Australian
National
University

New Applications of Sparse Methods in Physics

Ra Inta,

Centre for Gravitational Physics,

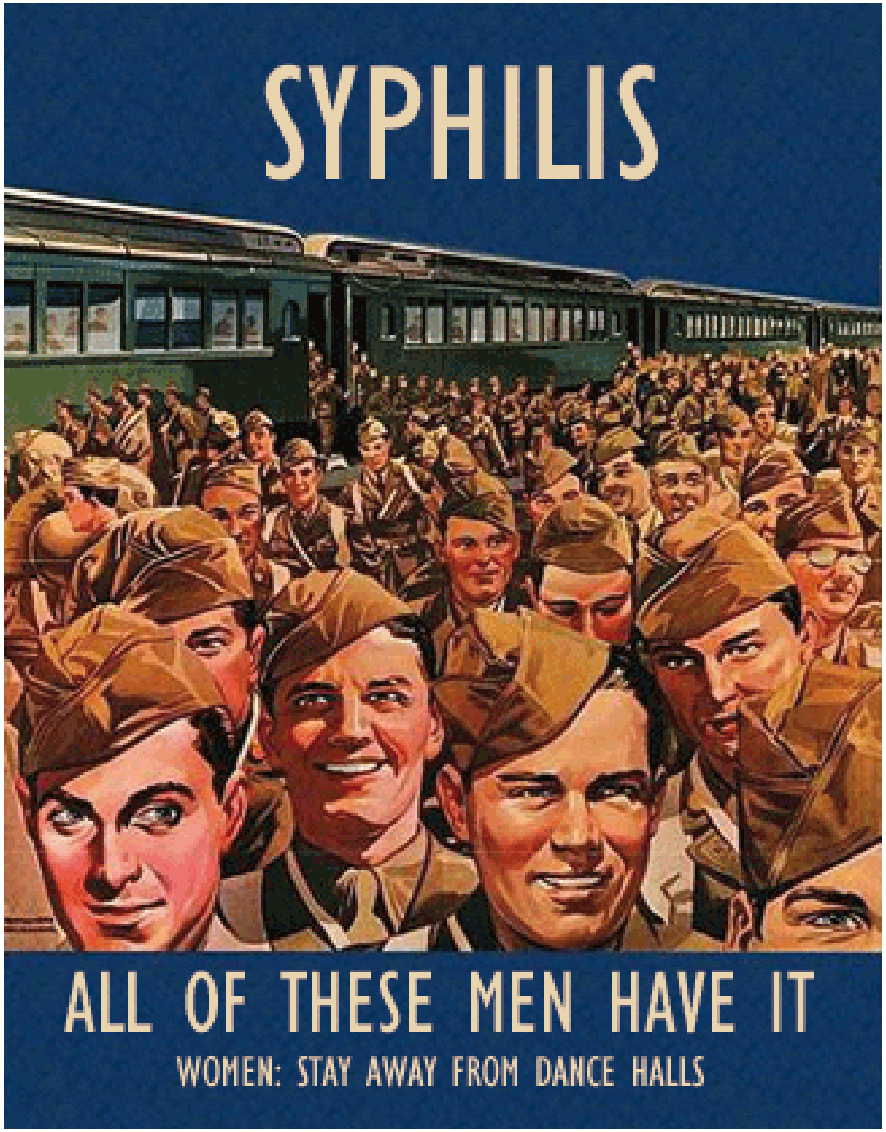
The Australian National University

Sparse methods

A vector is **S-sparse** if it has **at most S non-zero coefficients**. ('Compressible' if you can ignore many components)

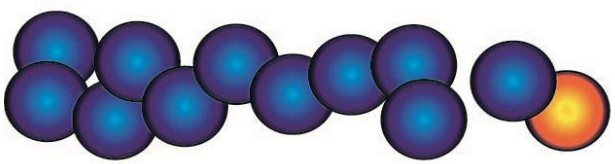
Wide application:

