Compressed sensing, random Fourier matrix and jittered sampling

Enrico Au-Yeung

(Joint work with Hassan Mansour, Rayan Saab and Ozgur Yilmaz)

SLIM

University of British Columbia

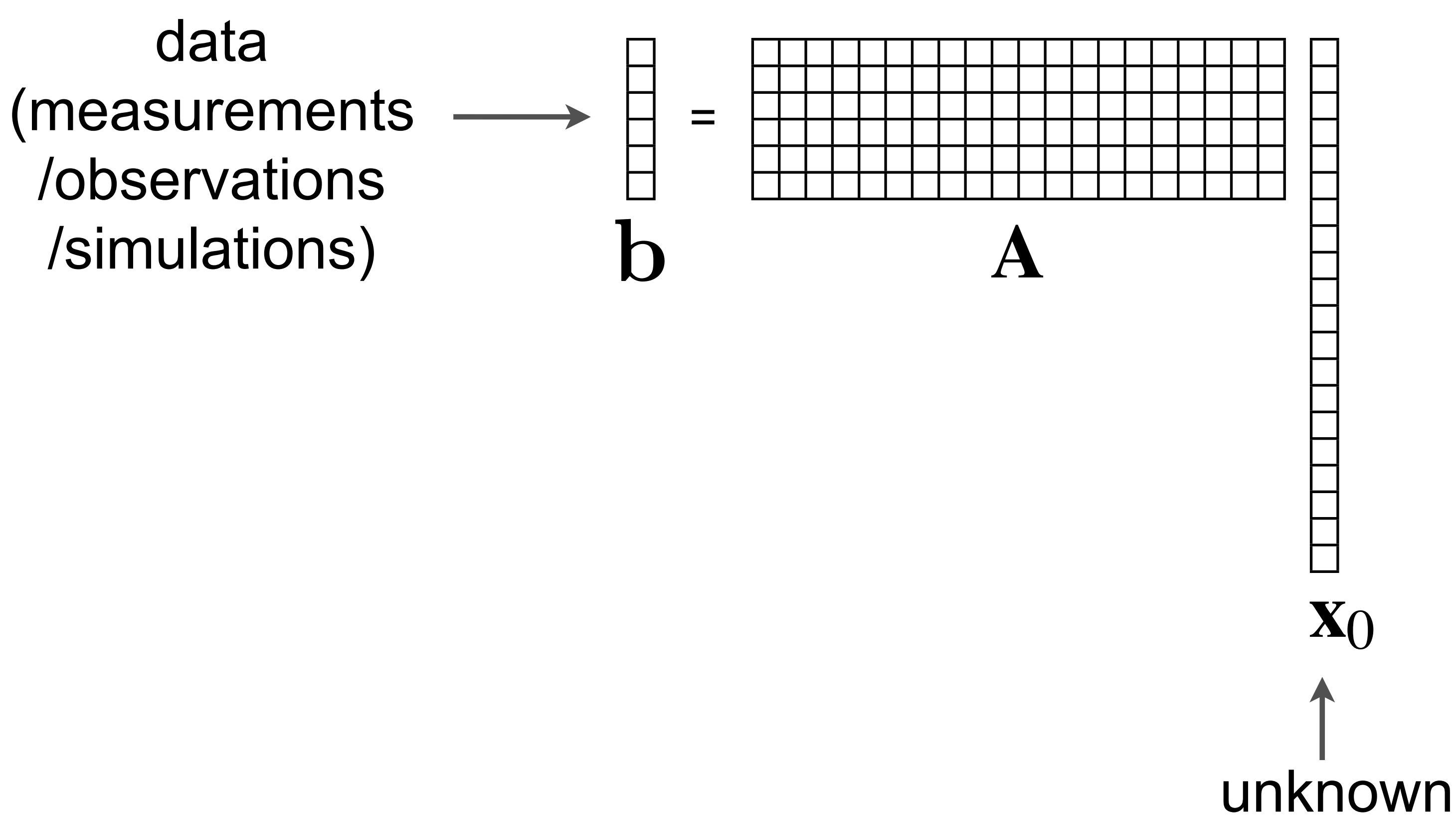
Introduction

Compressed sensing is a powerful technique to reconstruct sparse data.

- What is the main advantage of jittered sampling?
- What do we mean by better results?

Compressed sensing

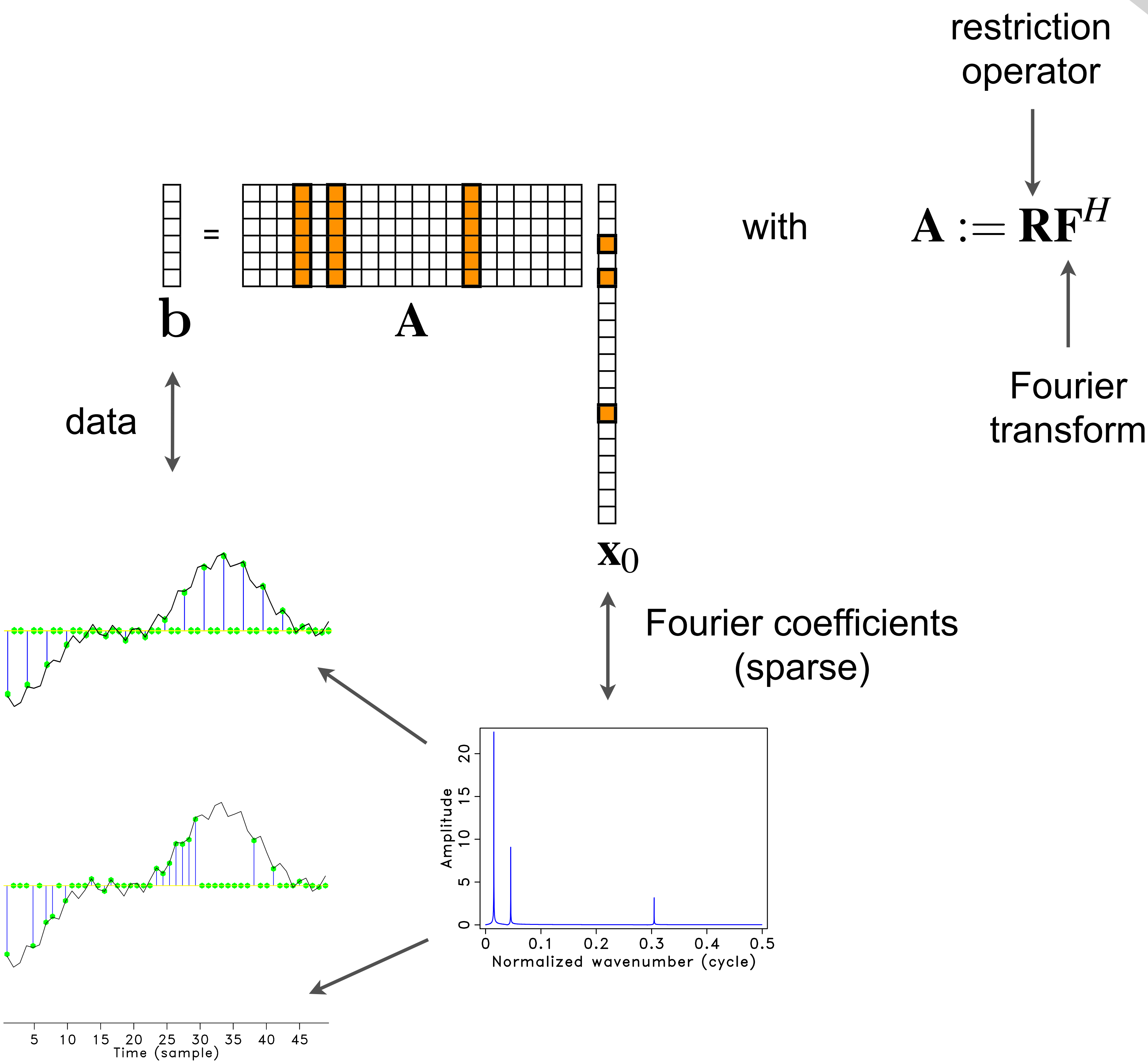
Consider the following (severely) *underdetermined* system of *linear* equations:



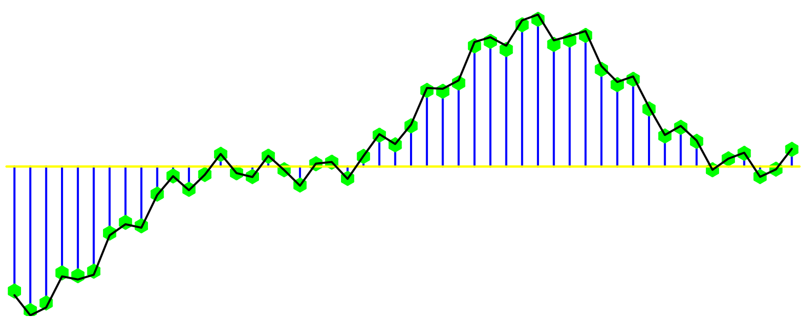
Is it possible to recover $\mathbf{x}_0$ accurately from $\mathbf{b}$?

Compressed Sensing attempts to answer this questions rigorously.

