Anti-leakage least-squares spectral analysis for data regularization

Ebrahim Ghaderpour
Postdoctoral Scholar
Department of Mathematics and Statistics
University of Calgary

Joint Work With
Dr. Wenyuan Liao, Dr. Michael Lamoureux, Da Li and Dr. Spiros Pagiatakis

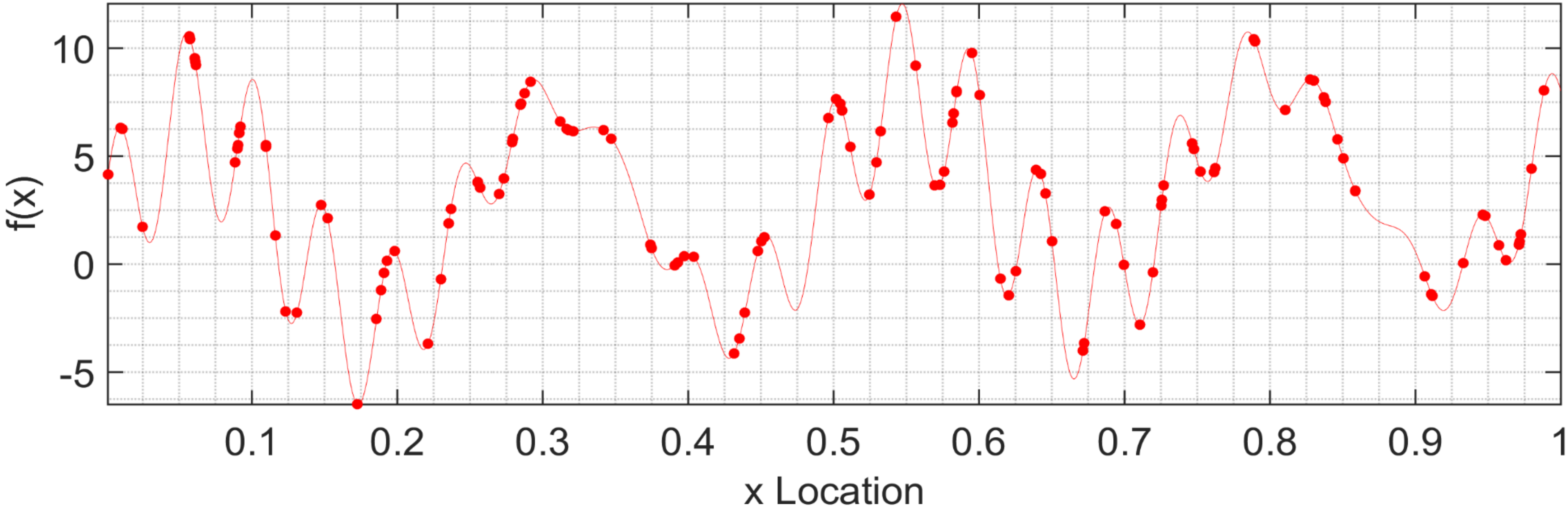
Outline

- Irregularly sampled (unequally spaced) data series
- Least-squares spectral analysis (LSSA)
- Anti-leakage LSSA
- Application in seismic trace interpolation
- Conclusion

An irregularly spaced data series (red dots)

Consider the following function, where the x_ℓ's are 128 random numbers in [0,1].

$$f(x_\ell) = 5\sin(25.6\,x_\ell) + 2.5\ \sin(128\,x_\ell + 1) + \sqrt{3}\ \sin(140\,x_\ell) + \sqrt{2} + \pi x_\ell$$



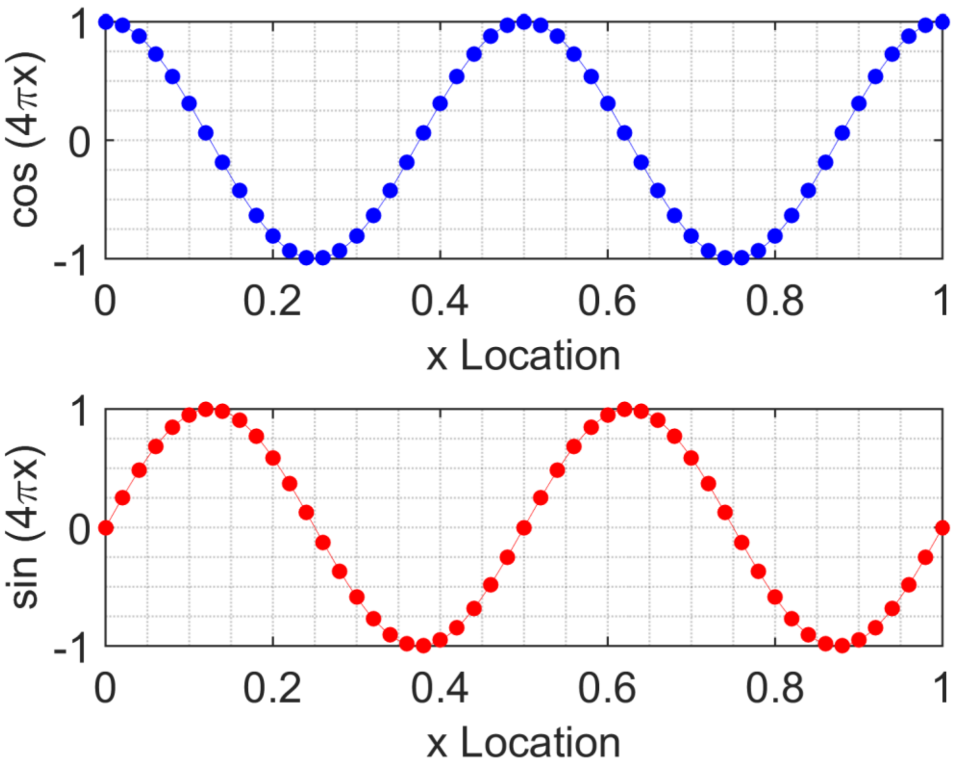
Least-squares spectral analysis (LSSA)

- The LSSA estimates a frequency spectrum based on the least-squares fit of sinusoids to data series.

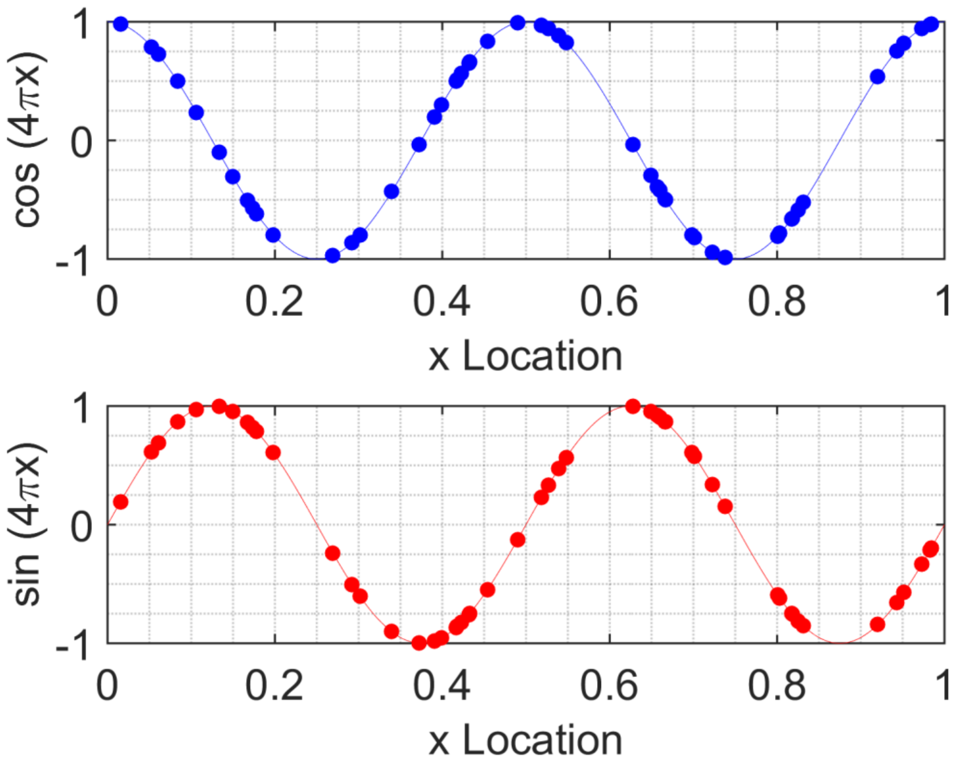
- Unlike the Fourier analysis, the LSSA considers the correlations among the sine and cosine basis functions for each frequency.

- In the LSSA, for each frequency, we fit the sinusoids of that frequency to the data series.