- Syphilis in WWII
- Image/file compression

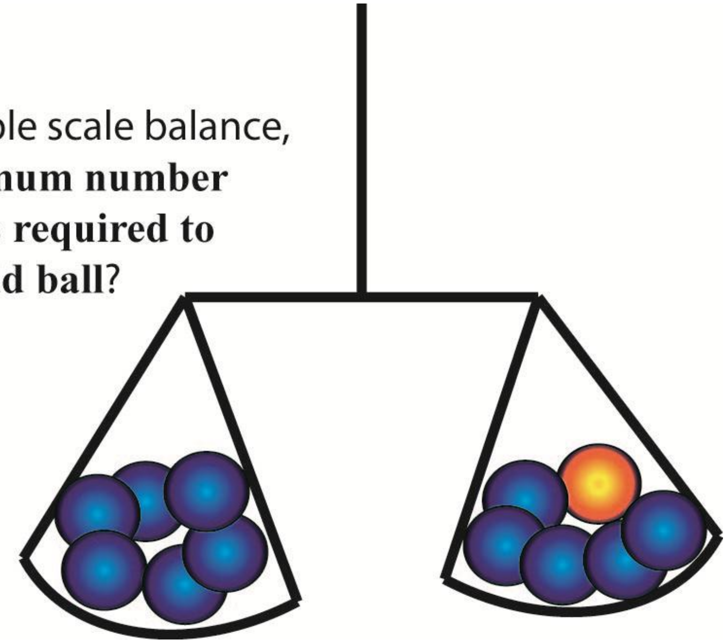


The twelve ball problem

You're given 12 identical balls, except exactly **one is heavier/lighter than the others.**



Using only a simple scale balance, what is the **minimum number of measurements required to determine the odd ball?**



Here, the ball mass representation has a sparsity of S=1

Success stories

- JPEG-2000 format. Only need about 10% coefficients
- The internet would be a beautiful but empty landscape without sparse-based compression

Modern sparse methods

- Advances in image processing
- Funny andedcdote (Nyquist, Nyquist Nyquist)
- Terry Tao (2006)

Compressive sampling (CS)

- Also known as ‘compressed sensing’
- Main result: quantitative circumvention of Shannon-Nyquist sampling theorem

Compressive sampling in a nutshell

- Unified framework for encoding/reconstruction of sparse signals
- Applications from radar, Herschel space observatory, to a single pixel camera

E. Candès, J. Romberg and T. Tao: *IEEE Trans. on Information Theory* **52**(2):489–509 (2006)

Image: M. F. Duarte et al., Rice University

Linear measurement system

$$y = \Theta \, x$$

Measurement vector (1 X M)

Original signal (1 X N)

$$\Theta = \Phi \, \psi$$

Sparse basis (N X N)

Sensing matrix (M X N)

Determined system: M = N

CS (‘under-determined’): M << N

Solve for x by explicitly imposing sparsity constraint:

$$\min \|x\|_0 \text{ subject to } \|y - \Theta x\|_2 = 0$$

($\|x\|_p = \sqrt[p]{\sum x_i^p}$ is the p-norm)

But it's a combinatorial problem!
(guess ${}^{N}C_{M}$ combinations for large N, moderate M)

CS: Almost as good (especially for large N):

