_k \cdot \hat{y}_k \approx i \cdot \sigma\left(\frac{|k|}{N}\right) \cdot \operatorname{sgn}(k) \cdot \hat{g}_k$$

A Discrete Variational Formulation

l_1 minimization

- Introduce matrices describing the components of the approximate equation

$$\begin{aligned} \Sigma &= \operatorname{diag}\left(\sigma\left(\frac{|-N|}{N}\right),\cdots,0,\cdots,\sigma\left(\frac{|N-1|}{N}\right)\right) \\ H &= \operatorname{diag}(\frac{\pi}{|-N|}\hat{h}_{-N},\cdots,0,\cdots,\frac{\pi}{|N-1|}\hat{h}_{N-1}) \quad \text{and} \\ F_{kj} &= \frac{1}{2N}(-1)^k\exp(\frac{-i\pi jk}{N}) \quad \text{where} \quad \hat{\mathbf{y}} = F\mathbf{y}((x)). \end{aligned}$$

- Then to find the discrete approximation to $y(x)$, given by vector $\mathbf{y}$, we can solve

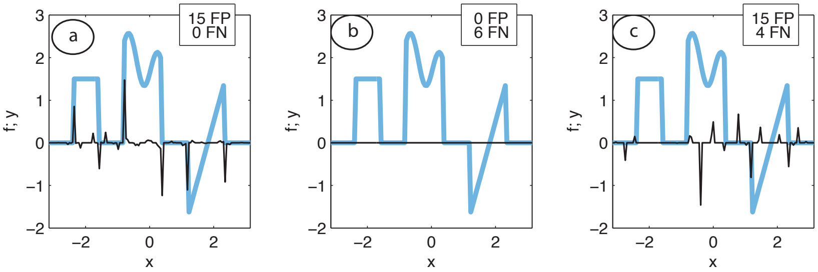
$$\mathbf{y} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 \qquad \text{subject to} \qquad \|\Sigma(HF\mathbf{u}-\mathbf{b})\|_2^2 \le \delta, \tag{10}$$

$\mathbf{b} = (-i\cdot\hat{g}_{-N},\cdots,0,\cdots,i\cdot\hat{g}_{N-1})$. Concentration weights the data fit term.

- This is a second order cone problem.
- Introduce λ and solve

$$\mathbf{y} = \arg\min_{\mathbf{u}}\{\lambda\|\mathbf{u}\|_1 + \frac{1}{2}\|\Sigma(HF\mathbf{u}-\mathbf{b})\|_2^2\}, \tag{11}$$

Experiments with $N = 64$ and under sampling but no noise and no blur.

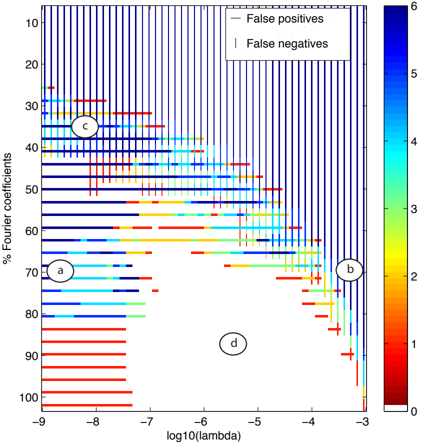


(a) $(\mathrm{FC}, \lambda) = 70\%, 10^{-9}$ (b) $(\mathrm{FC}, \lambda) = 70\%, 10^{-3}$ (c) $(\mathrm{FC}, \lambda) = 35\%, 10^{-8}$

Figure: Edge detection using the exponential concentration factor. FC is the percentage of Fourier Coefficients used, y is the thin line, unseen in (b). FP and FN are the count of misidentified edges, either false positive or false negative, using 5% threshold on $\mathbf{y}$.

One sees the effect of the regularization parameter comparing (a) and (b), and the effect of reducing the number of Fourier Coefficients comparing (a) and (c).