Orthogonality (correlation) of sinusoidal basis functions



$$\sum_{k=1}^{51} \cos(4\pi x_k)\sin(4\pi x_k) \cong 0$$



$$\sum_{k=1}^{51} \cos(4\pi x_k)\sin(4\pi x_k) \cong -3.5$$

Least-squares spectral analysis

Let $\boldsymbol{f} = [f(x_\ell)]$ be a column vector of n samples, where the x_ℓ's may be irregularly spaced. For each frequency ω_k, we minimize the cost function $\psi(\boldsymbol{c}_k) = (\boldsymbol{f} - \boldsymbol{\phi}_k \boldsymbol{c}_k)^T (\boldsymbol{f} - \boldsymbol{\phi}_k \boldsymbol{c}_k)$, where T is transpose and $\boldsymbol{\phi}_k$ is the following $n \times 2$ **design** matrix:

$$\boldsymbol{\phi}_k = [\cos(2\pi\omega_k x_\ell) \quad \sin(2\pi\omega_k x_\ell)].$$

The least-squares spectrum (frequency-amplitude) is defined as:

$$\widehat{\boldsymbol{c}_k} = \left(\boldsymbol{\phi}_k^T \boldsymbol{\phi}_k\right)^{-1} \boldsymbol{\phi}_k^T \boldsymbol{f}$$

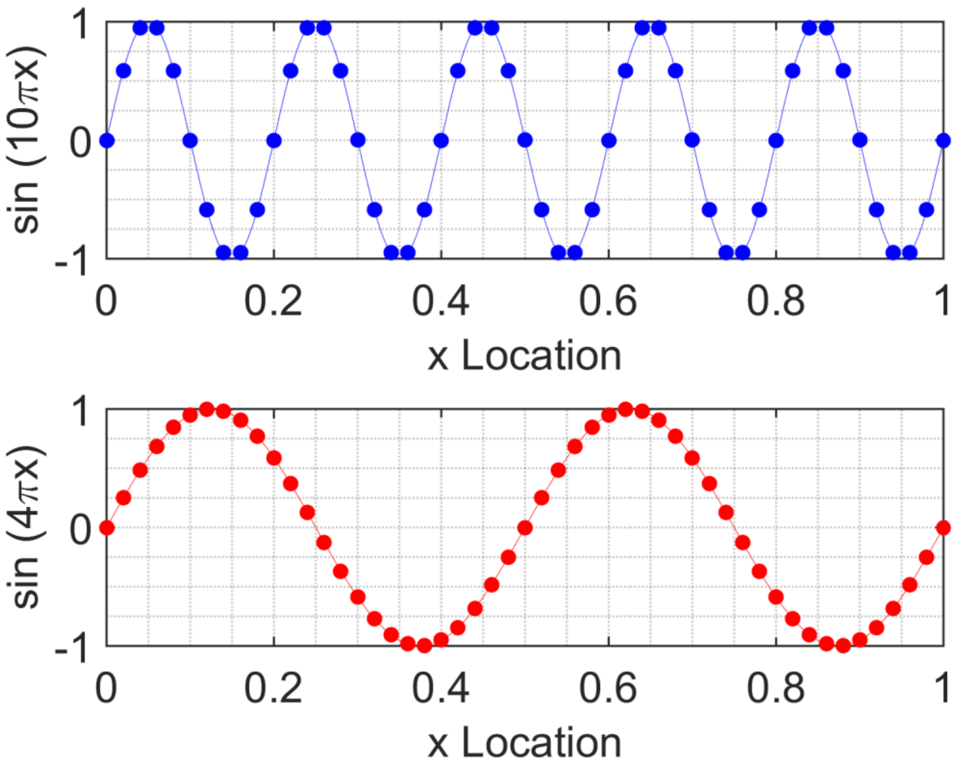
Least-squares spectral analysis

The least-squares spectrum (variance or energy) is defined as:

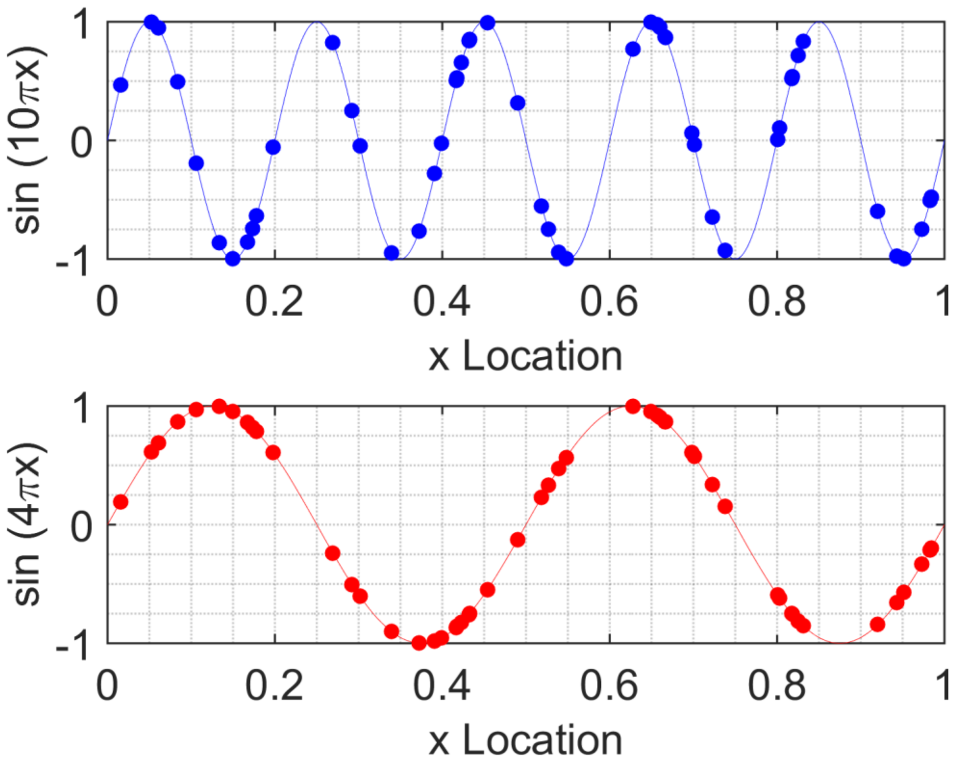
$$s(\omega_k) = \frac{\boldsymbol{f}^T \boldsymbol{\phi}_k \hat{\boldsymbol{c}}_k}{\boldsymbol{f}^T \boldsymbol{f}} \in (0,1)$$

that is the ratio of estimated signal to the total (the sum of estimated signal and noise) and follows the beta distribution.

Orthogonality (correlation) of sinusoids of different frequencies



$$\sum_{k=1}^{51} \sin(10\pi x_k)\sin(4\pi x_k) \cong 0$$



$$\sum_{k=1}^{51} \sin(10\pi x_k)\sin(4\pi x_k) \cong -0.5$$

Anti-leakage least-squares spectral analysis

In the LSSA, each frequency is examined separately!

Therefore, the correlations among the sinusoids of different frequencies are not considered.

What can we do to account for these correlations?

Anti-leakage least-squares spectral analysis

- We can add the sinusoids of different frequencies to the **design** matrix by an iterative method.

- In the iterative method, the sinusoids of a frequency that present highest energy in the spectrum will be added to the **design** matrix first.

- The iterative method stops when there is no significant peak at 95% confidence level (commonly used) in the residual data spectrum.

