Bandwidth extension for wave-based imaging

Laurent Demanet

Department of Mathematics, MIT

Nam Nguyen, IBM
Elita Li, NUS

May 2017 – IPAM – UCLA
Bandwidth extension...

- Which signal model?
- How far can we hope to extrapolate?
- What level of confidence?
\[ \hat{u}(\omega) = \hat{g}(\omega) \hat{v}(\omega) \]

\[ \hat{g}(\omega) \text{ is} \]
- PSF (image processing)
- pulse (radar)
- wavelet (seismic)

Blur at scale \( 2\pi/\omega_1 \)

Rayleigh diffraction limit

For illustration, reflection model

\[ \hat{v}(\omega) = \int G_c(x_S, x; \omega) \rho(x) G_c(x, x_R; \omega) dx + h.o.t., \]

background velocity \( c \), reflectivity \( \rho \)
High frequencies

Beyond $\omega_1$: super-resolution

Find small-scale details that seem absent from data

Credit: Papson & Narayanan

Credit: CO2CRC

Super-resolved microscopy: W. E. Moerner.
Low frequencies

Below $\omega_0$: inverse scattering

The dependence on $c$ is “least nonlinear” at low frequencies:

$$G_c(x, y; \omega) \sim e^{i\omega \tau(x,y)} \Rightarrow \frac{\delta^2 \hat{V}}{\delta c^2} \sim \omega^2$$

**FWI**: solve for $(c, \rho)$ starting with low frequencies only, slowly grow the window.

**EFWI**: seed FWI on synthesized low frequencies

Extrapolation to low $\omega$ $\Leftrightarrow$ Inversion of $c, \rho$
\[ \hat{\nu}(\omega) = \int G_c(x_S, x; \omega) \rho(x) G_c(x, x_R; \omega) \, dx + h.o.t. \]

but \( c, \rho, \hat{g}(\omega) \) are unknown.

1. Sparse superposition of complex exponentials (signal processing)

2. Wave-inspired model reduction (seismology)
1. Sparse model

In place of

\[ \hat{v}(\omega) = \int G_c(x_s, x; \omega) \rho(x) G_c(x, x_R; \omega) dx + h.o.t., \]

use instead

\[ \hat{v}(\omega) = \sum_{j=1}^{k} a_j e^{i\omega \theta_j} + e(\omega) \]

... because wave recordings can often be decomposed into coherent events:

- \[ G_c(x, y, \omega) \sim e^{i\omega \tau(x, y)} \]
- the singularities of \( \rho(x) \) matter most
In time: a spike train

\[ v(t) = \sum_{j=1}^{k} a_j \delta(t - \theta_j) + \text{error} \]

Super-resolution question

Minimum recoverable distance \( \tau = \min_{i \neq j} |\theta_i - \theta_j| \) vs.
- band limit \( \Omega = \omega_1 - \omega_0 \),
- sparsity \( k \),
- noise level \( \sigma = \|e\|_2 \).

Extrapolation: determine \( a_j, \theta_j \), then evaluate \( \hat{v}(\omega) \) for new \( \omega \)
Safe, pessimistic answer: Nyquist/Rayleigh sampling $\tau \geq \frac{2\pi}{\Omega}$.

Radioastronomy, ultrasound, NMR spectroscopy: super-resolution seems possible when $\tau < \frac{2\pi}{\Omega}$. Sparsity helps.

Classes of algorithms:

1. Subspace / MUSIC (MUltiple SIgnal Classification)
2. Matrix pencil / HSVD / FRI etc.
3. $\ell_1$ minimization to recover a sparse $x$ from $y$
4. Matching pursuits: remove spikes 1-by-1
\[ \hat{\nu}(\omega) = \sum_{j=1}^{k} a_j e^{i\omega \theta_j} + e(\omega), \quad |\omega| \leq \frac{\Omega}{2} \]

Parameters: \( \theta_j \in \tau \mathbb{Z} \), and \( \text{SRF} = \frac{2\pi}{\tau \Omega} \). (Nyquist: \( \text{SRF} = 1 \).)

**Theorem (Donoho)**

Let \( S_k = \{ k\text{-sparse } x \} \), and consider the minimax error

\[ E_{\text{minmax}} = \inf_{\hat{\nu}} \sup_{\nu \in S_k} \sup_{\|e\|_2 \leq \sigma} \|\hat{\nu} - \nu_0\|_2. \]

Then

\[ C_1,k (\text{SRF})^{2k-1} \sigma \leq E_{\text{minmax}} \leq C_2,k (\text{SRF})^{2k+1} \sigma \]

Donoho's original model: sparse clumps of cardinality \( k \).
An answer to Donoho’s question:

**Theorem (D., Nguyen, 2015)**

The recoverability scaling is

\[ E_{\text{minmax}} \asymp C_k (SRF)^{2k-1} \sigma \]

Additionally, *MUSIC reaches this bound*.