$$\min \|x\|_1 \text{ subject to } \|y - \Theta x\|_2 = 0$$

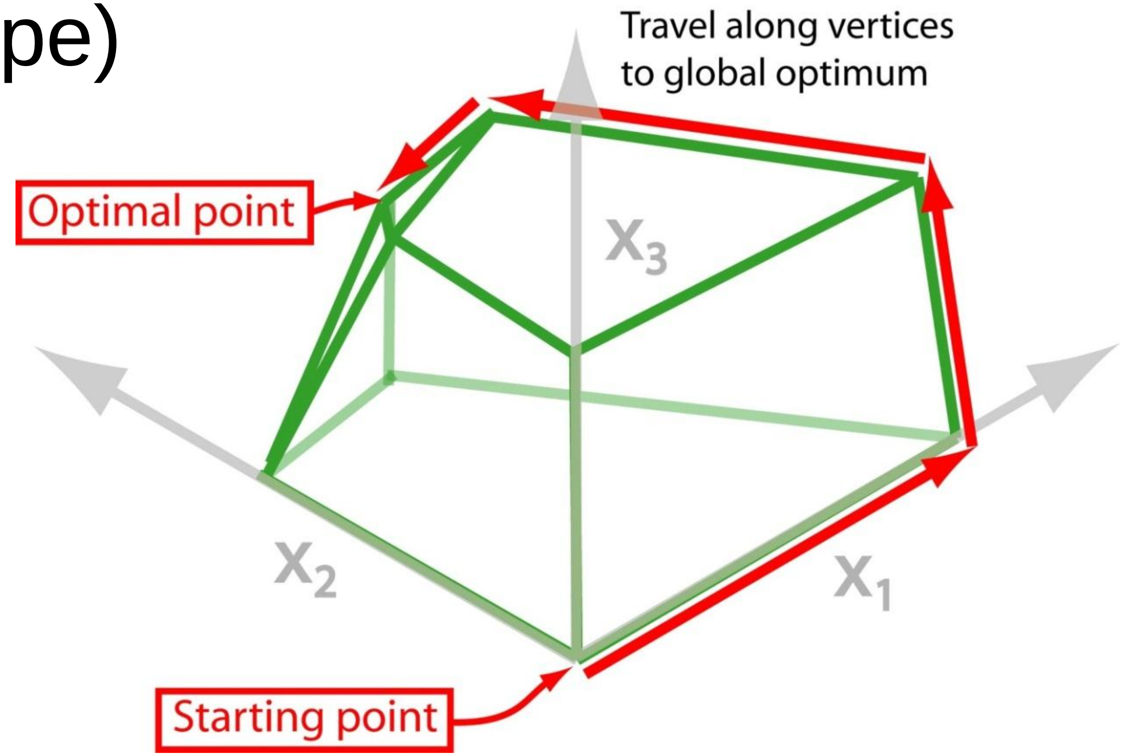
Problems with noise? Relax!