Example 1: The irregularly spaced data series

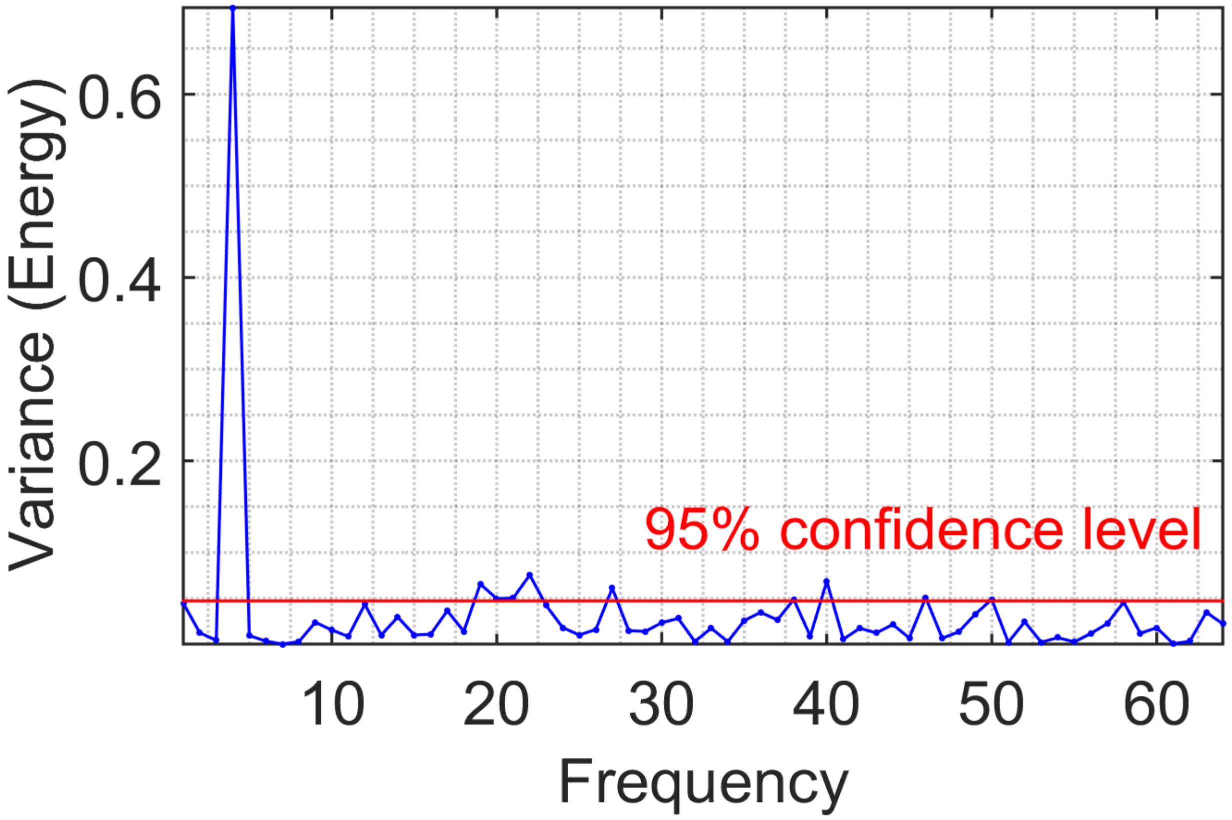
$$f(x_\ell) = 5\sin(25.6\, x_\ell) + 2.5\ \sin(128\, x_\ell + 1) + \sqrt{3}\ \sin(140\, x_\ell) + \sqrt{2} + \pi x_\ell$$