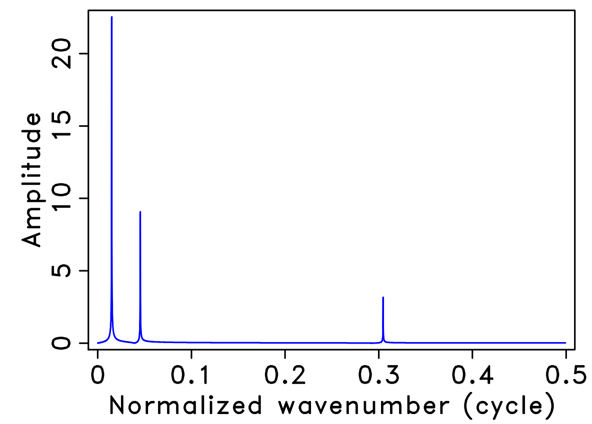
Sparse recovery



Coarse sampling schemes

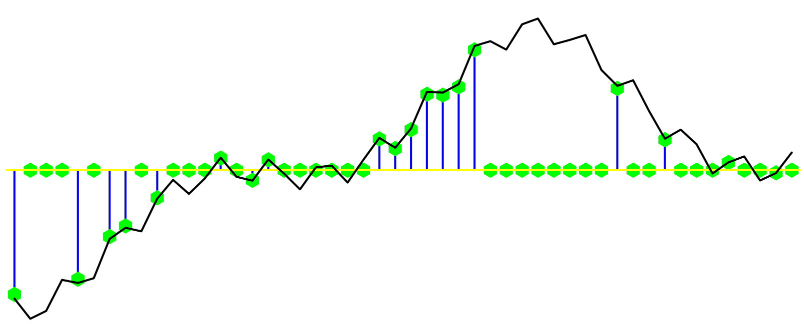


Fourier transform

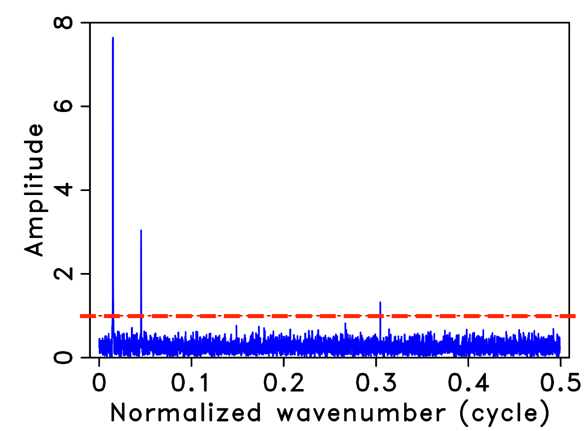


few significant coefficients

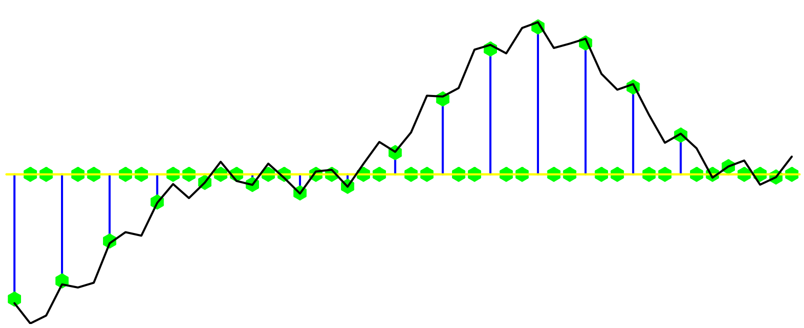
3-fold under-sampling



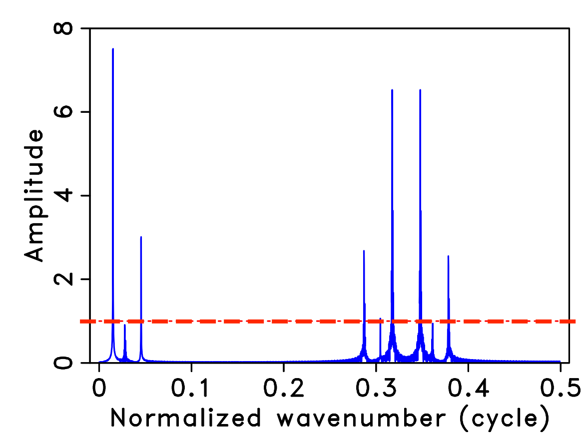
Fourier transform



✓ significant coefficients detected



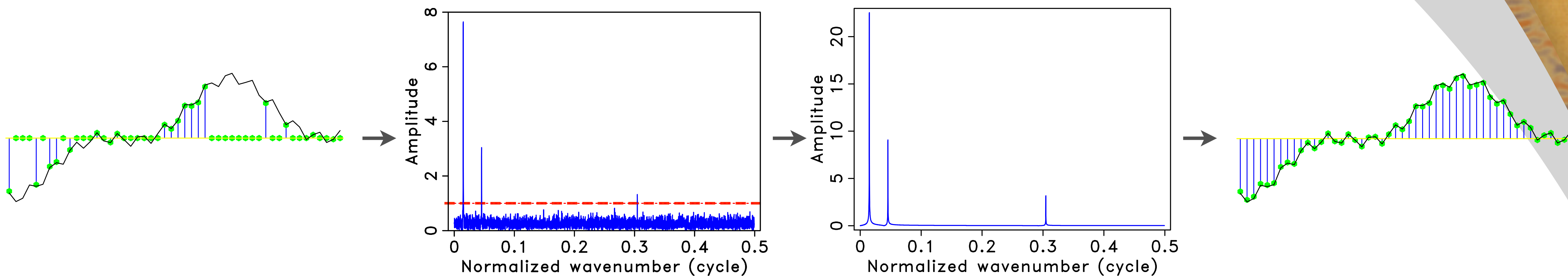
Fourier transform



✗ ambiguity

[Hennenfent & Herrmann, '08]