$$\min \|x\|_1 \text{ subject to } \|y - \Theta x\|_2 \leq \epsilon$$

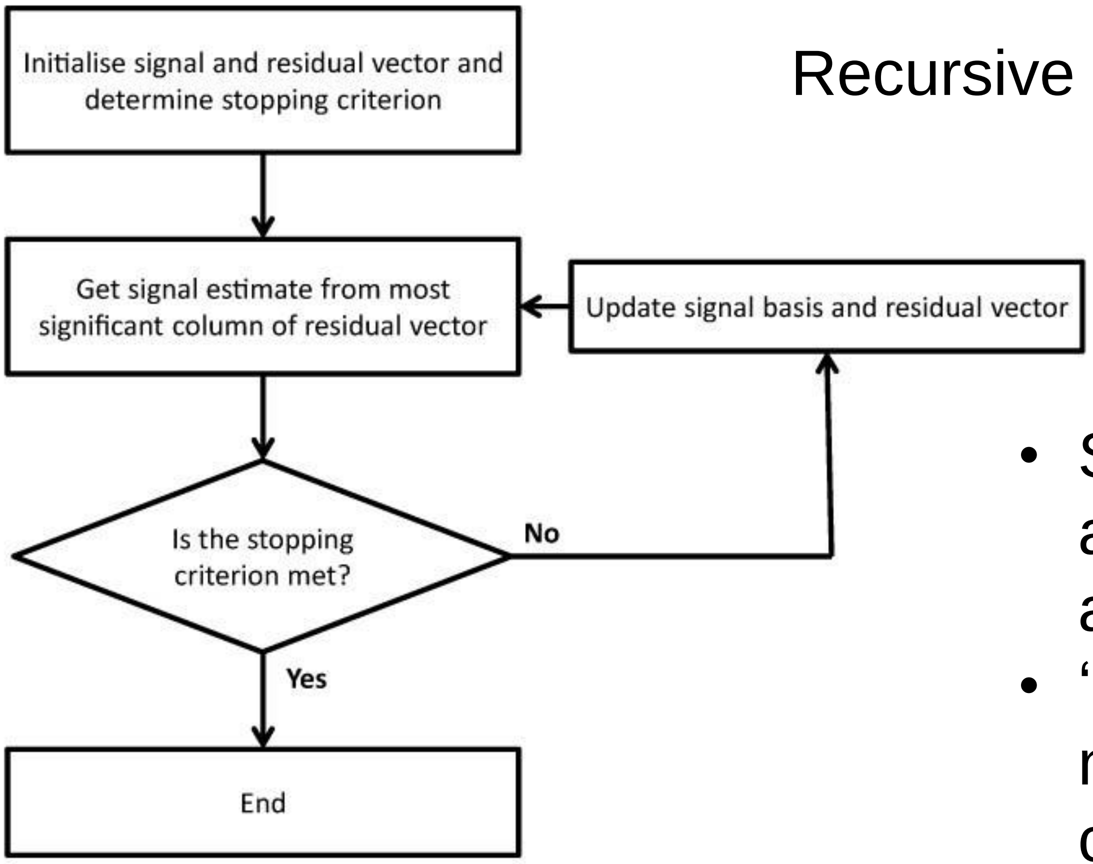
For some total noise/compressibility residual ϵ

Linear program

- Optimisation reconstruction methods
- E.g. Simplex method (linear constraints give convex polytope)



Orthogonal matching pursuit (OMP)



Recursive greedy algorithm

- Similar to CLEAN algorithm (radio astronomy)
- ‘Easy’ to implement noise-based stopping criteria

Headline CS results

Can determine minimum number of samples required for perfect reconstruction:

$$M > C\,S\,\mu^2\ \log(N)$$

Where *C* is small (~0.5) , and the mutual coherence:

$$\mu \stackrel{\text{def}}{=} \sqrt{N} \max_{i,j<N} | < \varphi_i, \psi_j > |$$

Measures how ‘spread out’ the signal is in the sampling domain

Australian National University

Mutual coherence

Guarantees sampled low dimensional subspace sufficiently covers sparse basis (Restricted Isometry)

e.g. Delta functions in Fourier domain are minimally coherent with time domain:

$$\mu = \sqrt{N} \max_k \int \delta(f - f_k) e^{i2\pi f t} dt = \sqrt{N} \frac{1}{\sqrt{N}} = 1$$

E. Candès and J. Romberg, *Inverse Problems* **23**:969–985 (2007)

CS applications

Basic ingredients:

1. Sparse/compressible representation
2. Low mutual coherence between sensing and sparsity bases

Sparse fast Fourier transform (sFFT)

- Implements non-recursive OMP
- Computational complexity for known sparsity S:
  $O(S \log(N))$

Compare to FFT [$O(N \log(N))$]:
Speed-up is $O(N/S)$, so e.g., for $N=10^6$, $S=100$:

Here **sFFT** is (theoretically) **10,000 times as fast as FFT!**

Hassanieh, H., Indyk, P., Katabi, D., and Price, E.: "Nearly Optimal Sparse Fourier Transform," *arXiv* **1201.2501v1** (12 Jan 2012)