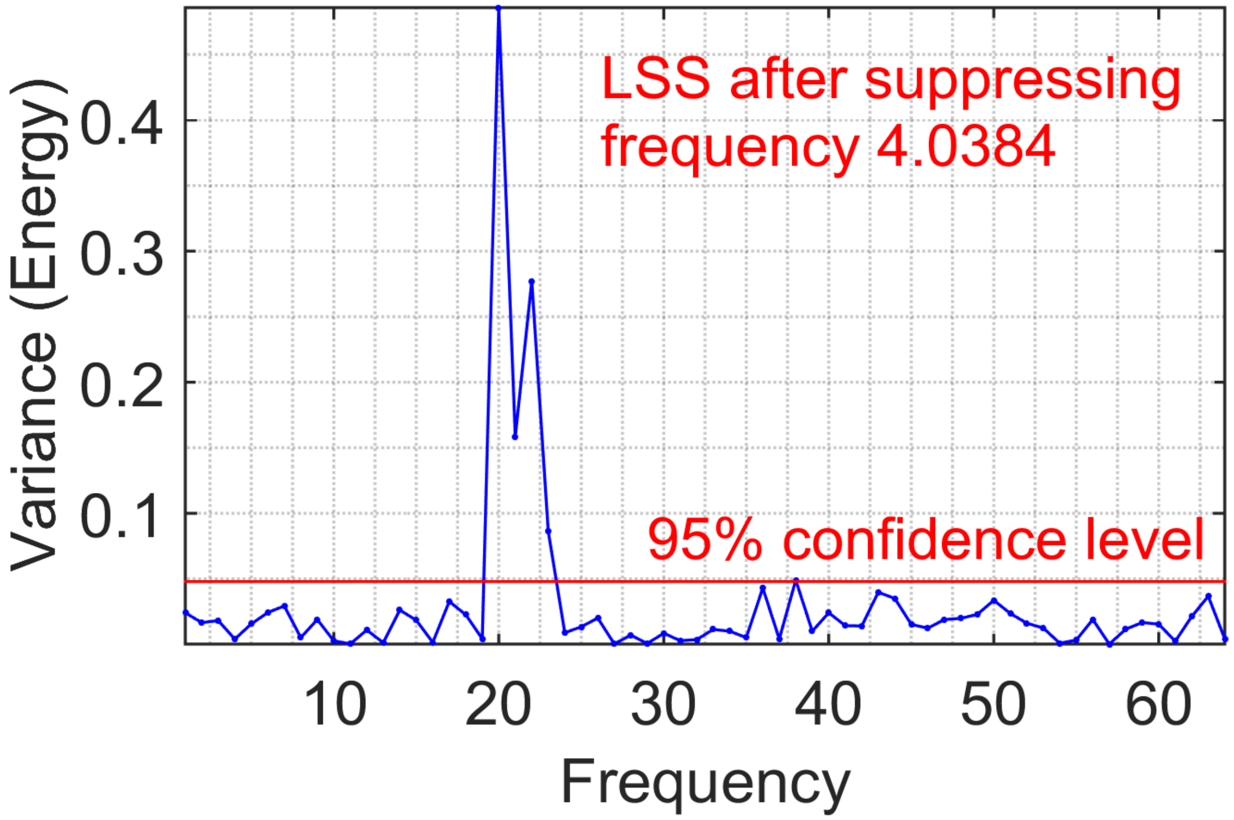
$\omega_1 \cong 4.0743665$

$\omega_3 \cong 22.2816920$

$\omega_2 \cong 20.3718327$

Example 1: Least-squares spectrum (variance) of the data series



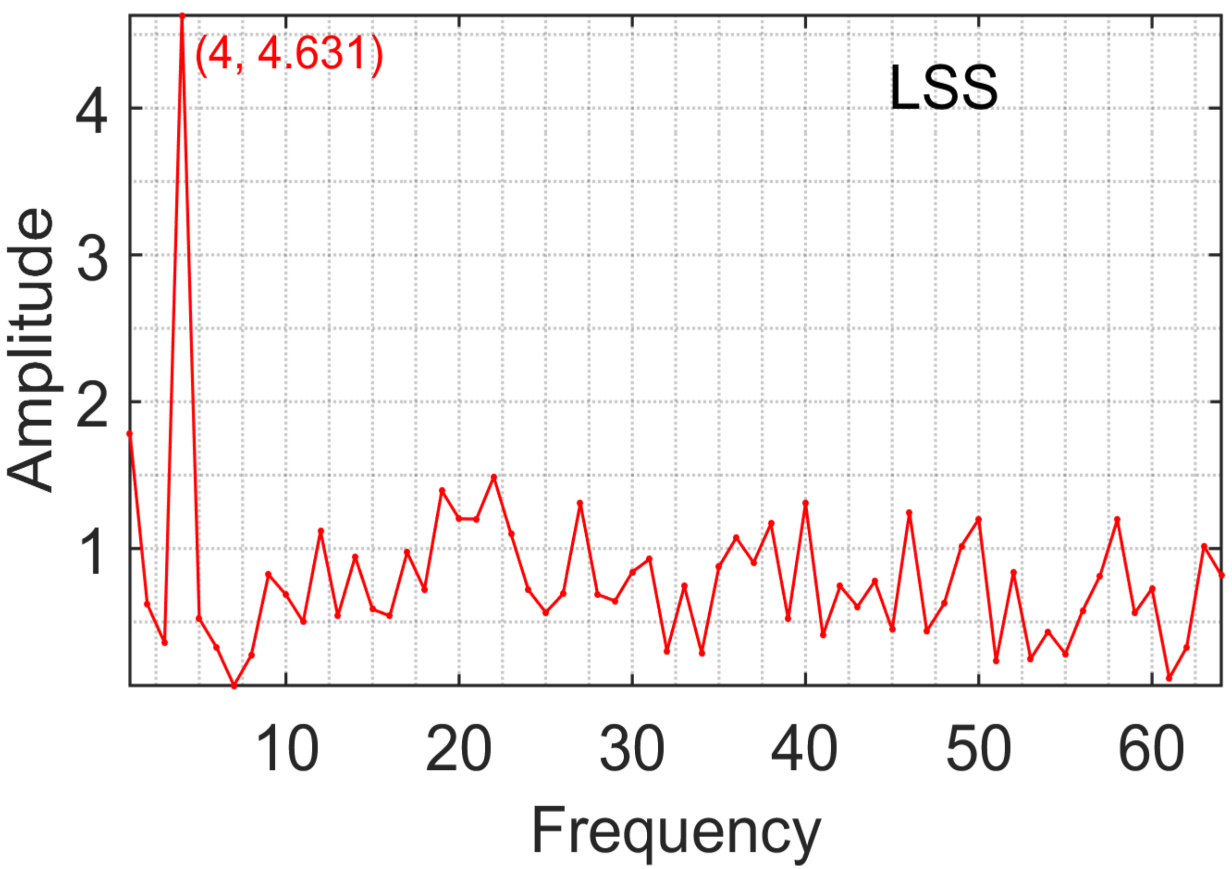


Example 1: Iterations in the anti-leakage LSSA

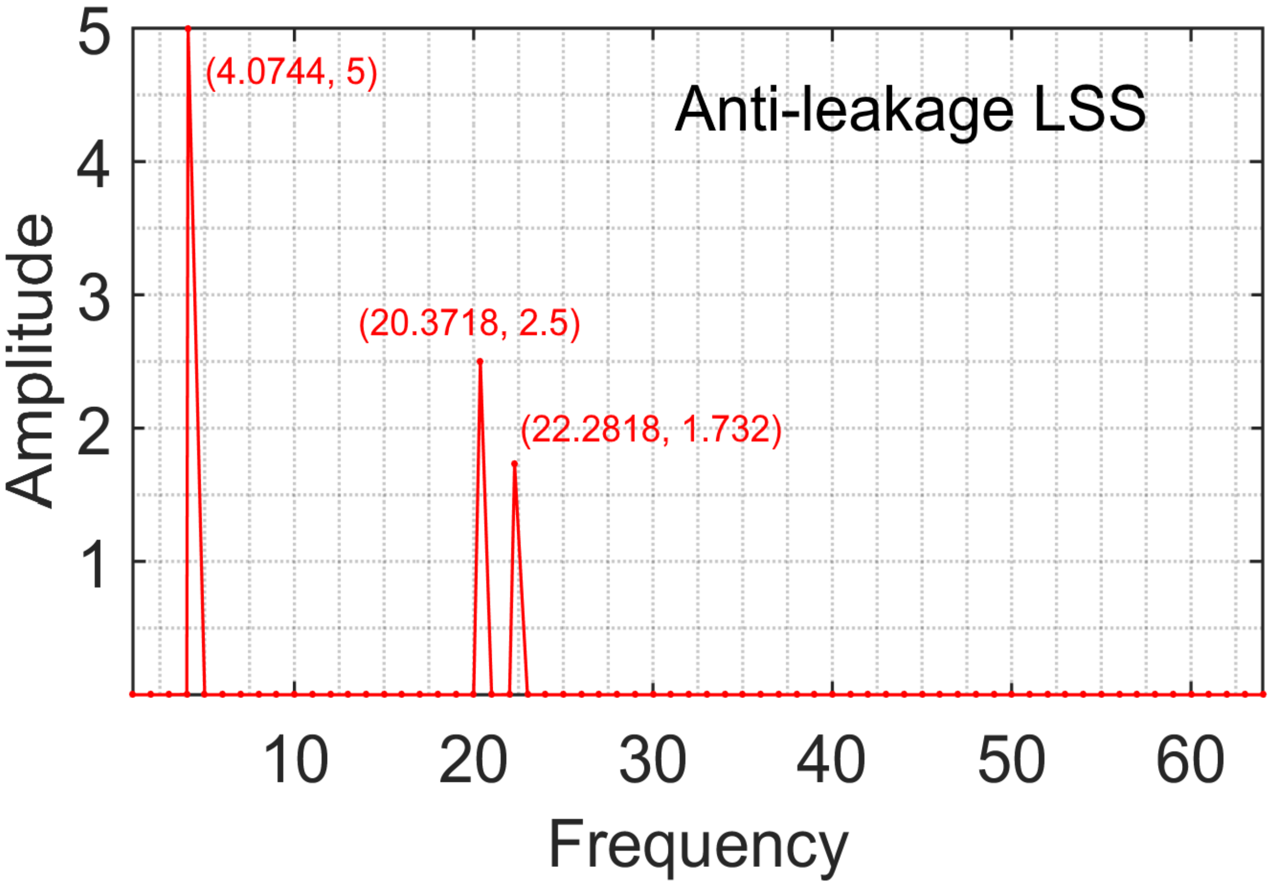
Iteration number	1st frequency	2nd frequency	3rd frequency	Norm of residual
First	4.0384			23.5040
Second	4.0384	20.3057		13.3170
Third	4.0384	20.3057	22.2947	3.5079
Fourth	**4.0748**	20.3057	22.2947	2.2501
Fifth	4.0748	**20.3708**	22.2947	0.3068
Sixth	4.0748	20.3708	**22.2818**	0.0468
Seventh	4.0748	**20.3718**	22.2818	0.0319
Eighth (final)	**4.0744**	20.3718	22.2818	0.0036

Example 1: Least-squares spectrum (LSS) and Anti-leakage LSS

Before

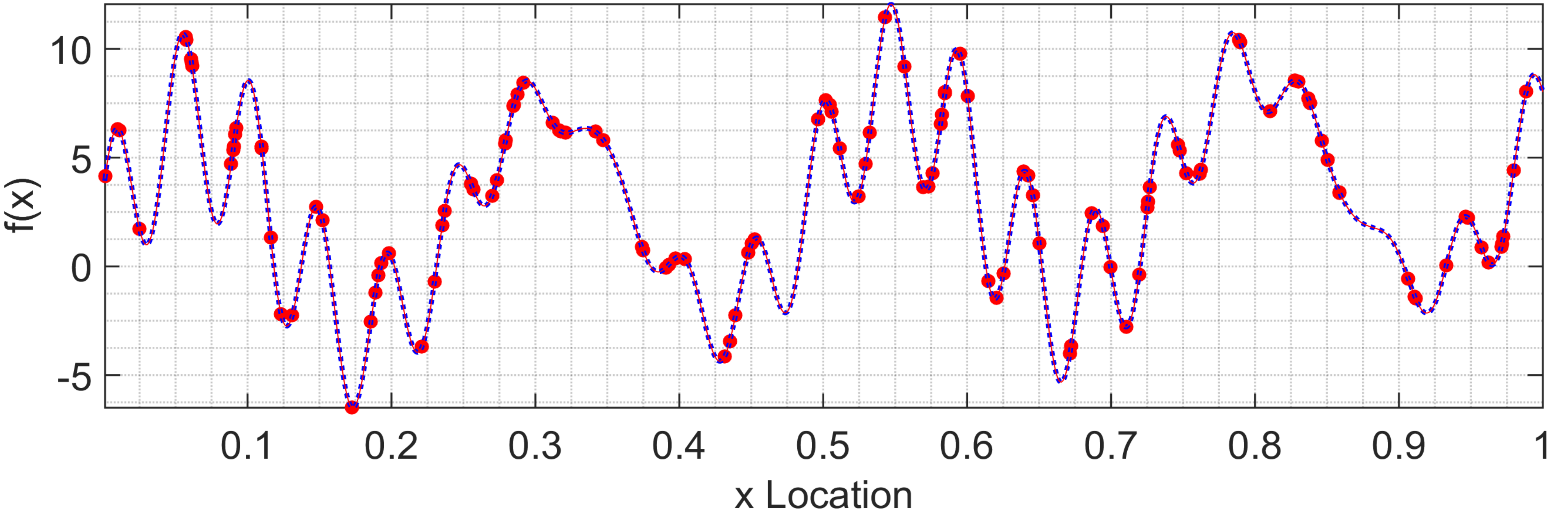


After



Example 1: The interpolation result

The irregularly spaced data series (red dots) and its interpolation result (blue dots)



Application (Seismic trace interpolation)

Transform each trace from time domain to frequency domain using the Fast Fourier Transform (FFT)

→

Generate a data series using the available traces for each frequency (a frequency slice)