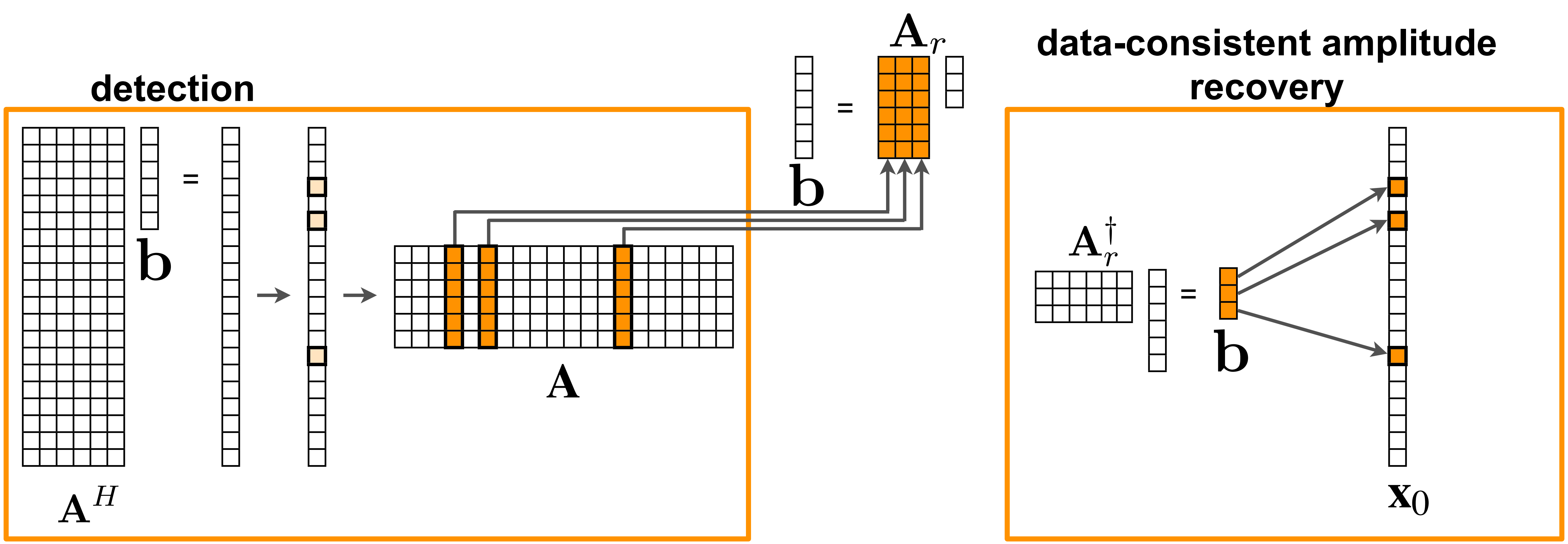
NAIVE sparsity-promoting recovery



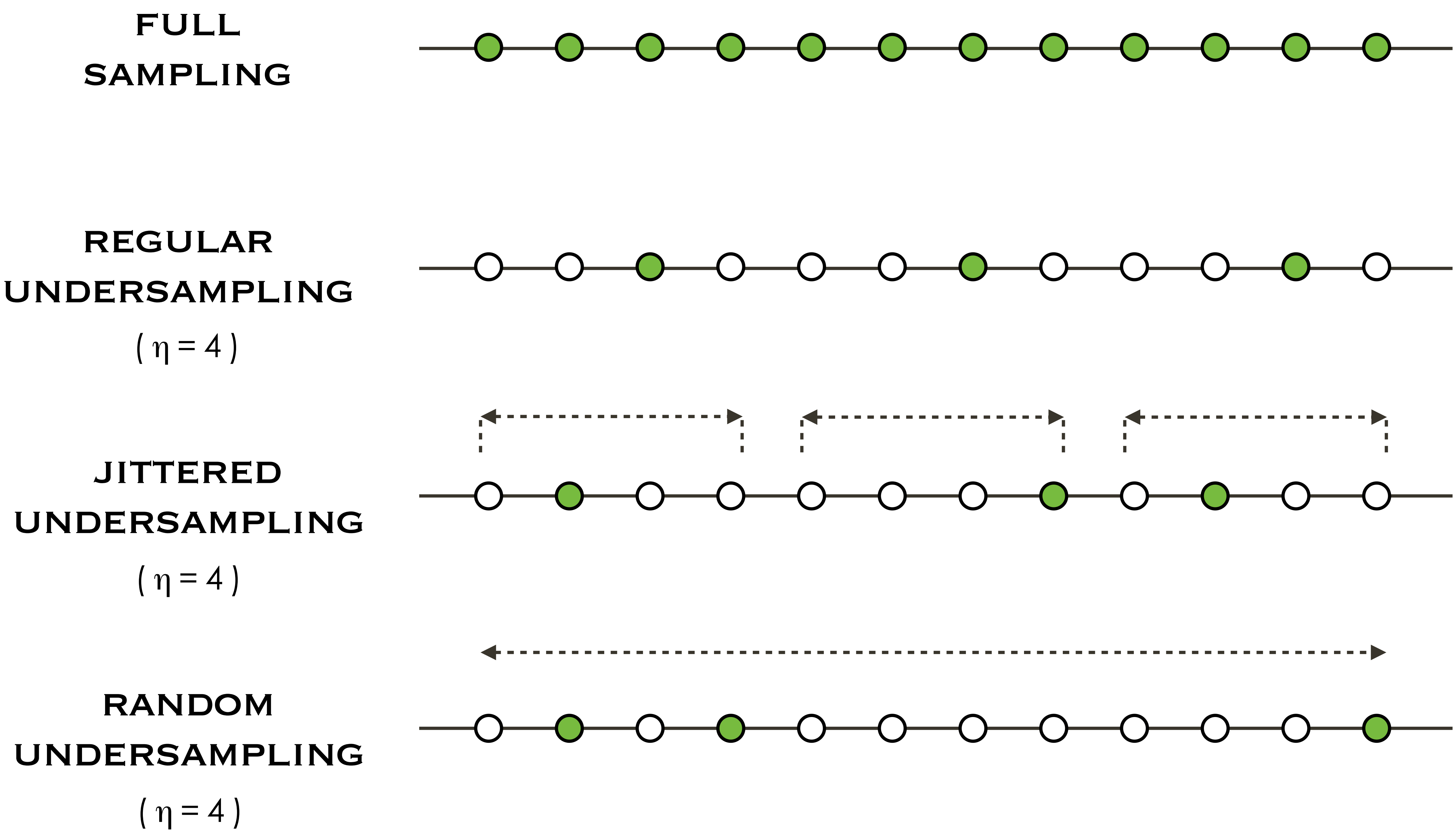
inverse Fourier transform

detection + data-consistent amplitude recovery

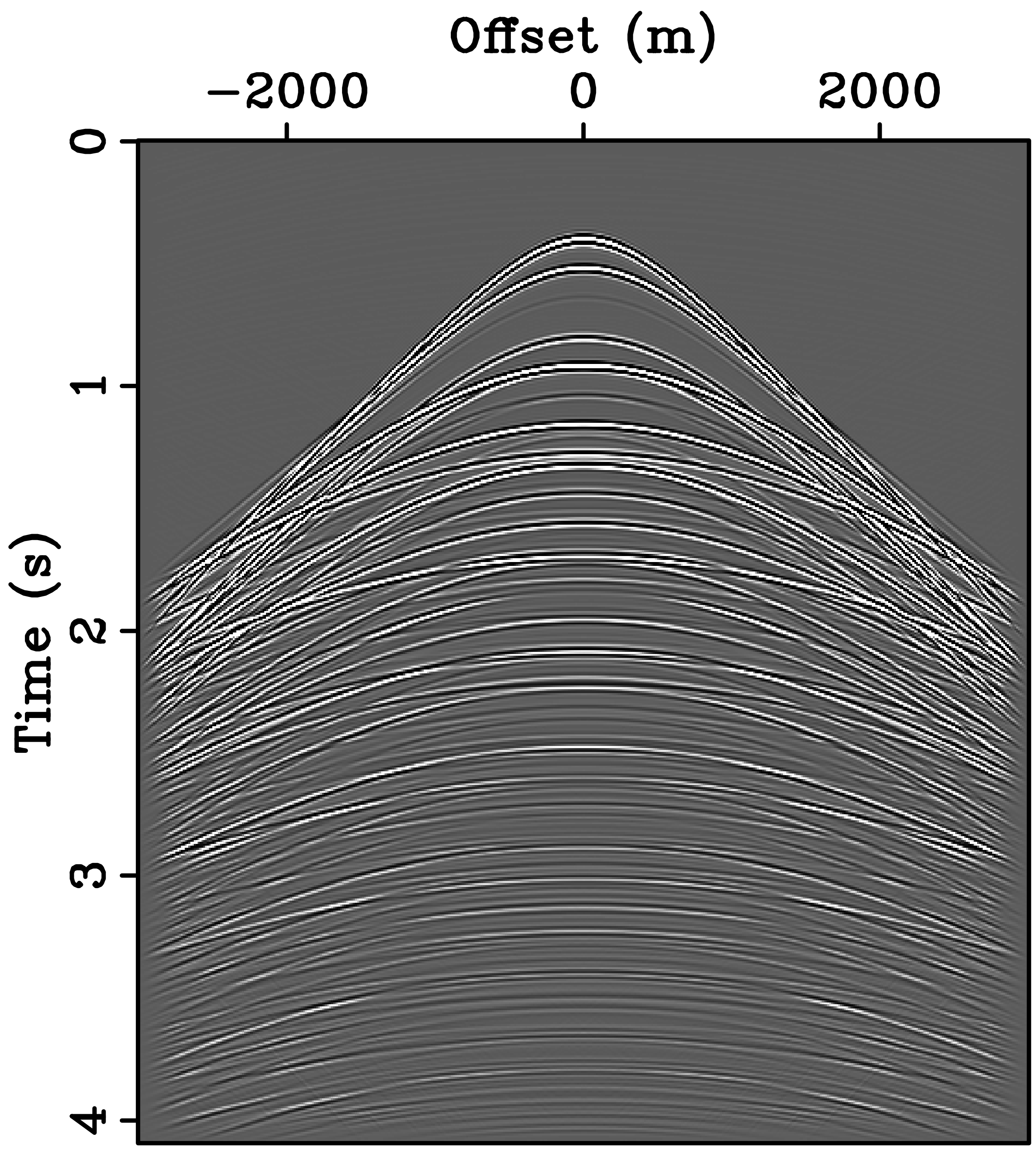
Fourier transform

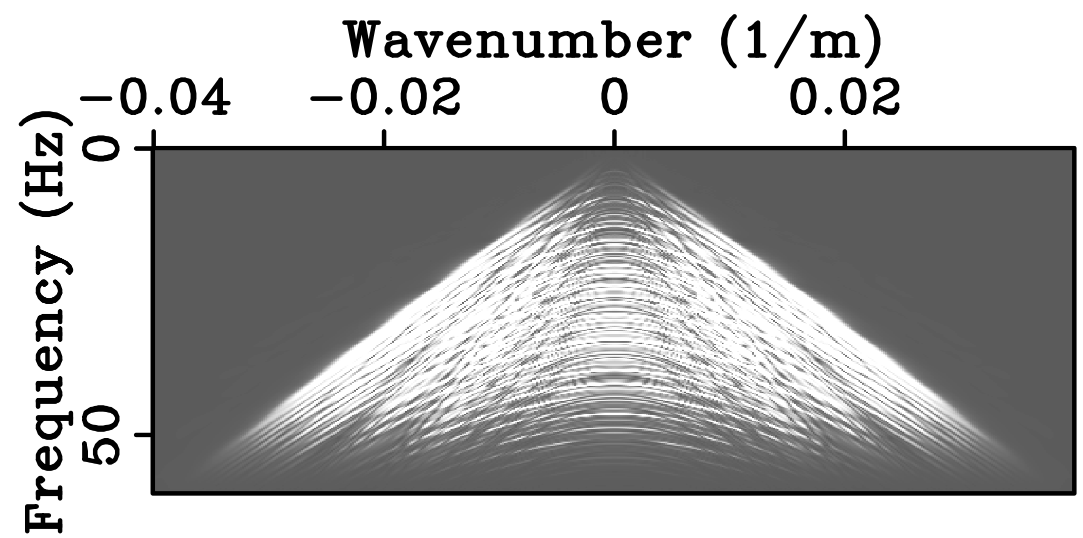


Sampling schemes

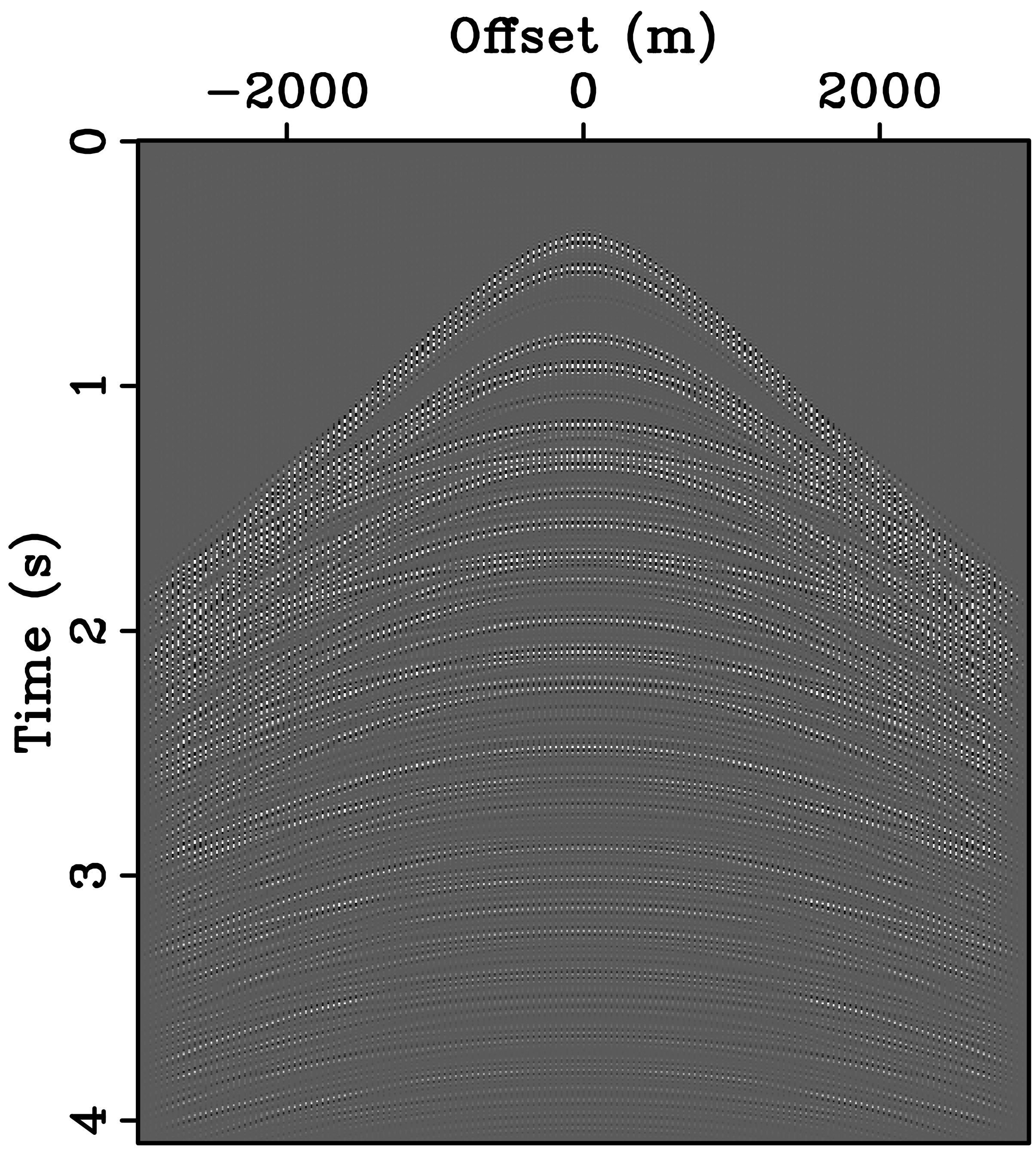


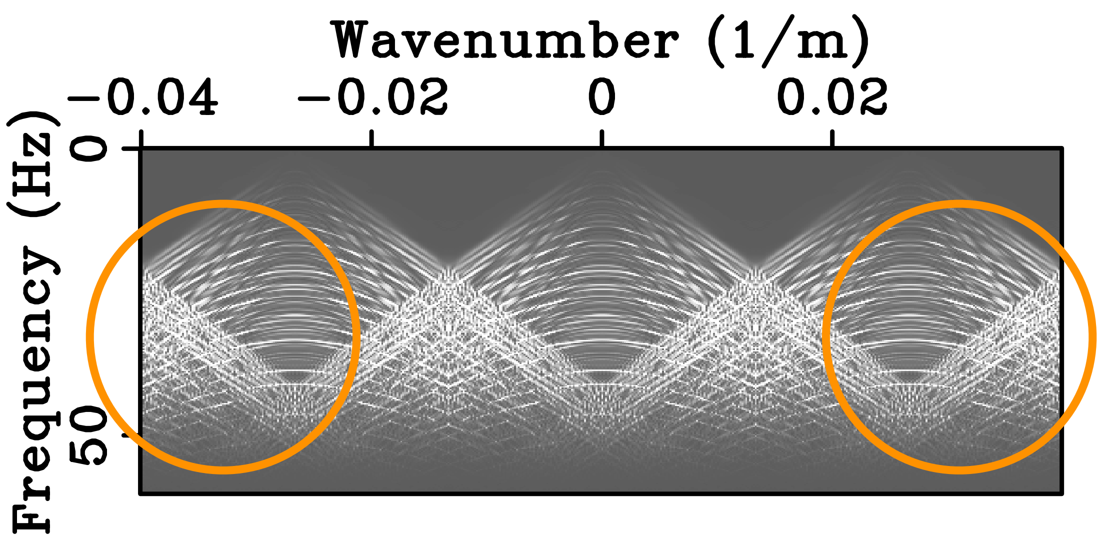
Model


Offset (m)
−2000
0
2000
Time (s)
0
1
2
3
4