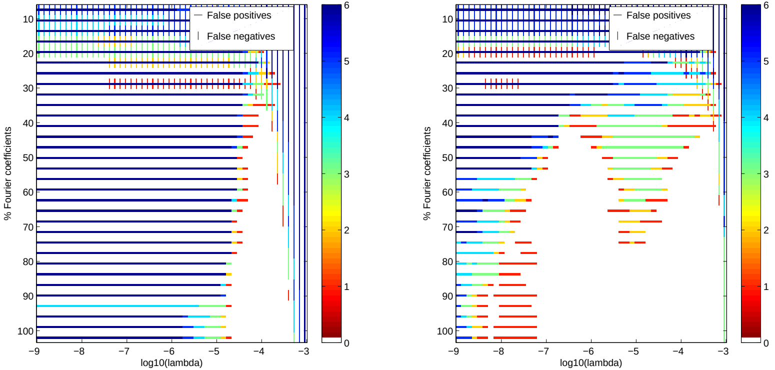
Under sampling, no noise, no blur. False Positives and False Negatives with Waveform



Region (d) shows that there is a range of regularization parameters for which the method is robust with respect to correct identification of edges provided up to about 70% of coefficients are retained.

Figure: Illustrating the impact of the regularization parameter choice in relation to the number of Fourier Coefficients that are sampled, and the impact on the number of False Positives and False Negatives

Using the Polynomial Concentration Factors. $N = 64$

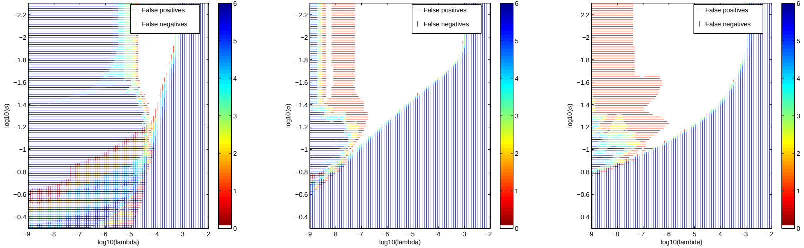


(a) N=64, σ_1 (b) N=64, σ_2

Figure: No blur and no noise. Edge detection using the polynomial concentration factor with varying amounts of Fourier data missing.

Higher order concentration factors perform better at capturing the edges correctly for a wider range of regularization parameters.

Edge detection in the presence of blur in the coefficients. $N = 64$.

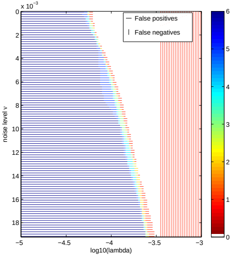


(a) σ_1 (b) σ_2 (c) σ_{exp}

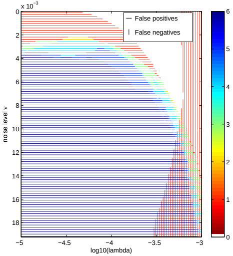
Figure: Edge detection in blurred signals using σ_p, for $p = 1, 2$, and σ_{exp}. All plots show that the method can handle blurring where the traditional CF method fails.

Gaussian blur of variance $\tau = 0.05$, for point spread function coefficients $\hat{h}_k = e^{-\frac{k^2\tau^2}{2}}$. The higher order concentration factors again perform better at capturing the edges correctly for a wider range of regularization parameters.

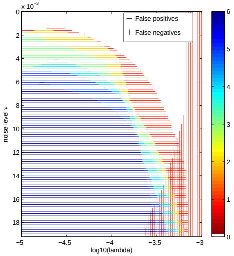
Edge detection in the presence of additive noise in the coefficients. $N = 64$.



(a) σ_1



(b) σ_2

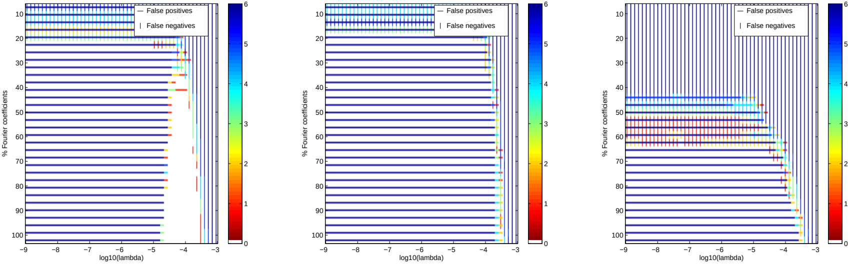


(c) σ_{exp}

Figure: Edge detection in signals with noise of variance .015 applied to Fourier Coefficients. All plots show that the method can handle noise where the traditional CF method fails.