How the sFFT works

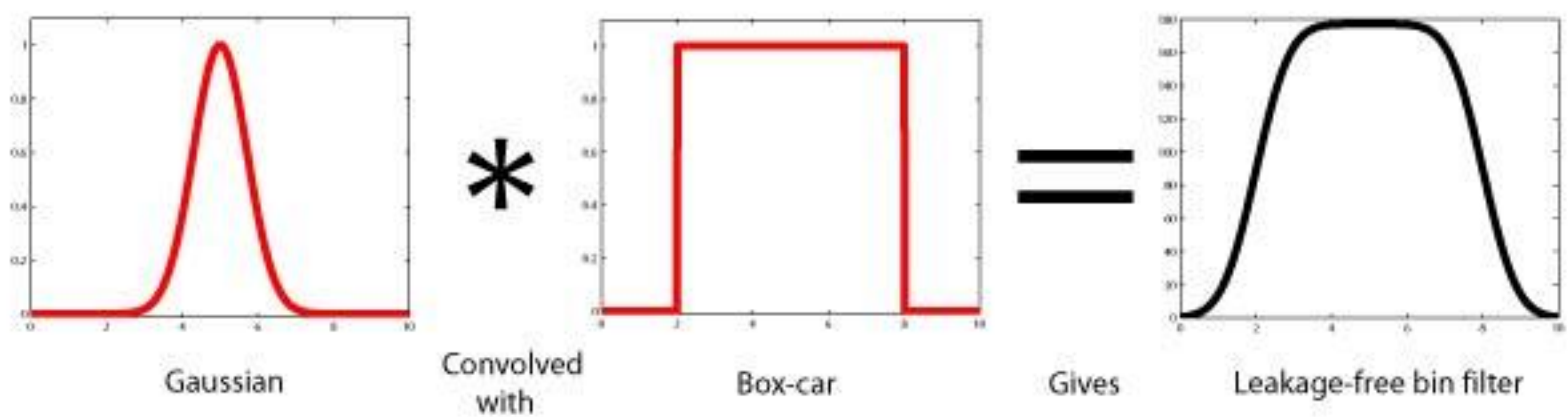
Recall:

FFT: output proportional to N: O(N log(N))

sFFT: output proportional to S: O(S log(N))

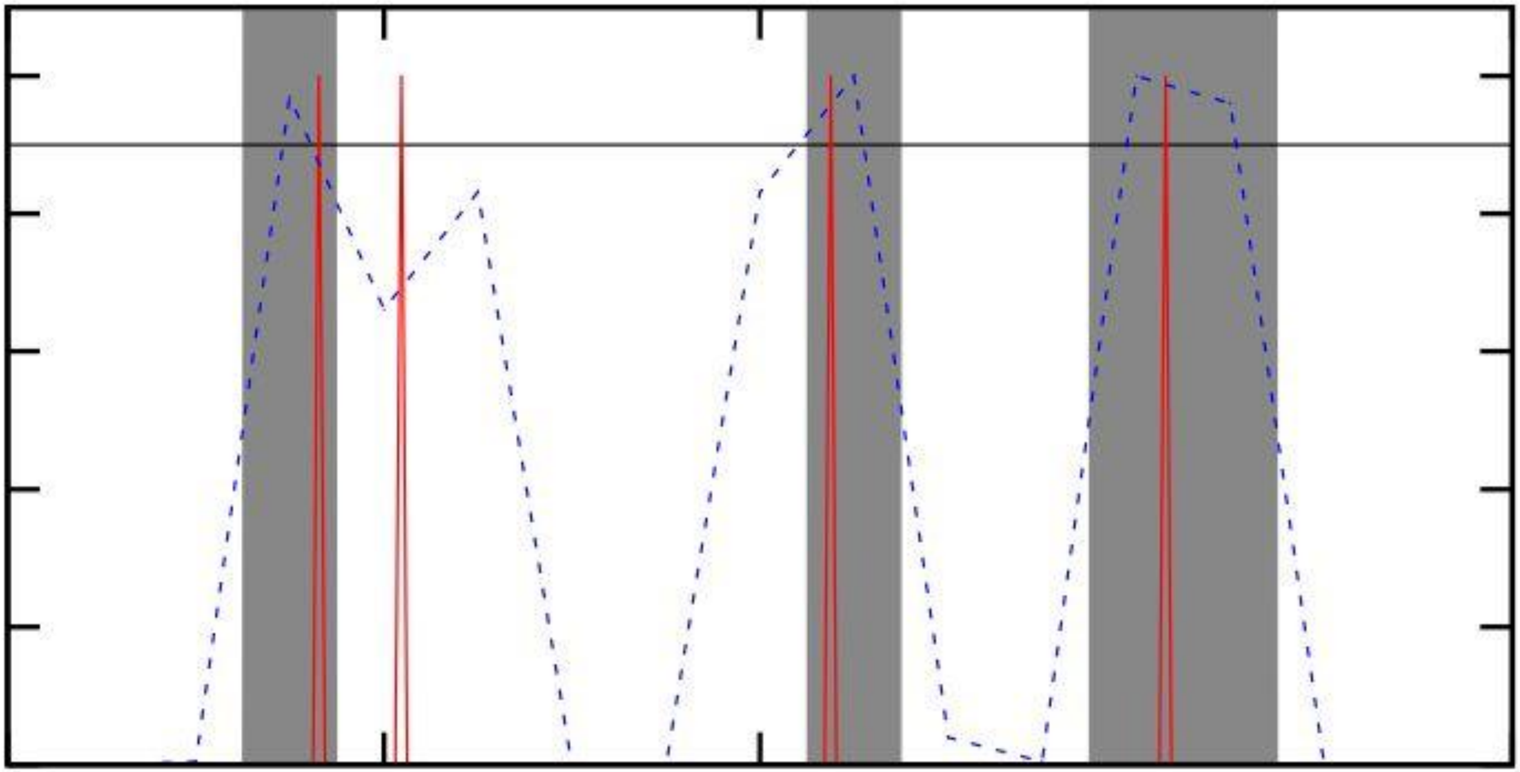
sFFT: very wide bins, permuted so each only contain (at most) single large coefficient