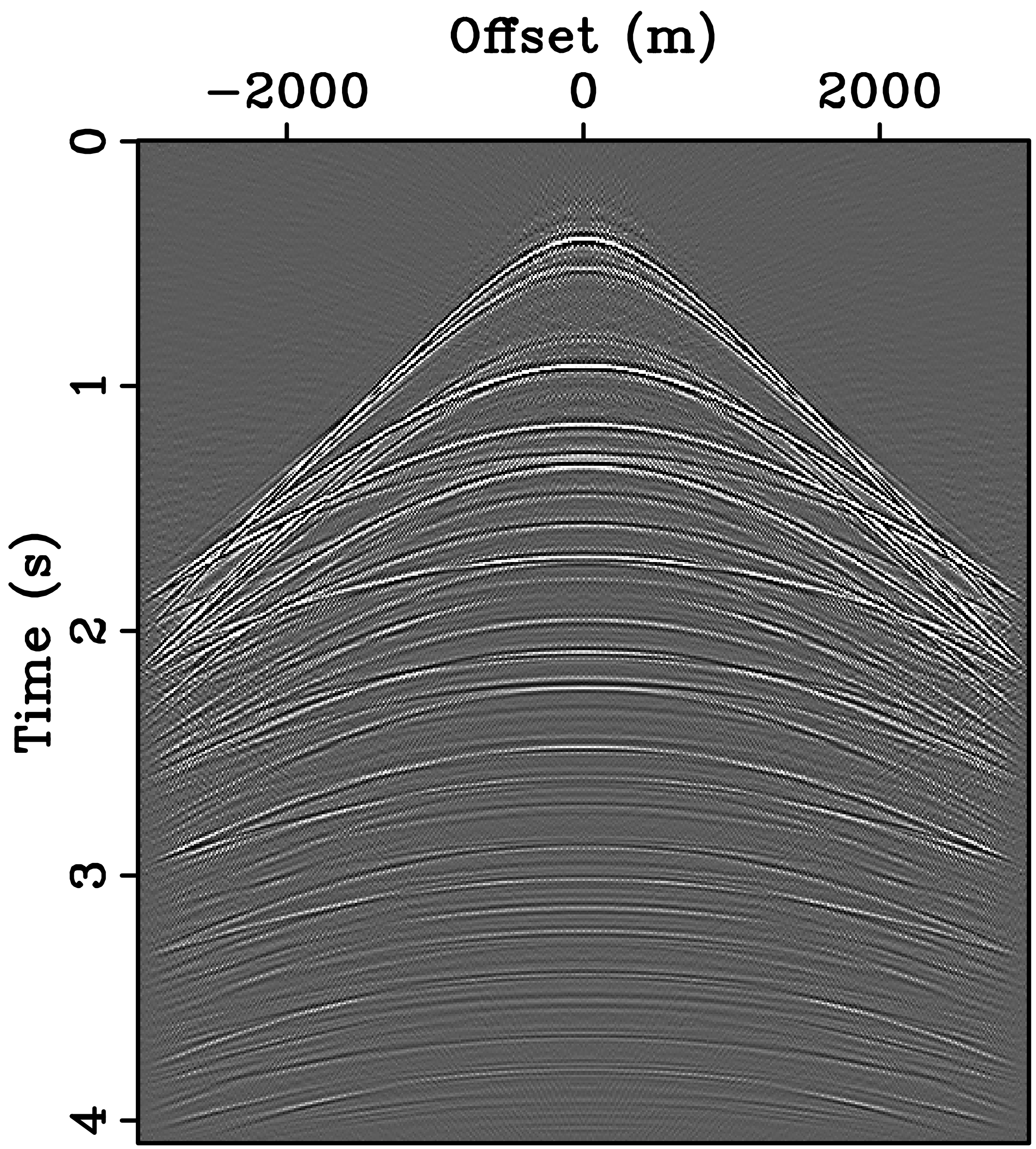

Wavenumber (1/m)
−0.04
−0.02
0
0.02
Frequency (Hz)
0
50

Regular 3-fold undersampling

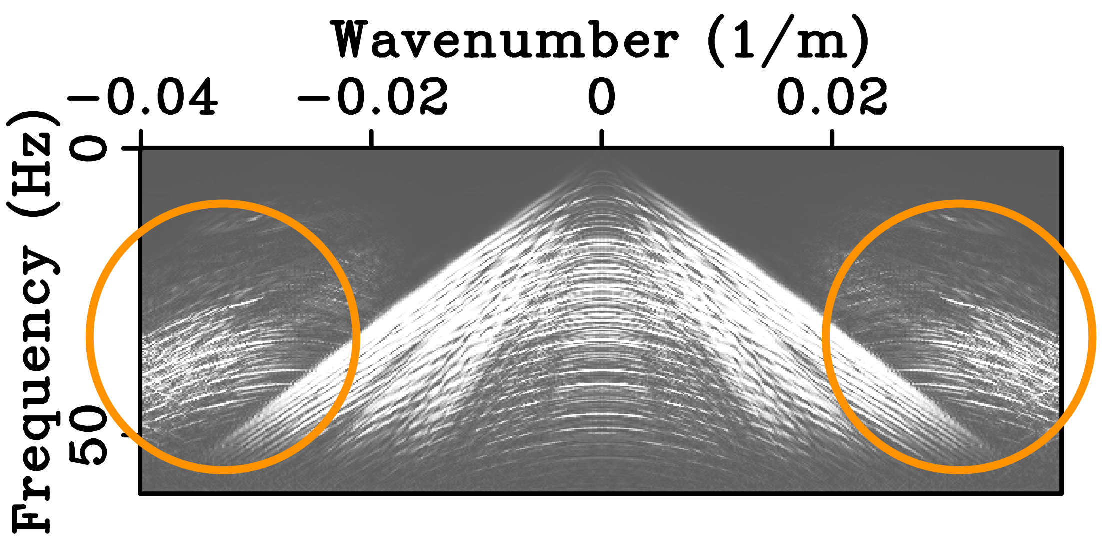




CRSI from regular 3-fold undersampling

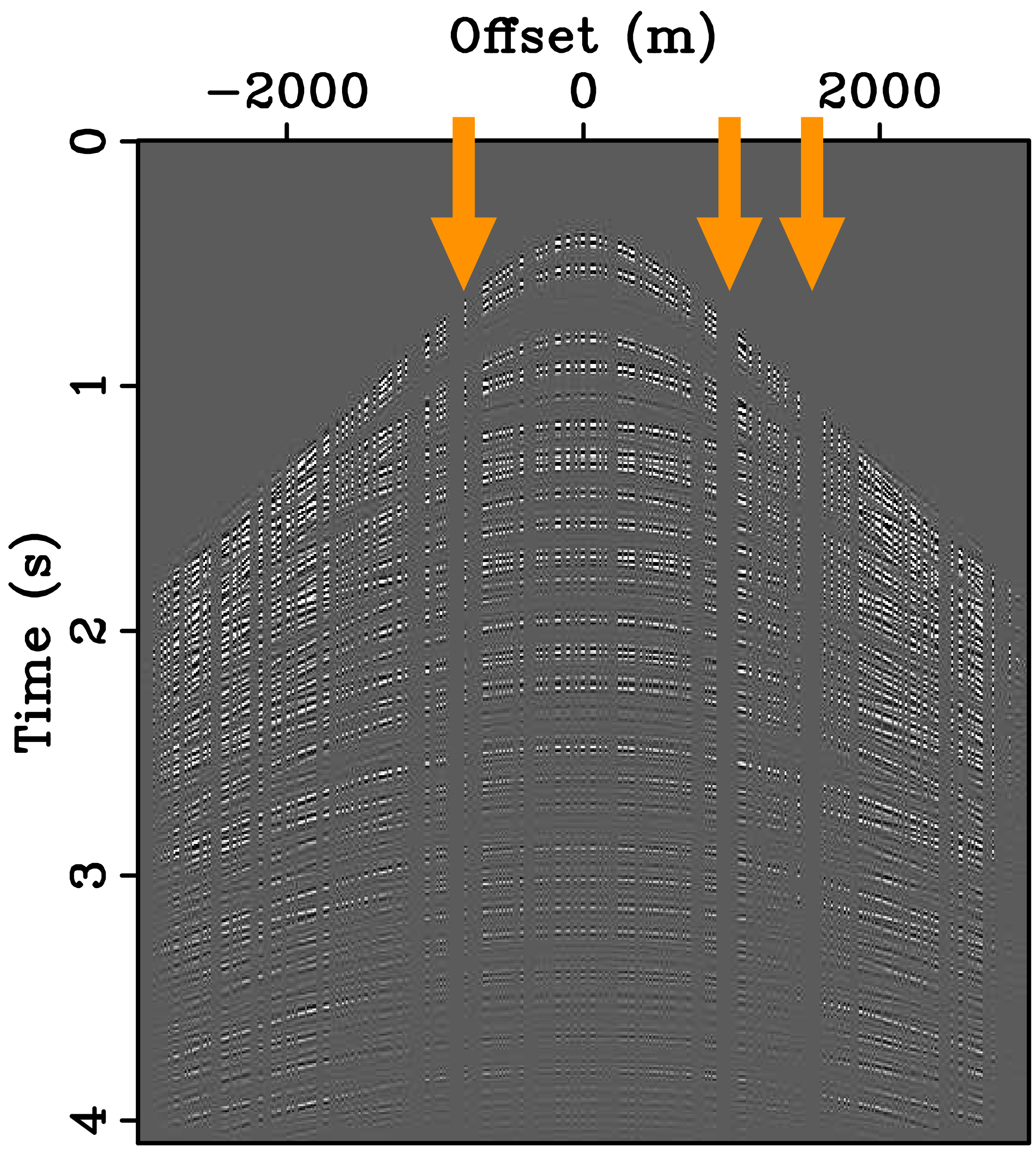


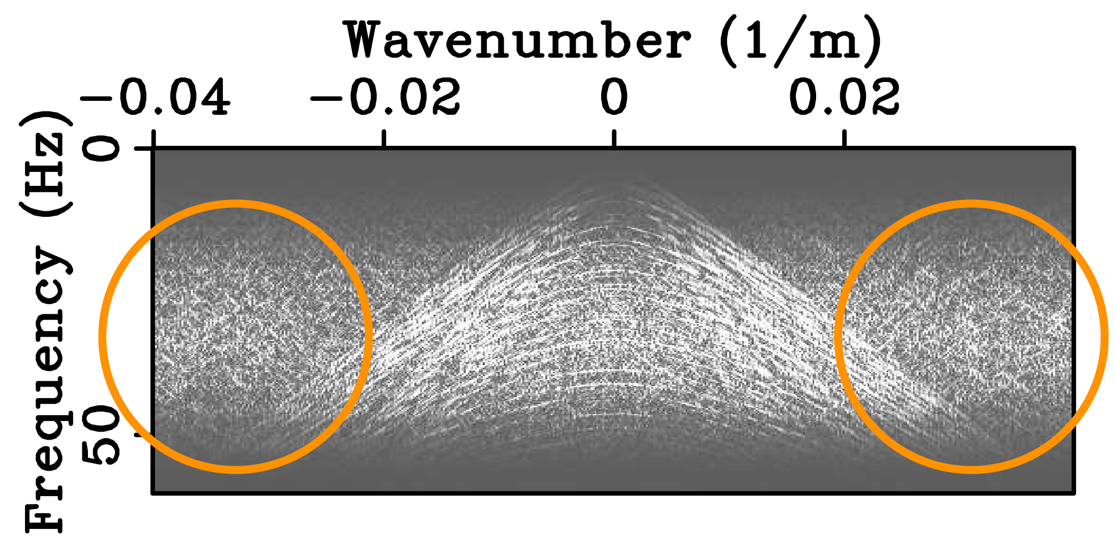
SNR = 6.92 dB



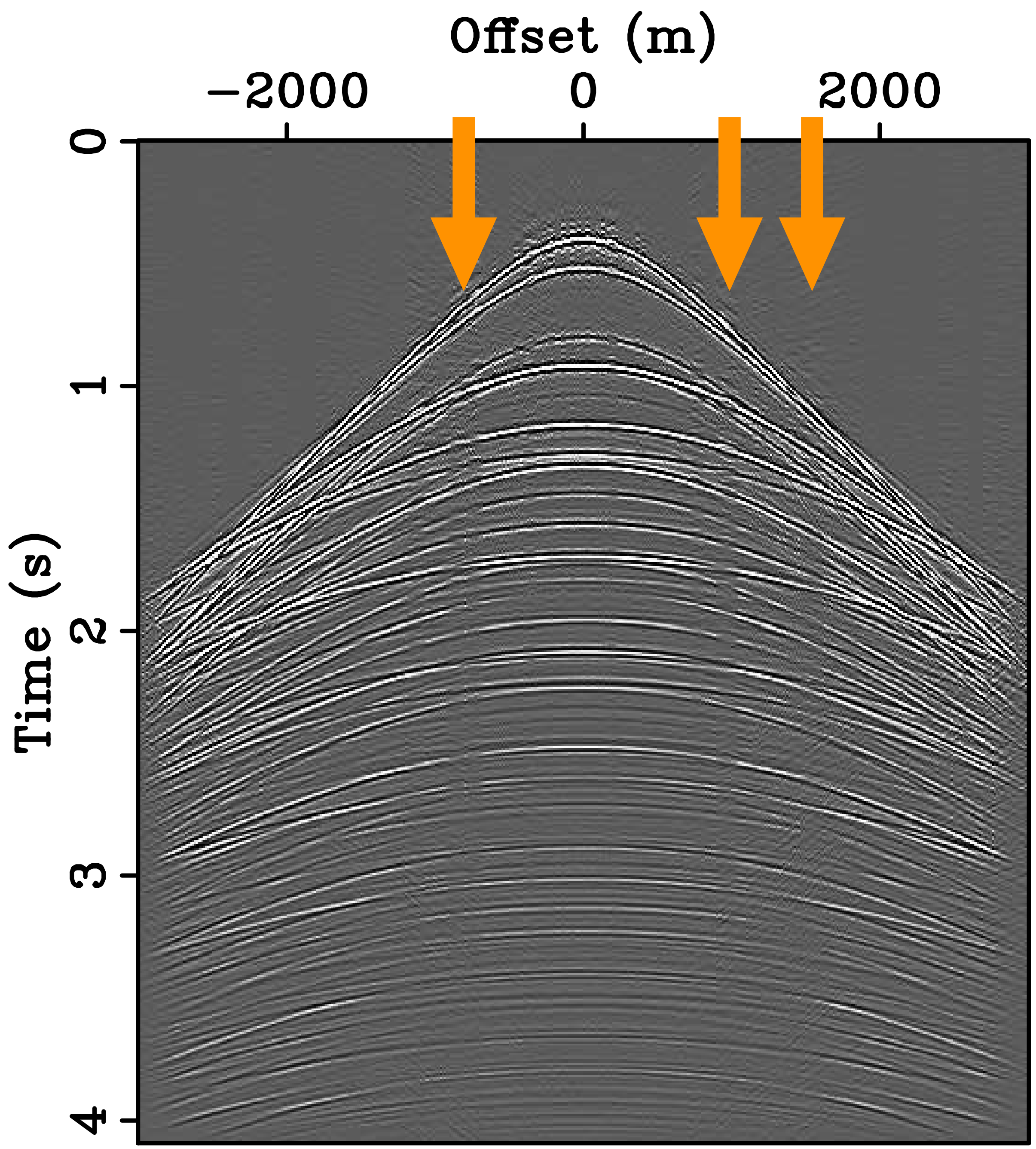
$$\mathrm{SNR} = 20 \times \log_{10}\left(\frac{\|\text{model}\|_2}{\|\text{reconstruction error}\|_2}\right)$$