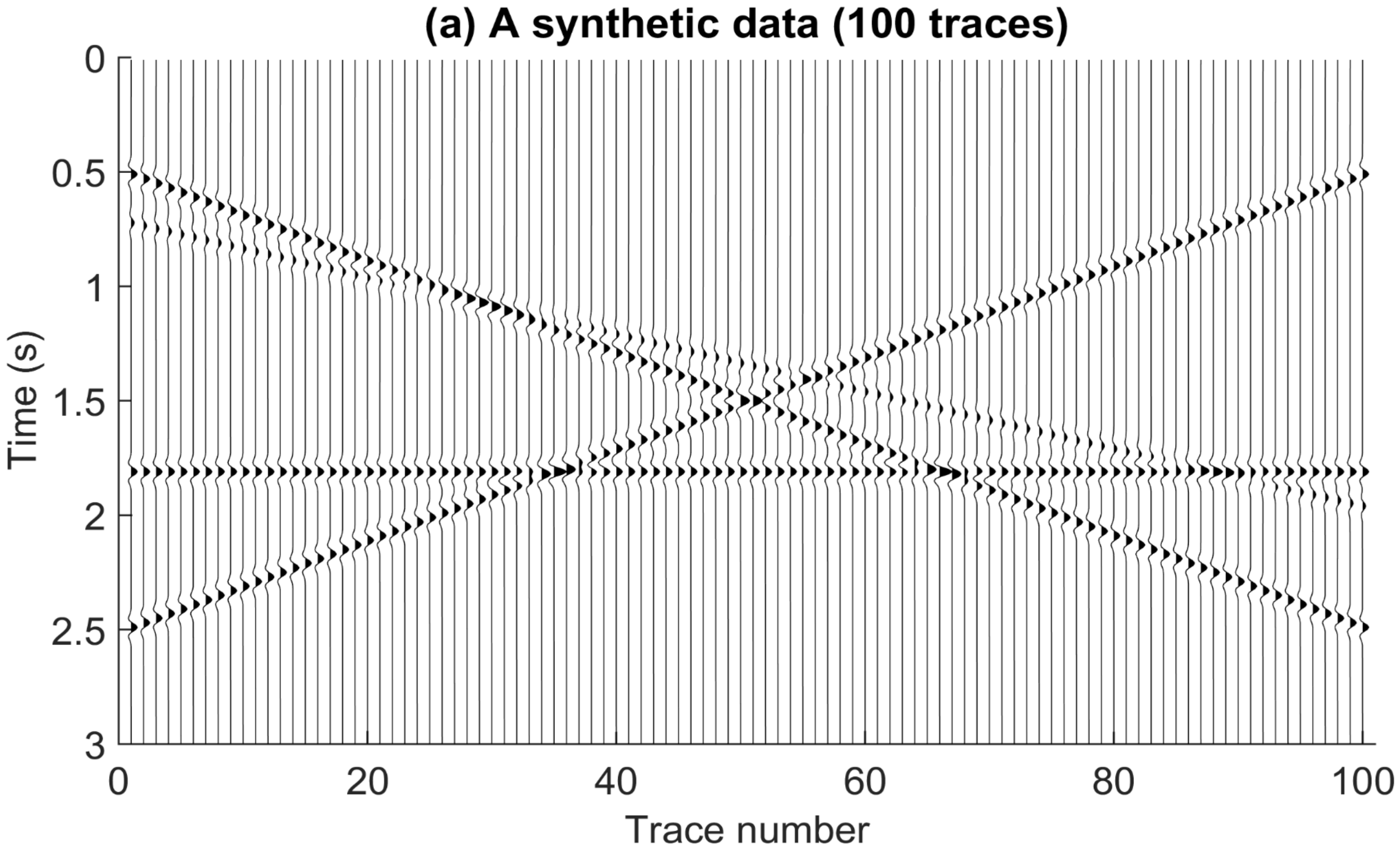
↓

Use the anti-leakage LSSA to regularize each frequency slice

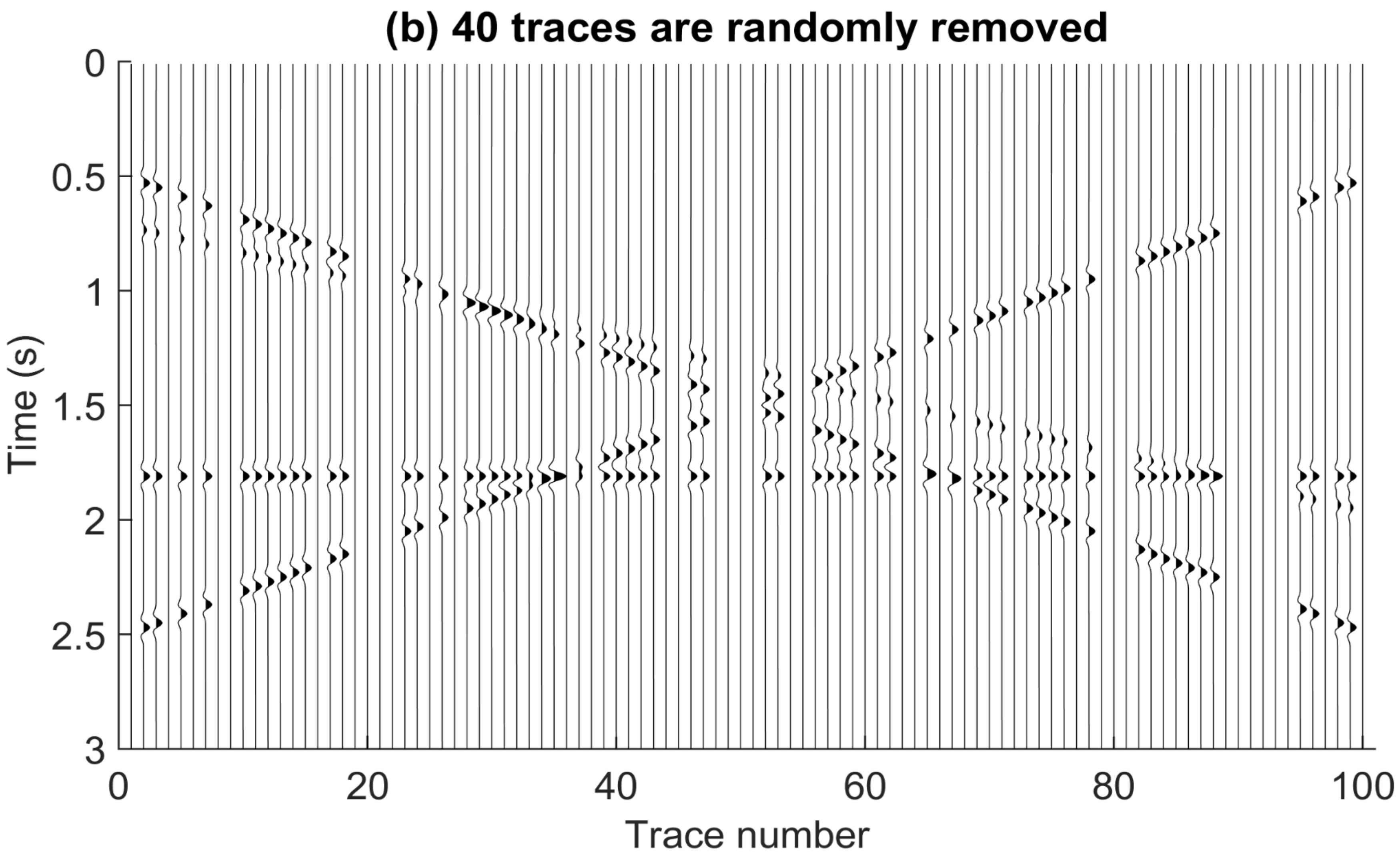
→

Transform each trace from frequency domain to time domain using the Inverse FFT

Example 2: A synthetic seismic data containing four events

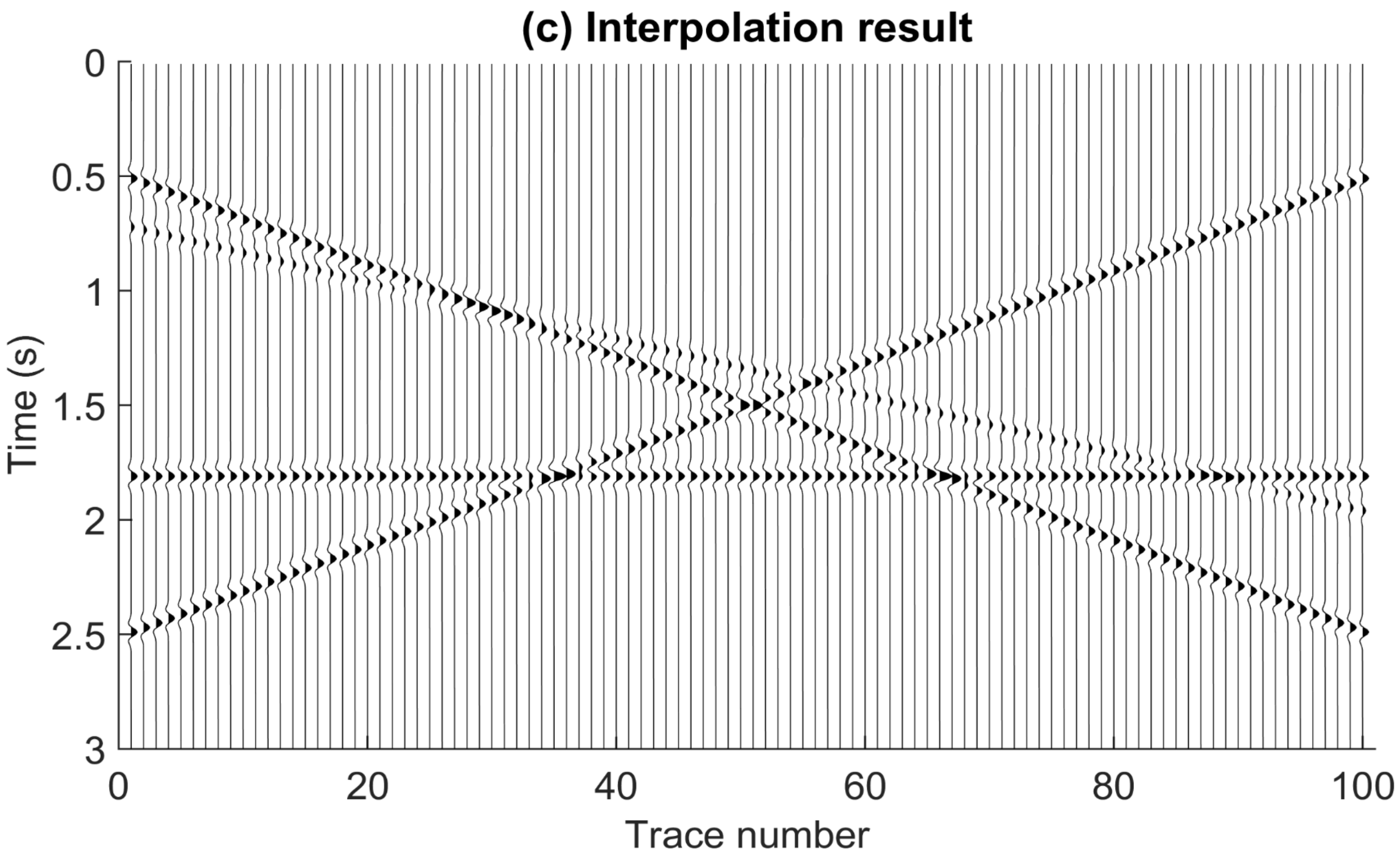