In this case the higher order exponential concentration factor performs better than the quadratic, perhaps due to its inherent filtering of coefficients contaminated with noise.

Is the waveform correlation required? Examples without the waveform for $N = 64$



(a) σ_1 (b) σ_3 (c) σ_{exp}

Figure: The figures illustrate for no blur and no noise success of edge detection for correctly finding edges, as Fourier data are removed, and robustness to choice of the regularization parameter λ.

When using the low order polynomial concentration factor the method is quite robust, but for higher order concentration factors the method is more sensitive to the choice of λ and the ability to correctly detect edges is limited.

Non-harmonic Fourier data

Motivation

- Modern MRI scanners optimize data collection strategies by collection of Fourier data on non-cartesian representations of the k-space. The non-harmonic Fourier data, $\hat{f}(\omega_k)$, for piecewise-analytic $f \in L^2(\mathbb{R}(-\pi,\pi))$ are defined by

$$\hat{f}(\omega_k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-i\omega_k x}dx, \ \omega_k \notin \mathbb{Z}. \tag{12}$$

- Consider an immediate extension of the convolution form of the generalized conjugate partial Fourier sum (2)

$$\tilde{S}_N^{\sigma}[f](x) = (f * \tilde{C}_N^{\sigma})(x) \tag{13}$$

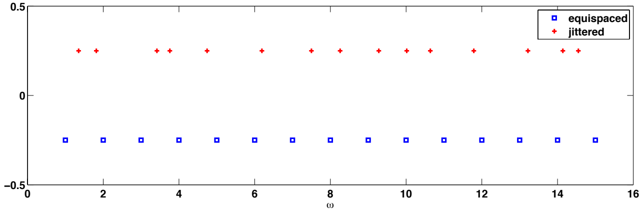
$$:= i\sum_{k=-N}^{N} \alpha_k \hat{f}(\omega_k)\mathrm{sgn}(\omega_k)\sigma\left(\frac{|\omega_k|}{N}\right)e^{i\omega_k x}. \tag{14}$$

The coefficients α_k are weights for the non-uniform trapezoidal rule approximation of the inverse Fourier integral. (*convolutional gridding*).

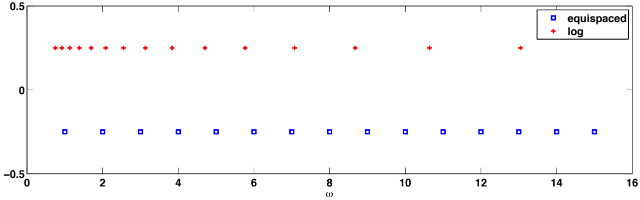
- Example sampling

$$\textit{jittering} \quad \omega_k = k \pm \zeta_k, \quad \zeta_k \sim U[0,\theta], \ k = -N, -(N-1), ..., N. \tag{15}$$

Example Distributions



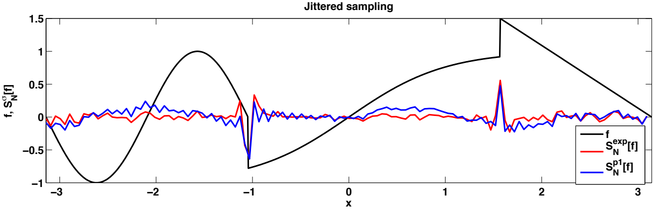
(a) Jittered sampling



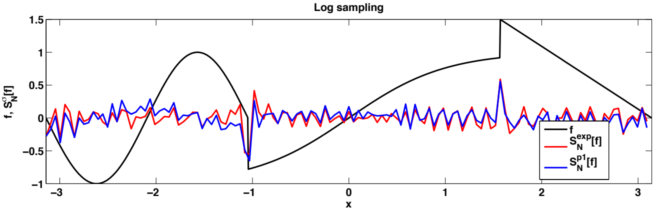
(b) Log sampling

Figure: Non-harmonic sampling distributions (right half plane), $N = 16$

Applying the Edge Detector with the non-harmonic concentration sum



(a) Edges from jittered sampling using σ_1 and σ_{exp} with $\alpha = 2$.