Random 3-fold undersampling

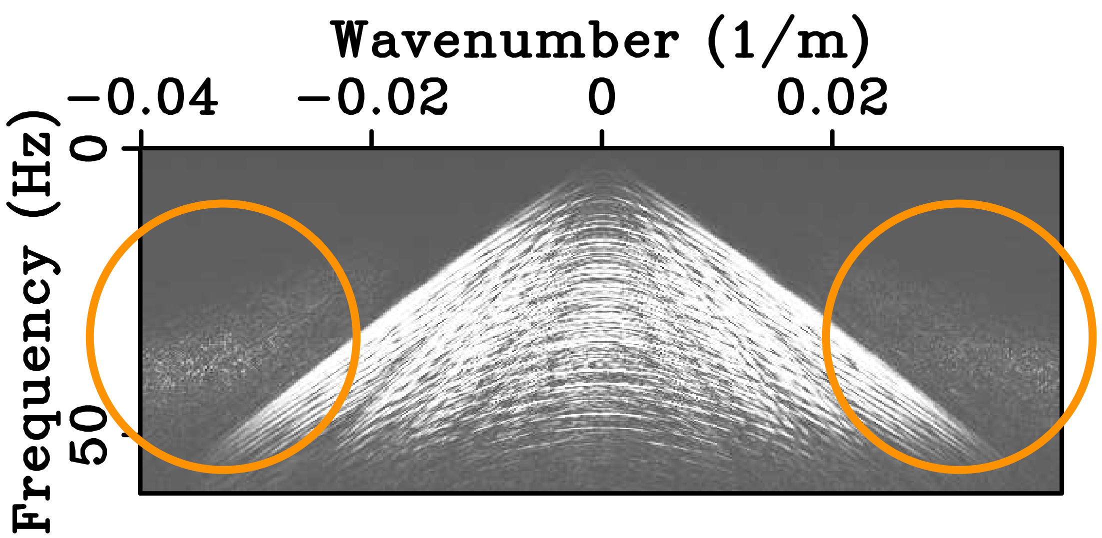




CRSI from random 3-fold undersampling

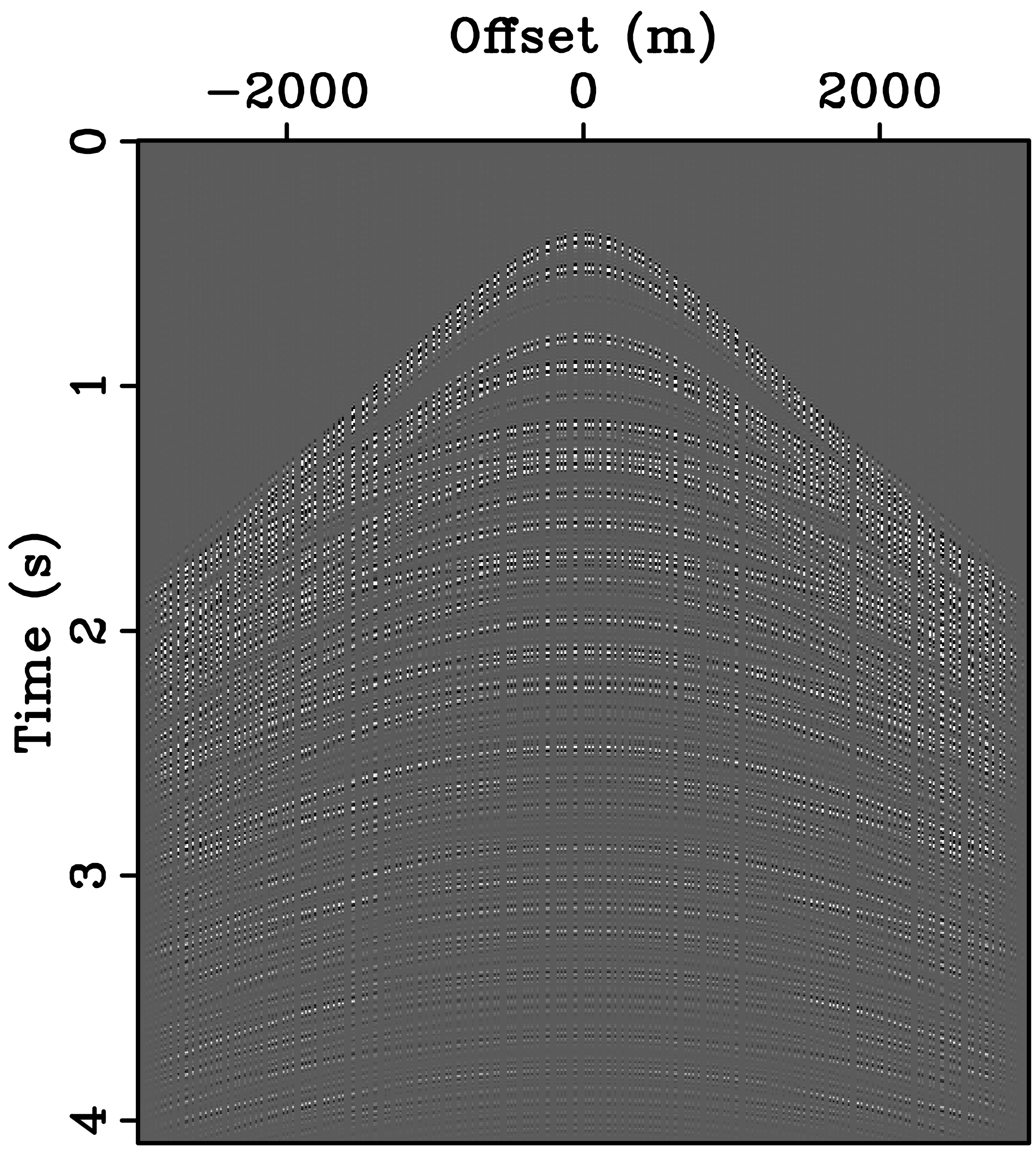


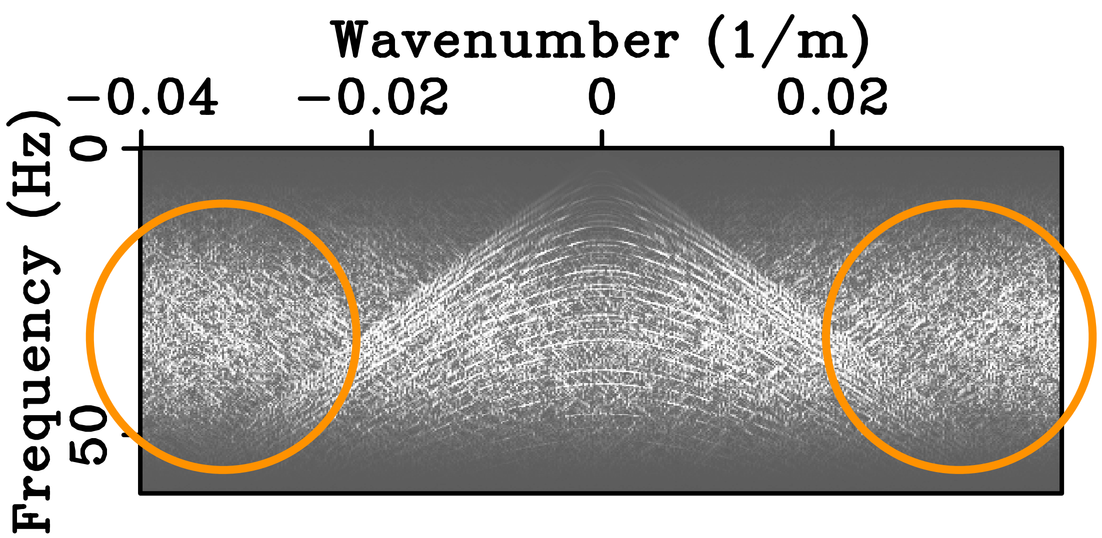
SNR = 9.72 dB



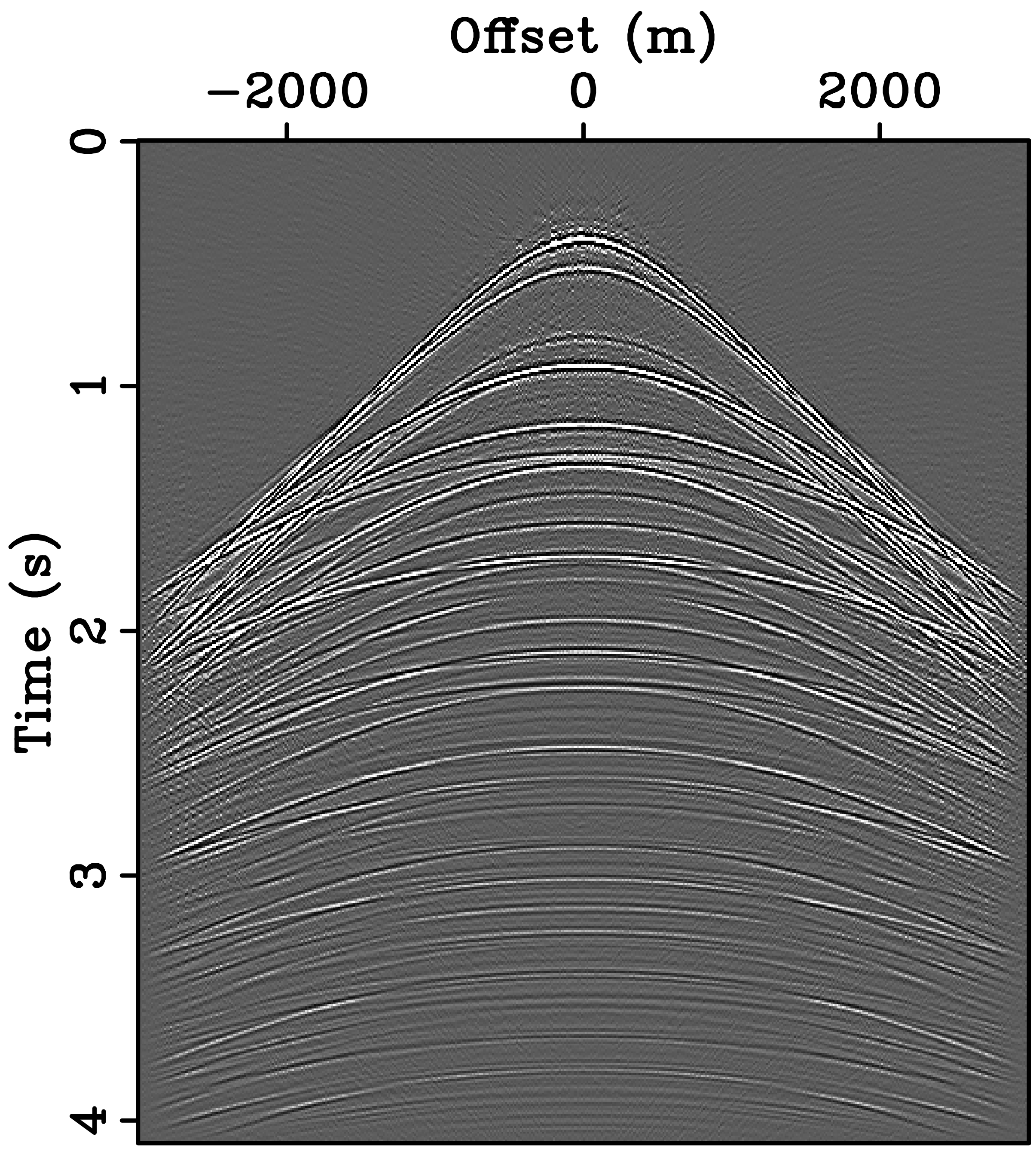
$$\mathrm{SNR} = 20 \times \log_{10}\left(\frac{\|\mathrm{model}\|_2}{\|\mathrm{reconstruction\ error}\|_2}\right)$$