Example 2:



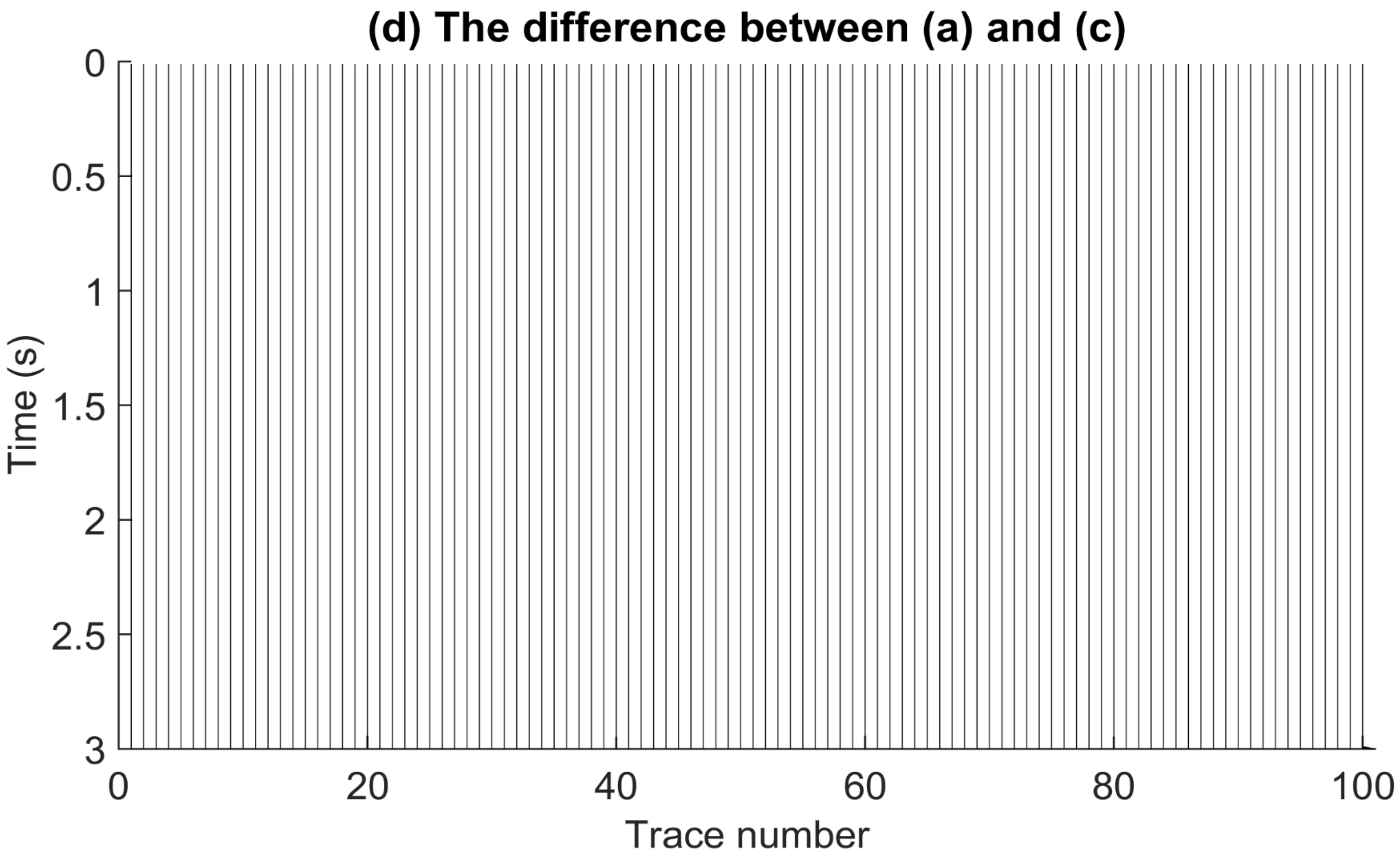
(b) 40 traces are randomly removed
Time (s)
0
0.5
1
1.5
2
2.5
3
0
20
40
60
80
100
Trace number

Example 2:



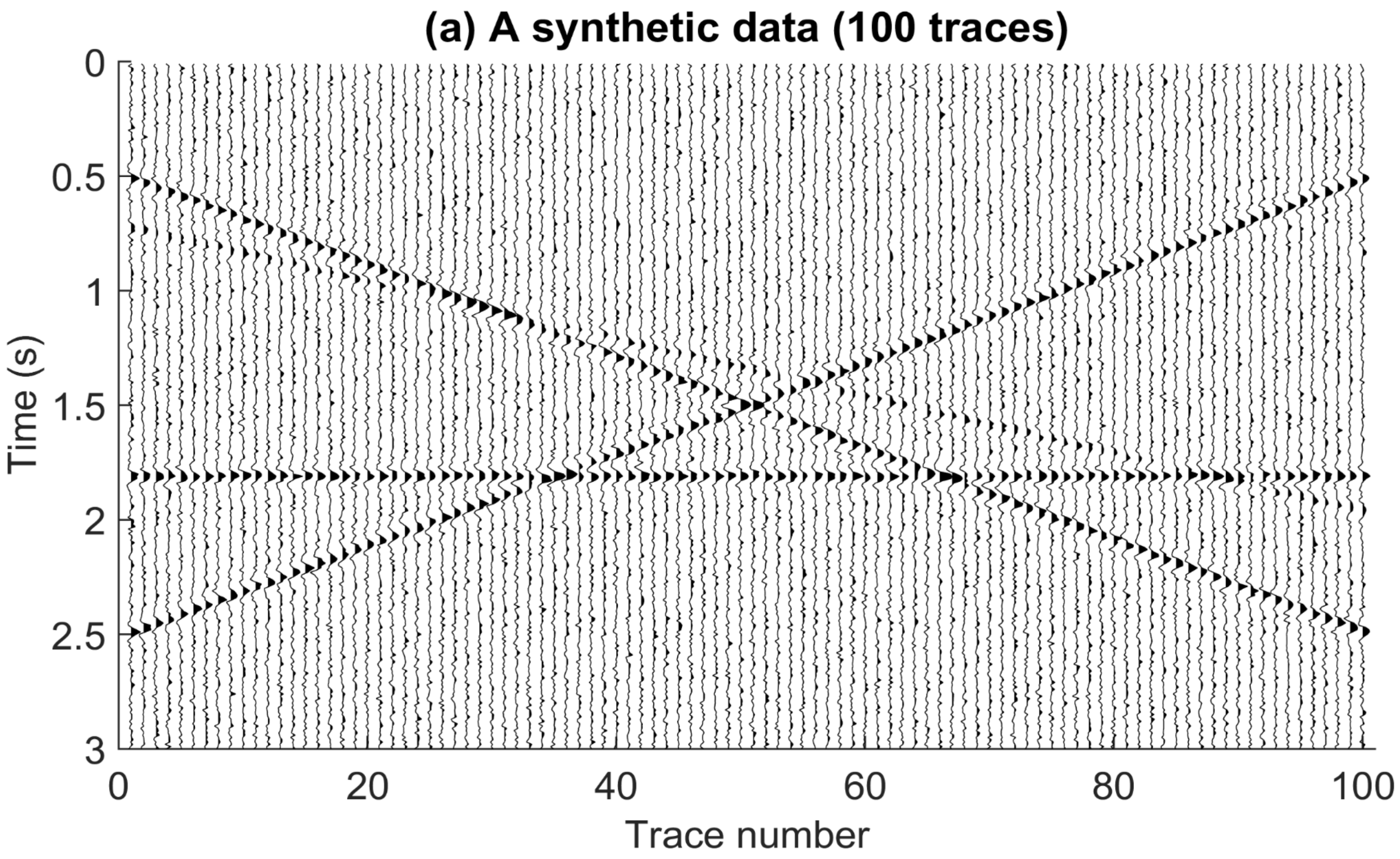
(c) Interpolation result
Time (s)
0
0.5
1
1.5
2
2.5
3
0
20
40
60
80
100
Trace number