(b) Edges from log sampling using σ_1 and σ_{exp} with $\alpha = 2$.

Figure: Jump approximations: non-harmonic Fourier data using non-harmonic concentration sum, $N = 64$

Extending the Sparsity Approach

Find a Sparse Approximation of the Jump Function

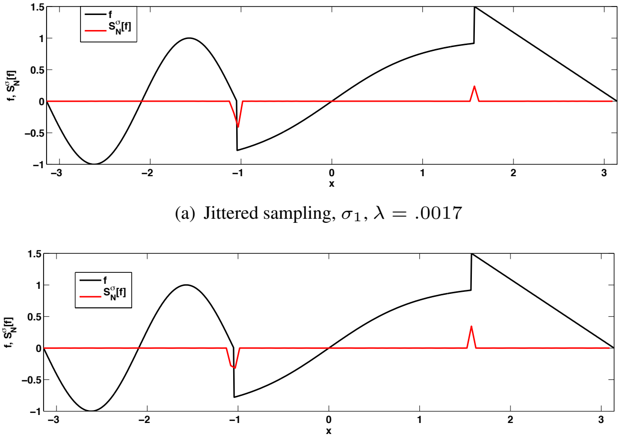
- Assume $\mathbf{g} = (\hat{g}(\omega_{-N}), ..., \hat{g}(\omega_{N-1}))^T$ is the vector of non-harmonic measurements.
- Assume $\mathbf{y}$, approximates $[f]$ on equispaced grid $x_j = \frac{\pi j}{N} - \pi$, $j = 0, \ldots, 2N-1$.
- Introduce the necessary matrices on the non-harmonic modes, Σ the diagonal matrix of concentration factors, H the diagonal matrix of blur coefficients, $F \in \mathbb{C}^{2N \times 2N}$ the discrete non-harmonic Fourier matrix, and W a Toeplitz matrix whose rows contain shifted replicates of the jump waveform $W_N^\sigma(x)$

$$\begin{aligned}
\Sigma &= i \cdot \text{diag}\left(\text{sgn}(\omega_{-N})\,\sigma\left(\frac{|\omega_{-N}|}{N}\right), ..., 0, ..., \text{sgn}(|\omega_{N-1}|)\,\sigma\left(\frac{|\omega_{N-1}|}{N}\right)\right)\\
H &= \text{diag}\left(\hat{h}(\omega_{-N}), ..., \hat{h}(\omega_{N-1})\right),\\
F_{kj} &= \exp\left[i\left(-\pi + \frac{\pi j}{N}\right)\omega_k\right], \quad k = -N, ..., N-1,\ j = 0, ..., 2N-1.
\end{aligned}$$

- Compute the jump approximation by solving

$$\mathbf{y} = \arg\min_{\mathbf{u}}\{\lambda\|\mathbf{u}\|_1 + \frac{1}{2}\|HFW\mathbf{u} - \Sigma\mathbf{g}\|_2^2\}. \tag{16}$$

Example for exact data: Detects the location but not the height

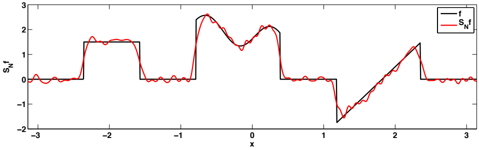


(a) Jittered sampling, σ_1, $\lambda = .0017$

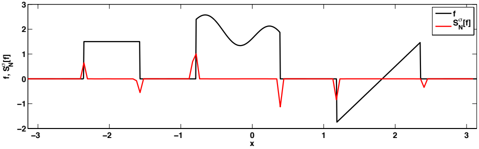
(b) Log sampling, $\sigma_{\exp}$, $\alpha = 2$, $\lambda = .00091$

Figure: Jump approximations from non-harmonic Fourier data using the variational formulation, $N = 64$

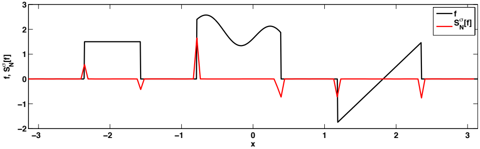
Example for blurred and noisy non-harmonic Fourier data, $N = 64$



(a) Fourier reconstruction of blurred noisy data



(b) Jittered spectral data using σ_1



(c) Log spectral data using $\sigma_{\exp}$, $\alpha = 2$

- Gaussian blur variance $\tau = .05$.
- Additive white complex Gaussian noise, variance .015.
- Regularization parameters .002 in (b) and .0013 in (c).
- Solution is more sensitive to choice of regularization parameter λ than for the harmonic case.
- Determination of λ is harder for the log than jittered data.