Jittered 3-fold undersampling

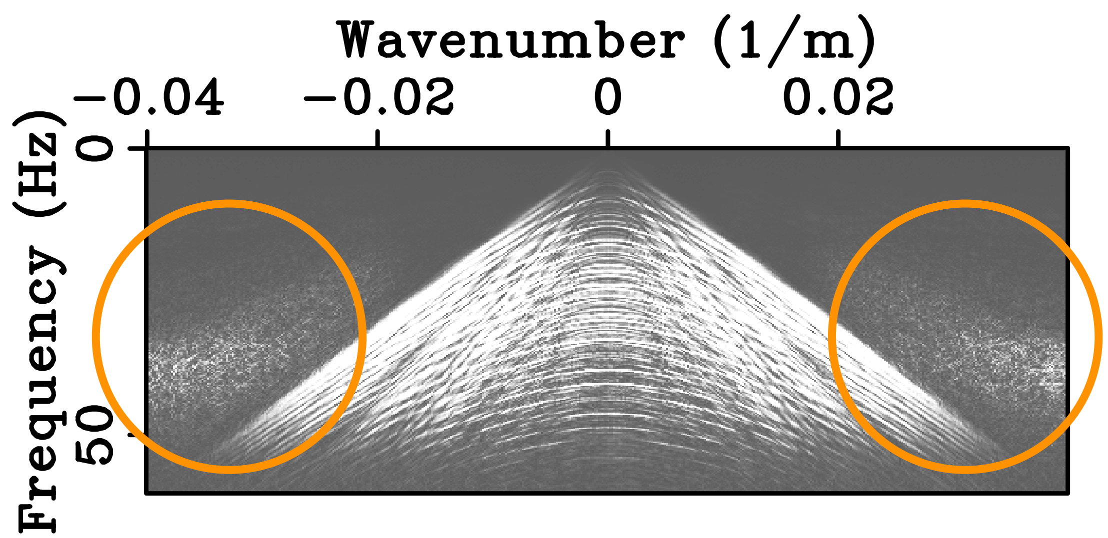




CRSI from jittered 3-fold undersampling



SNR = 10.42 dB



Compressed sensing

How many measurements do we need to make?
Far less than what Shannon tells us.

$$y = Ax$$

x is a vector in $\mathbb{R}^N$
y is a vector in $\mathbb{R}^n$
y is observation
A is a measurement matrix
A has n rows and N columns, where $n \ll M$.

Measurement matrix

$$y = Ax \qquad A \text{ is a matrix}$$

The matrix satisfies the RIP property:
nearly preserve length of sparse vectors.

Fourier matrix is more practical.

Gaussian matrix is easier to analyze .

Random Fourier matrix

- Start with a DFT matrix (N by N)
- Pick n rows at random
- Column entries are not independent
- How to pick the n rows?

Regular sampling:

row 10, 20, 30, 40, 50, 60, ... , 990, 1000

Jittered sampling:

7, 16, 22, 36, 44, 53, 67, ... , 994

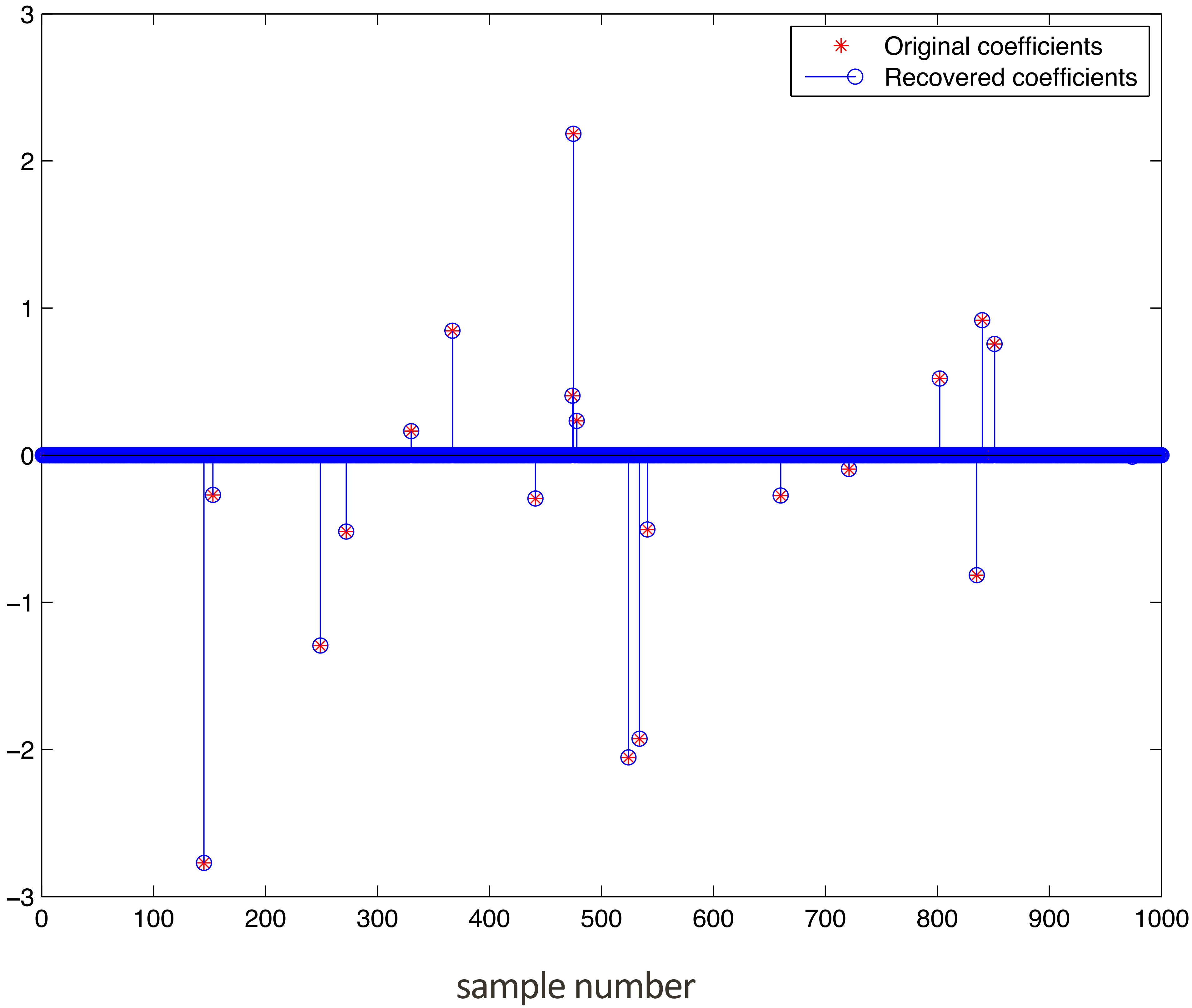
from row 1 to row 10, pick one row

from row 11 to row 20, pick the second row

from row 21 to row 30, pick the third row

Recovery of sparse signal

N = 1000
n = 200
k-sparse
k = 20



SLIM

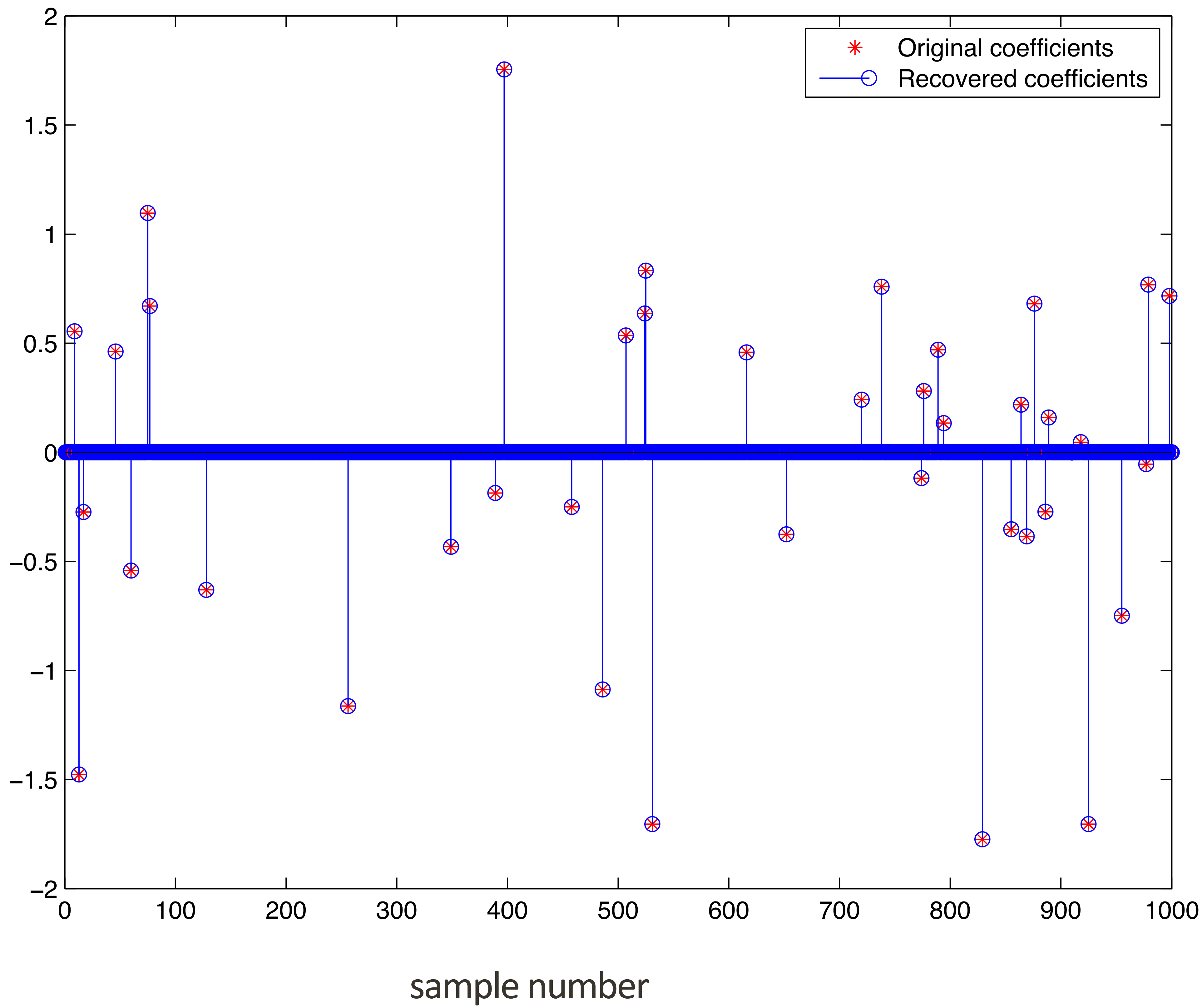
Sparsity

- Compressed sensing can be applied when a signal is sparse.