Example 2:

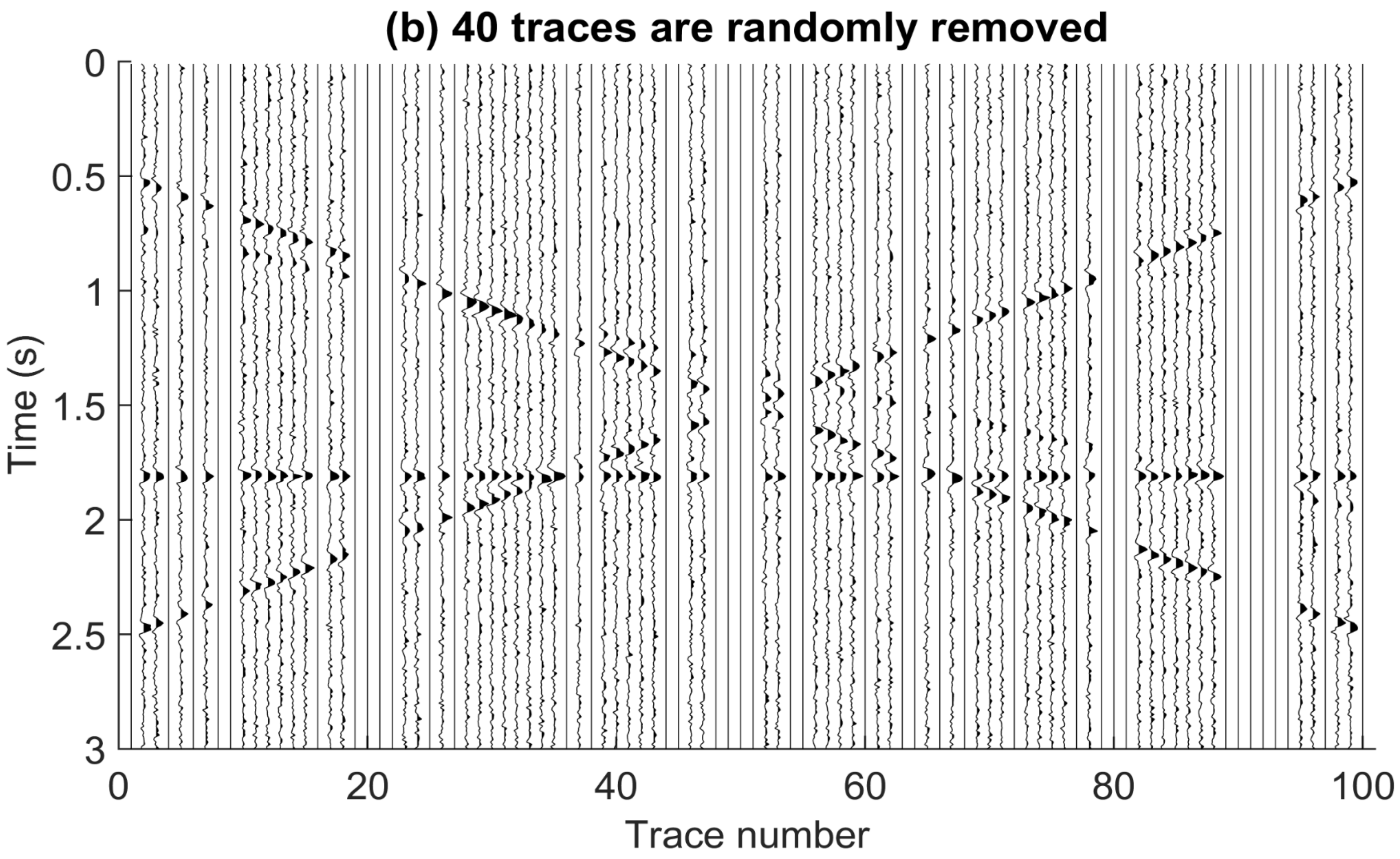


(d) The difference between (a) and (c)
0
0.5
1
1.5
2
2.5
3
Time (s)
0
20
40
60
80
100
Trace number

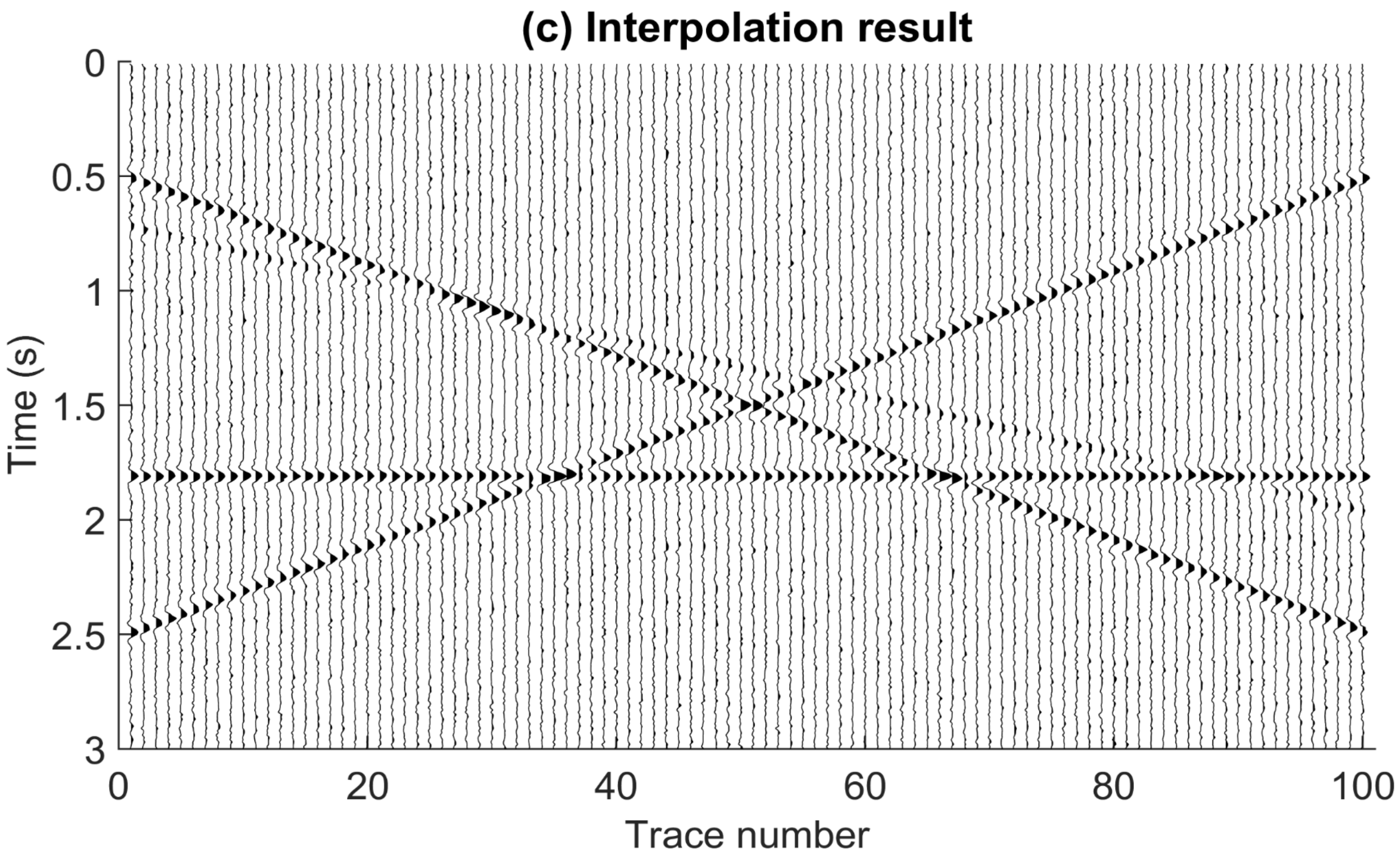
Example 3: A synthetic seismic data containing four events