Conclusions

- Use of the variational formulation which employs sparsity in the jump function approximation yields a robust approach for both noisy and blurred signals.
- The approach requires the matching waveform to improve robustness with respect to choice of the regularization parameter.
- Method is successful in the presence of missing Fourier data. (here sampled from the middle of the spectrum).
- The approach is a regularized deconvolution of the approximate jump function.
- Higher order exponential concentration function outperforms low order polynomial concentration functions.
- Method can be extended for non-harmonic data, edges are detected but the heights are not correct.
- Algorithm has been extended for two dimensional examples, Stefan and Yin (2010).
- Can be useful for accurate classification of edges in signals.

References

1. A. GELB AND E. TADMOR, *Detection of Edges in Spectral Data*, in Appl. Comp. Harmonic Anal., 7 (1999), pp. 101–135.
2. A. GELB AND E. TADMOR, *Detection of Edges in Spectral Data II Nonlinear Enhancement*, in SIAM J. Numer. Anal., Vol. 38, 4 (2000), pp. 1389–1408.
3. A. GELB AND E. TADMOR, *Adaptive edge detectors for piecewise smooth data based on the minmod limiter*, in J. Sci. Comput., 28(2-3): (2006), pp. 279–306.
4. A. GELB AND D. CATES, *Detection of Edges in Spectral Data III -refinement of the concentration method*, in J. Sci. Comput., 36, 1 (2008), pp. 1-43.
5. E. TADMOR AND J. ZOU, *Novel edge detection methods for incomplete and noisy spectral data*, in J. Four. Analy. App. 14(5) (2008), pp 744-763.
6. W. STEFAN, A. VISWANATHAN, A. GELB, AND R. A. RENAUT, *Sparsity enforcing edge detection method for blurred and noisy Fourier data*, (2010).

A Two Dimensional Example (Stefan and Yin)

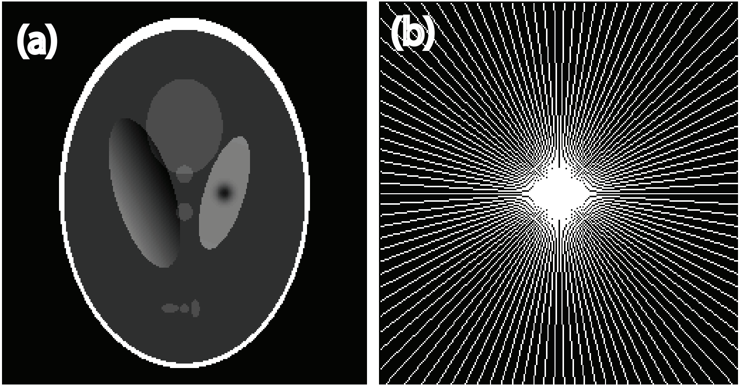


Figure: A modified Shepp logan phantom with gradients and a radial sampling pattern.

Some Two Dimensional Results (Stefan and Yin)