- What if your signal is not sparse enough?

Recovery of sparse signal

N = 1000
n = 200
k-sparse
k = 40



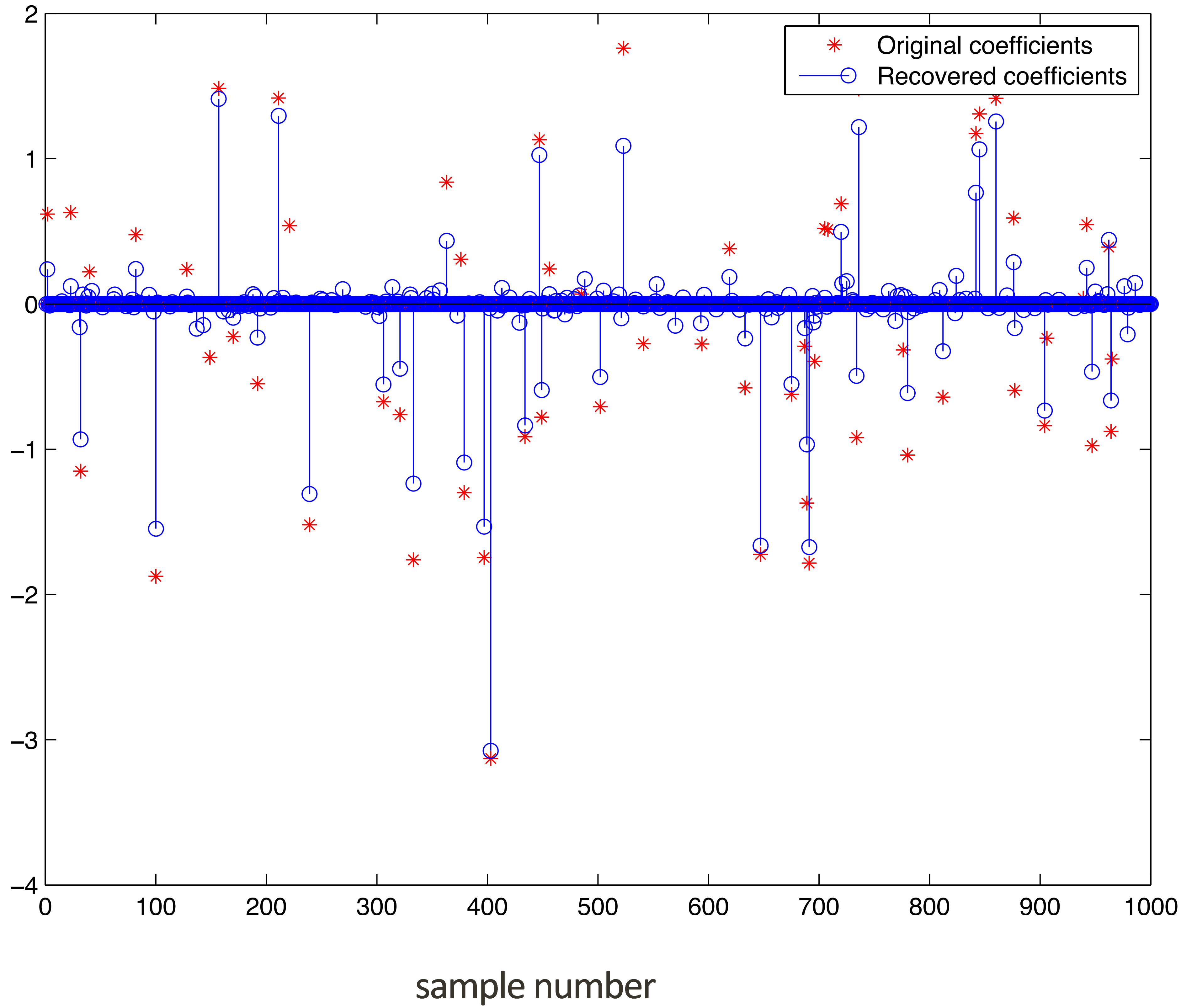
Partial recovery

N = 1000

n = 200

k-sparse

k = 60



When signal is not sparse enough, there can be reconstruction error.

But if most non-zero entries are in the low pass region of the signal, can we recover the signal?

What if most non-zero entries are in low-pass?

Signal in 1000 dimensions is 60-sparse, with 50 entries in the low pass and 10 entries in the high pass.

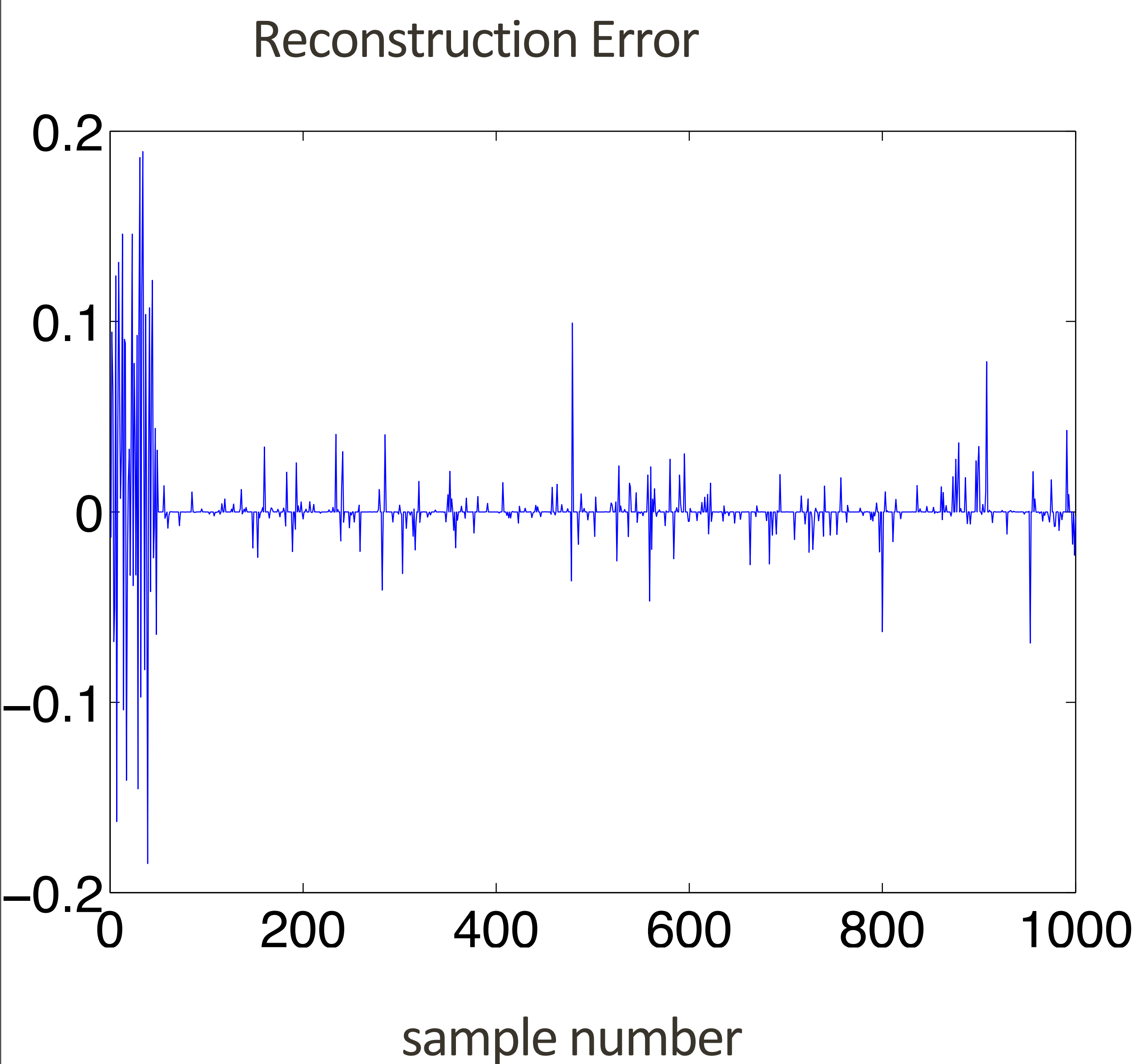
Suppose we make 200 measurements.

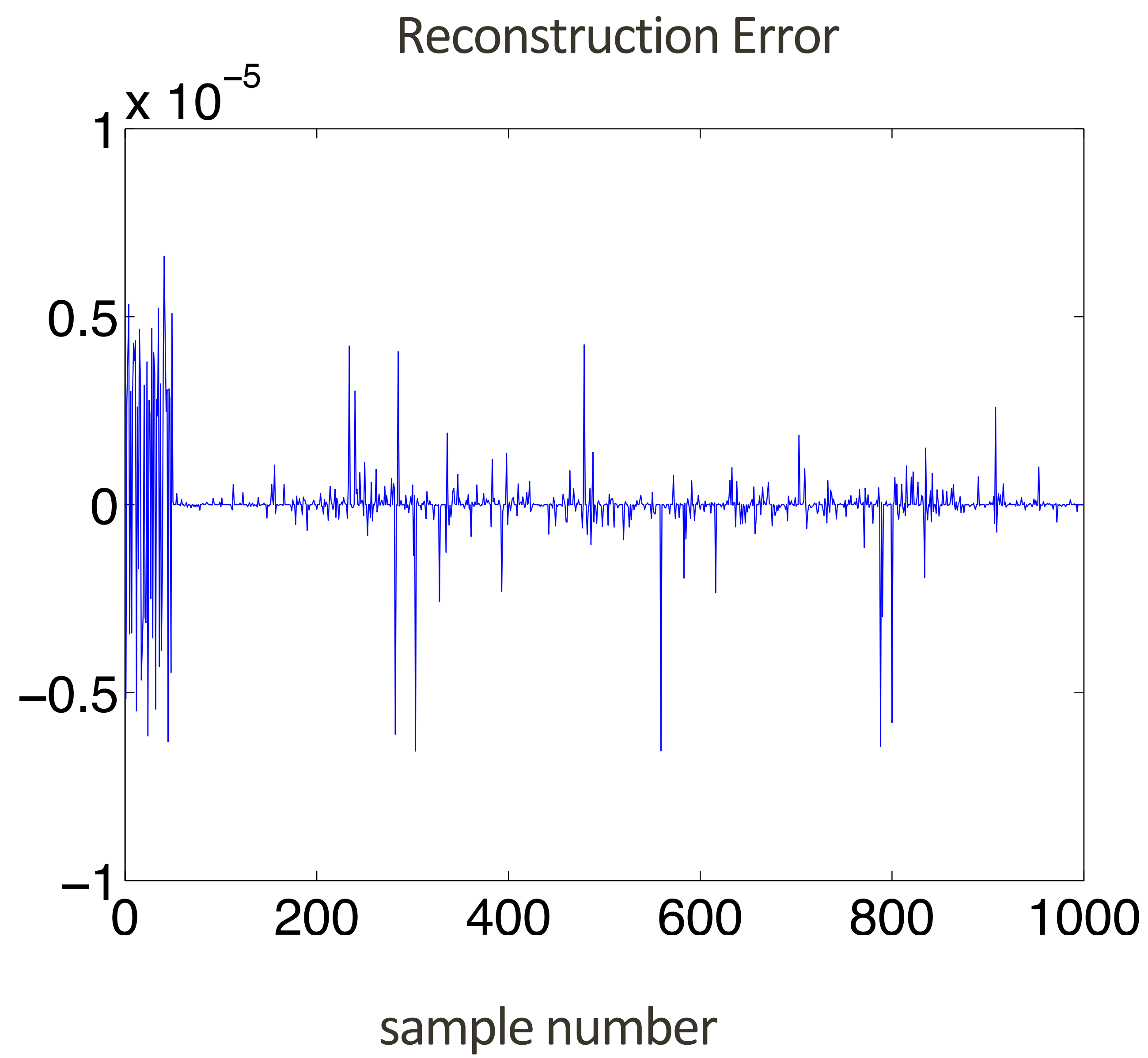
Goal: Compare the reconstruction error in Uniform sampling versus Jittered sampling.