Consequences for extrapolation:

- Go from band of size \( \Omega \) to band of size \( (SRF) \times \Omega \),
- Conditioning penalty factor \( (SRF)^{2k-1} \).
2. Wave-inspired model reduction

In place of

\[ \hat{v}_{R,S}(\omega) = \int G_c(x_S, x; \omega) \rho(x) G_c(x, x_R; \omega) \, dx + h.o.t., \]

use a sparse superposition of **atomic phases**

\[ \hat{v}_{R,S}(\omega) = \sum_{j=1}^{k} a_j(x_R, x_S; \omega) e^{i(\omega \tau_j(x_R, x_S) + \phi_j(x_R, x_S; \omega))} \]

where all functions are as smooth as possible (not oscillatory)

- \( a_j(x_R, x_S; \omega) \): geometrical spreading, wavelet inaccuracies
- \( \tau_j(x_R, x_S) \): arrival times
- \( \phi_j(x_R, x_S; \omega) \): dispersive corrections
Model reduction

In principle: minimize LS data misfit, while keeping $\phi_j$, $\nabla \phi_j$, $\nabla a_j$, $\nabla \tau_j$ small.

Strong nonconvexity!

Best handled by the tracking heuristic:

- Initialize with a good linear-phase finder (like MUSIC)
- Iterate to fit the nonlinearities in $a_j$ and $\phi_j$.
- Carefully grow a trust region in $\omega$ and $x_R, x_S$.

Extrapolation: polynomial fits for $a_j, \phi_j$, and evaluate formula for new values of $\omega$. 
Band: [16, 70] Hz.
[10, 85] Hz (true)  
[0.5, 120] Hz (extrap).
[10, 16] Hz (true)  [10, 16] Hz (extrap)  [0.5, 10] Hz (extrap)
Example 2: Marmousi at [5, 15] Hz

Original vs. initial
Example 2: Marmousi at $[5, 15]$ Hz

Synthesized data, $[5, 15]$ Hz, vs. phase tracking fit (9 events)
Example 2: Marmousi at [5, 15] Hz

Unavailable data, [1, 5] Hz, vs. extrapolated data
Example 2: Marmousi at [5, 15] Hz

(Cheating) (New method) (Old method)
Conclusions

Recent results help understand what is possible with extrapolation:

- **$k$-sparsity** enables provable bandwidth extension

- **model reduction with phases** is robust enough to handle seismic shot records

Outlook / other topics:

- **smoothness** enables extrapolation up to a fraction of the characteristic smoothness length

- **model extension** can be used in place of model reduction for frequency extrapolation
Bonus slides
1. Analytic model

\[ \| \frac{d^n}{d\omega^n} \hat{\nu}(\omega) \|_{L^\infty(-1,1)} \leq Q R^{-n} n! \]

... because the scattering poles of \( G_c(x, y; \cdot) \) are hopefully far from the real axis:

- radiation condition
- no resonant cavity
- attenuation
Main result

Noise level $\sigma = \|e\|_\infty$, $N$ equispaced points, Bernstein radius $\rho$. Least-squares fit by a polynomial $p_M(\omega)$ of degree

$$M = \min\left\{ \frac{\sqrt{N}}{2}, \log_{\rho}(1/\sigma) \right\}$$

**Theorem (D., Townsend, 2016)**

Let $r(\omega) = \omega + \sqrt{\omega^2 - 1}/2$, so that $r \to 1^-$ when reaching the ellipse. Then

$$|\hat{\nu}(\omega) - p_M(\omega)| \leq C_{\rho,\sigma} \frac{1}{1 - r(\omega)} \sigma^{\alpha(\omega)},$$

where

$$\alpha(\omega) = -\log_{\rho}(r(\omega)).$$

This rate is minimax optimal, and $N$ does not enter the rate.
In short, the news are bad.
1. LSERTM: solve in the ULS sense

$$\min_{\bar{\rho}} \int \int |\bar{\rho}(x, h)|^2 dx dh \quad \text{s.t.} \quad \hat{\nu} = F_e(\bar{\rho})$$

2. Better: TV regularization

$$\min_{\bar{\rho}} \lambda \int \int |\nabla_{x, h}\bar{\rho}(x, h)| dx dh + \| \hat{\nu} - F_e(\bar{\rho}) \|_2^2$$

Extrapolation: run the $F_e$ model with a different wavelet $\hat{w}(\omega)$. Results similar to model reduction.
Model extension
Model extension
Model extension
Model extension
FWI [2, 50] Hz, vs. EFWI [6, 50] Hz, vs. FWI [6, 50] Hz
(Cheating) (New method) (Old method)