Example 3:

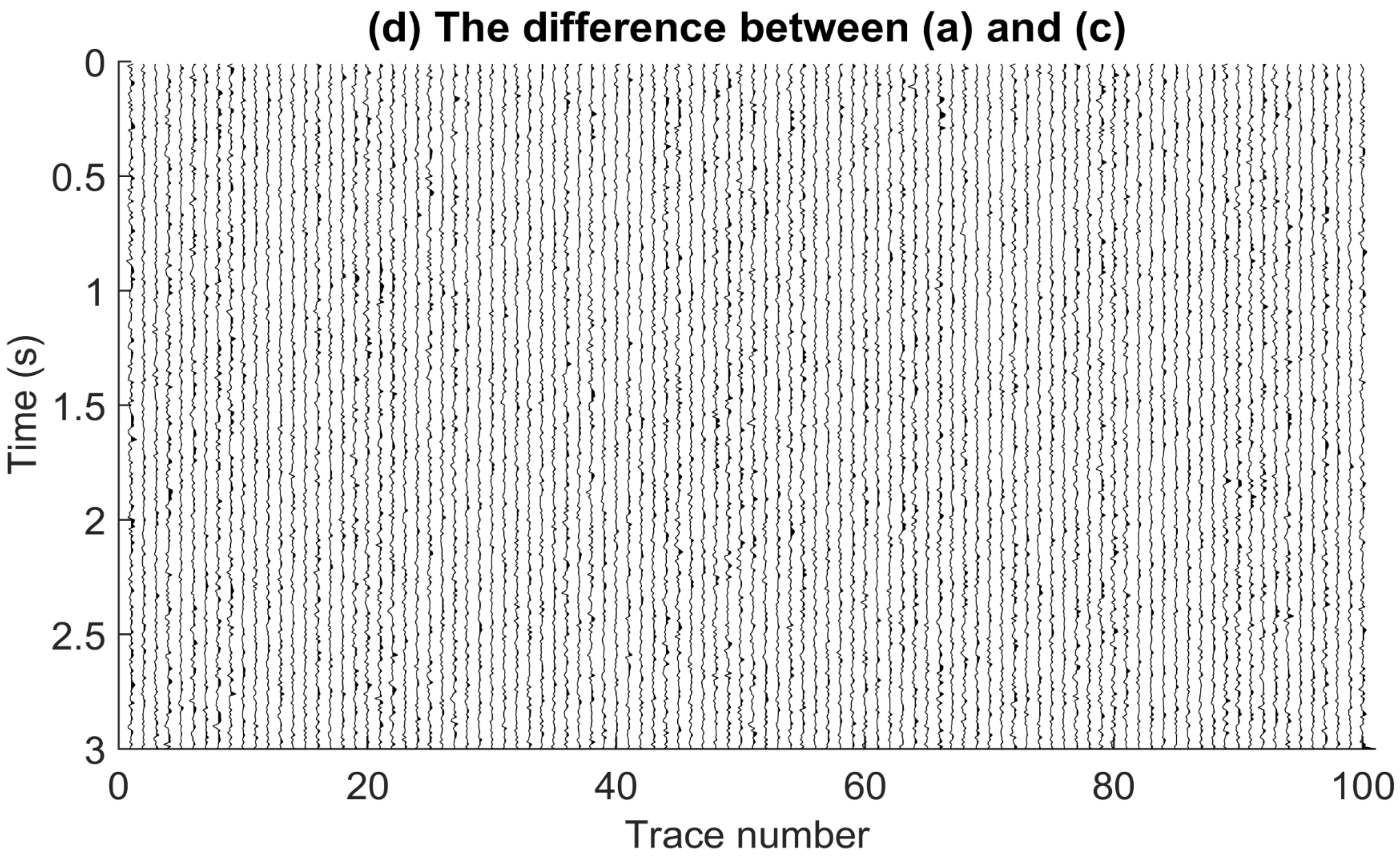

(b) 40 traces are randomly removed
Time (s)
0
0.5
1
1.5
2
2.5
3
0
20
40
60
80
100
Trace number

Example 3:



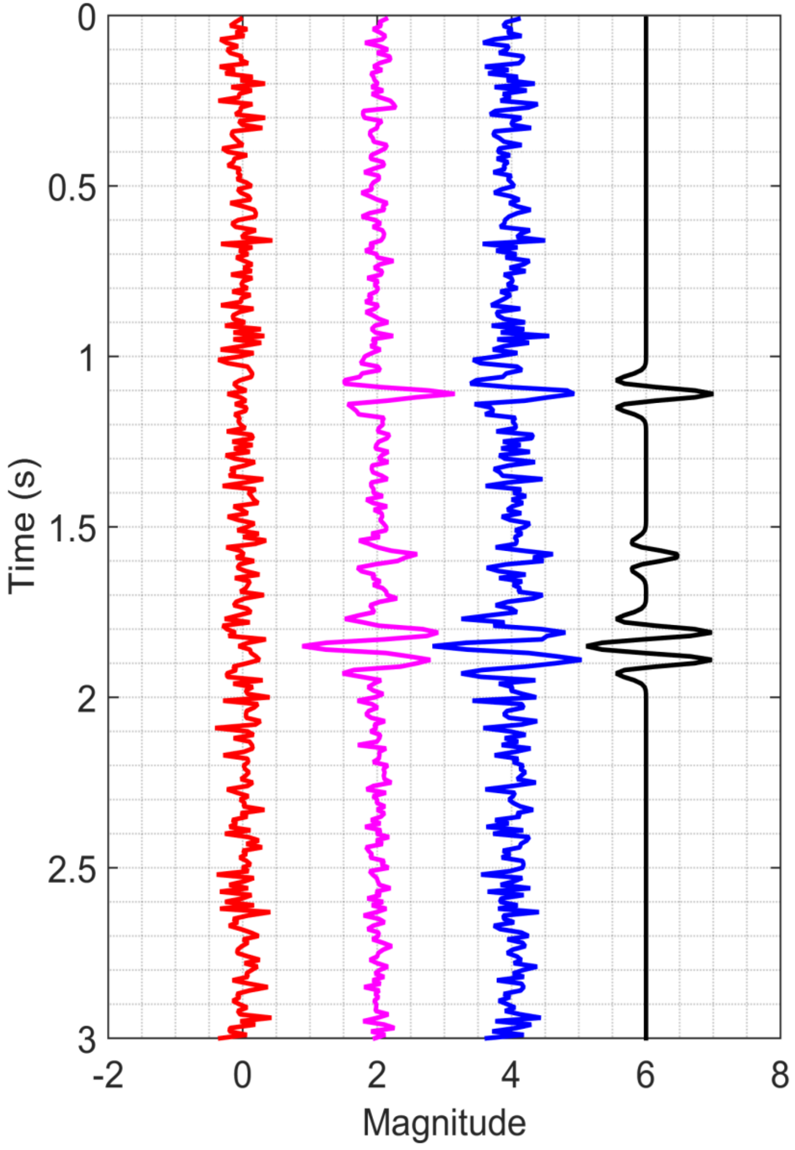
(c) Interpolation result
Time (s)
0
0.5
1
1.5
2
2.5
3
0
20
40
60
80
100
Trace number

Example 3:



(d) The difference between (a) and (c)
0
0.5
1
1.5
2
2.5
3
Time (s)
0
20
40
60
80
100
Trace number

Example 3: Examining trace #70 (as an example)



Without noise

With Gaussian white noise

The Interpolation result

The difference

Conclusion

- The anti-leakage LSSA is a very accurate method to regularize an irregularly spaced data series when it is ***stationary***.

- The proposed method is able to detect signals from noise up to a certain confidence level (usually 95%).

- The FFT is computationally faster than the anti-leakage LSSA; however, we cannot use the FFT here as our data series are irregularly spaced.

Acknowledgement

PIMS

NSERC
CRSNG

Postdoctoral Program
Department of Math and Stat

UNIVERSITY OF
CALGARY

Thank you for your attention!

Questions ???