Reconstruction error
= Difference between actual and reconstructed signal

Uniform

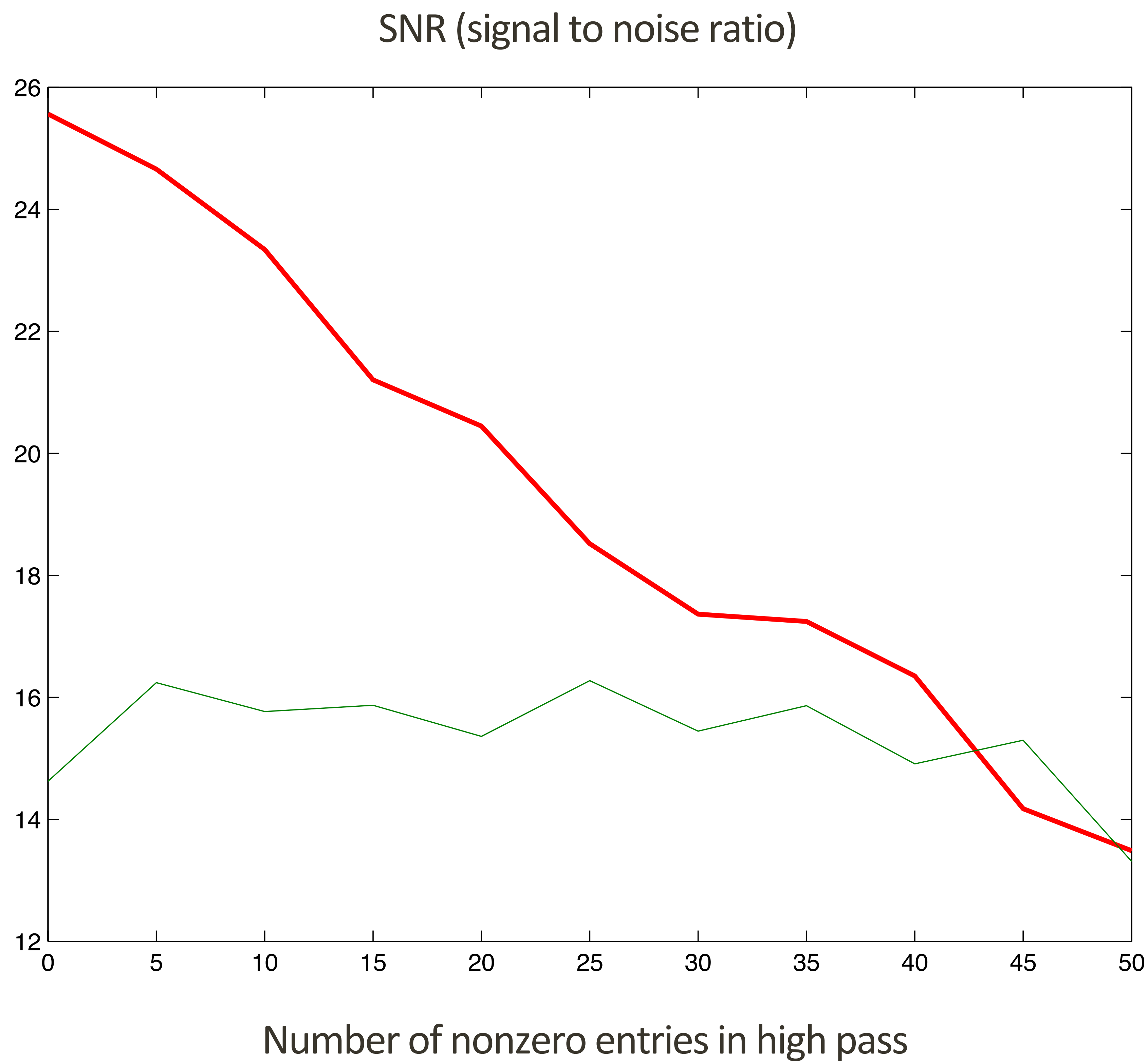
Jittered





Better recovery on lowpass

Jittered and Uniform sampling



N = 1000
n = 200
k-sparse
k = 60

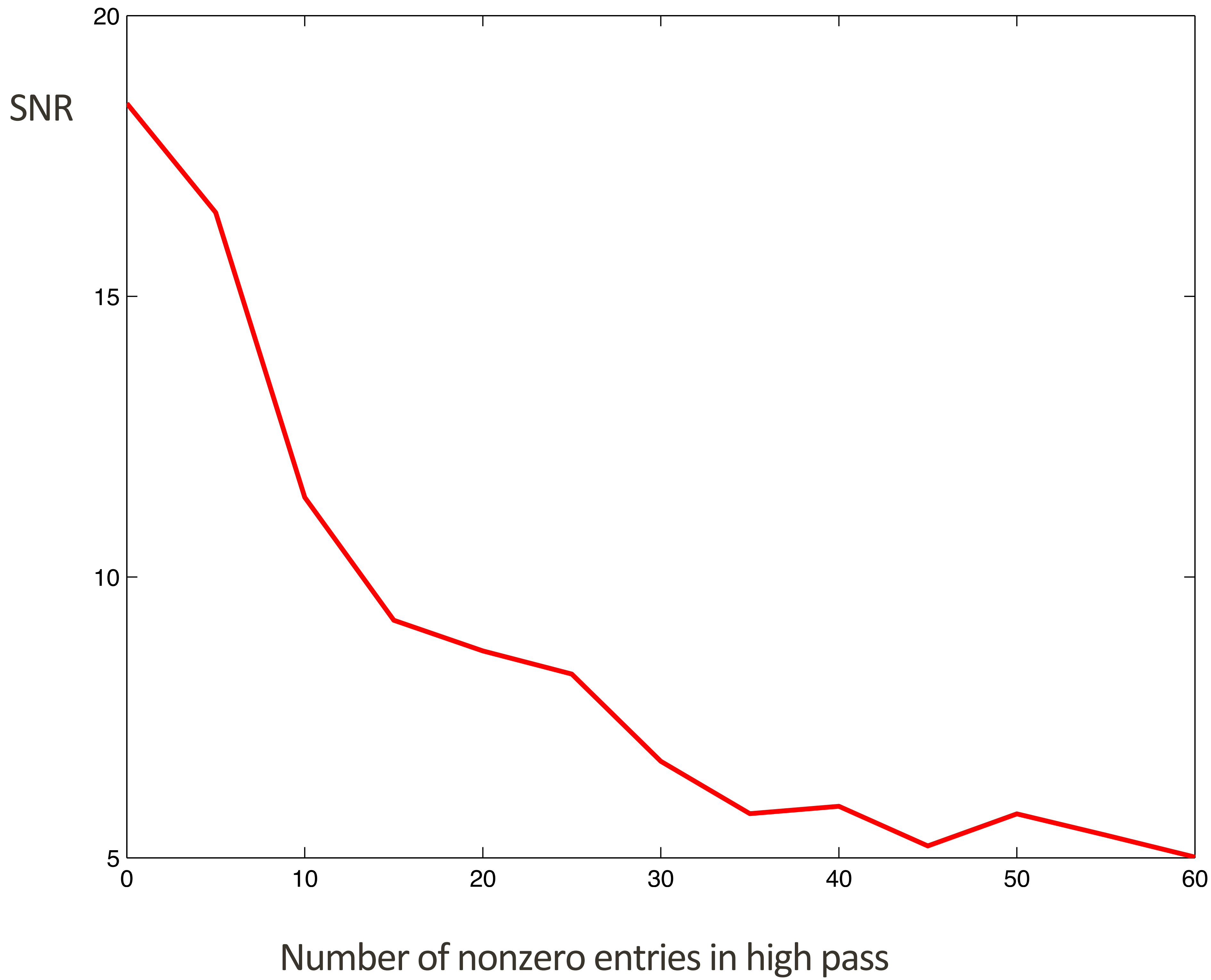
SNR decreases as frequency increases

N = 1000

n = 200

k-sparse,

k = 70



What do you gain?

Uniform: flat error as function of frequency

Jittered: SNR increases as frequency decreases

Moral of the story

What is one main advantage of jittered sampling?

Better reconstruction on low pass and without doing worse on high pass.

SLIM

Potential improvement

- Jittered sampling is better than uniform sampling
- Understand why
- Do educated jittered sampling instead of random jitter

Exploit sparsity structure

Data = sum of wave atoms

Each wave atom has the same shape.

$$X = c_1\psi_1 + c_2\psi_2 + \ldots + c_N\psi_N$$

Suppose coefficients are organized in a tree structure.

Better reconstruction

- What does it mean to reconstruct a signal with high probability?
- Uniform recovery implies non-uniform recovery, but converse is not true

Uniform recovery

Uniform recovery means that once the random matrix is chosen, then with high probability, all k-sparse signals can be recovered.

Non-uniform recovery states only that each k-sparse signal can be recovered with high probability using a random draw of the DFT matrix.

Theory behind the scene

$$B = \{x \in \mathbb{R}^N : \|x\|_2 = 1\}$$

Theorem: Let x, y be any two vectors in B.
Let A be a random projection into k-dimensional subspace.
Then the following holds.

$$Pr\left(| \; \|Ax\|_2 - \|Ay\|_2 \; | > u\sqrt{\frac{k}{n}} \; \|x - y\|_2 \right) \leq \exp(-cu^2 k)$$

Dimensional reduction

What does the theorem say in plain English?

$$N = 1000 \qquad k = 60 \qquad \mathbb{R}^N \to \mathbb{R}^k$$

Points in very high dimension that were near to each other, when you project them to lower dimension, (most likely) there will not be too much distortion.

SLIM

Theorem (on Dimensional reduction) applies to every random projection. It is stronger than the following theorem:

Given any two nearby points x and y, there exists a random projection matrix A such that Ax and Ay will not be distorted too much.

SLIM

Theorem (on Dimensional reduction) applies to every random projection. It is stronger than the following theorem:

Given any two nearby points x and y, there exists a random projection matrix A such that Ax and Ay will not be distorted too much.

This A can be taken to be a Gaussian matrix.
[Concentration of Measure argument does not apply to random Fourier matrix.]

Conclusion

Jittered sampling leads to better reconstruction on low pass, without doing worse on high pass.

Thank You!

Acknowledgement

- Ozgur Yilmaz
- Felix Herrmann
- Haneet Wason, Hassan Mansour
- All siblings from the SLIM family

This work was in part financially supported by a PIMS fellowship from the Pacific Institute for Mathematical Sciences and by a NSERC Accelerator Discovery award (Yilmaz - AID 411944-2011). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, Chevron, ConocoPhilips, Petrobras, PGS, Total SA, and WesternGeco.