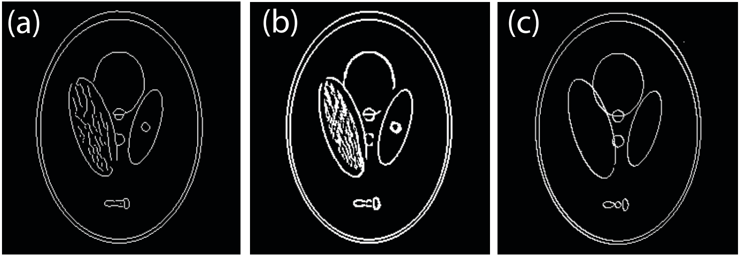
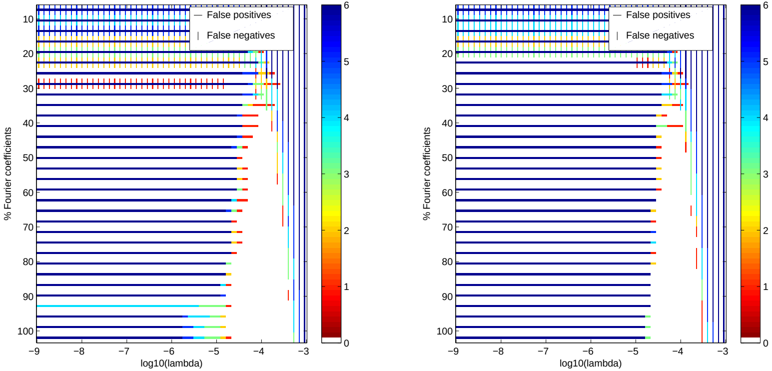


Figure: Edge detection using (a) Canny edge detector (matlab) after reconstruction from the radial samples using TV. (b) Wavelet edge detector on TV reconstruction. (c) A 5th order FD edge detector

Comparing the performance of the waveform correlation $N = 64$, for σ_1

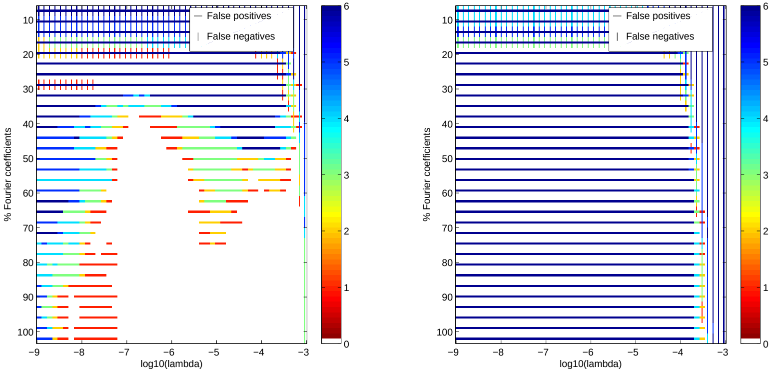


(a) With W (b) Without W

Figure: No blur, no noise. Edge detection using the polynomial concentration factor with and without the waveform weighting, varying amounts of Fourier data missing.

The two approaches are comparable when using the low order polynomial concentration factor.

Comparing the performance of the waveform correlation $N = 64$, for σ_3



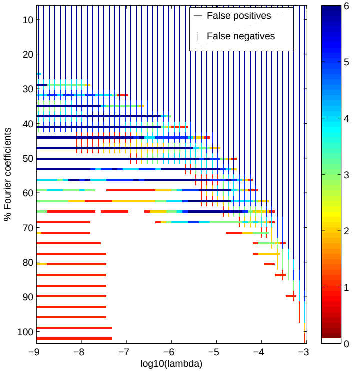
(a) With W

(b) Without W

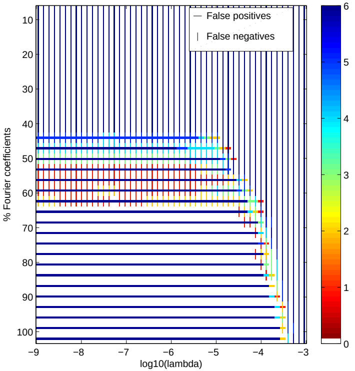
Figure: No blur, no noise.. Edge detection using the polynomial concentration factor with and without the waveform weighting, varying amounts of Fourier data missing.

The waveform is required when using a higher order polynomial concentration factor, which introduces more oscillations that need to be suppressed by the waveform.

Comparing the performance of the waveform correlation $N = 64$, for σ_{exp}



(a) With W



(b) Without W

Figure: No blur, no noise. Edge detection using the exponential concentration factor with and without the waveform weighting, varying amounts of Fourier data missing.

Again, the waveform is required when using a higher order concentration factor, which introduces more oscillations that need to be suppressed by the waveform.