Leakage-free bin filter:



- Very wide, tuned, bin size (B ~ $\sqrt{(NS)}$ initially, decreasing as coefficients identified)
- Leakage-free bin filters
- Use block recovery (i.e. not individual coefficients)
- Iteratively remove identified coefficients from *bins*, not the *signal*

Threshold and identify non-zero bins



Code available from http://groups.csail.mit.edu/netmit/sFFT/code.html

Application 1: localisation of signals in the time-frequency plane

- Gravitational waves (GWs) produced by binary compact objects (neutron stars and black holes)
- Sparse in ‘chirp’ domain
- Position information from GW signals depends on timing
- Nyquist limited

Detection pipelines

- Omega, continuous wave-burst
- Look for excess energy in Gabor (time-frequency) plane
- Omega: Sine-Gaussian wavelets:

$$\psi(\tau) = A\ \exp\left(\frac{-(2\pi f)^2}{Q^2}(\tau - t)^2\right)\ \exp(2\pi i[f(\tau - t)])$$

$$A = \left(\frac{8\pi f^2}{Q^2}\right)^{1/4}$$

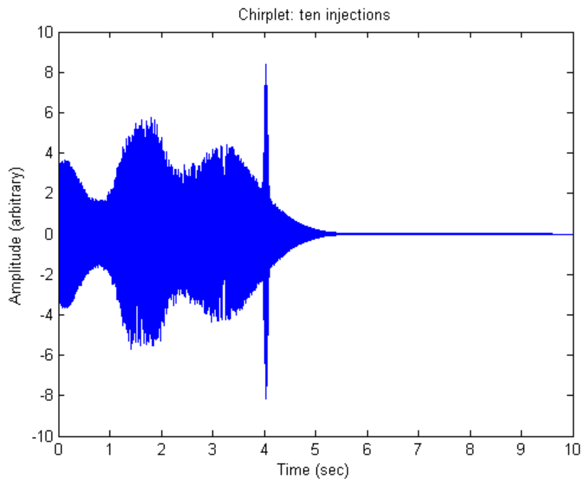
An improvement: chirplets

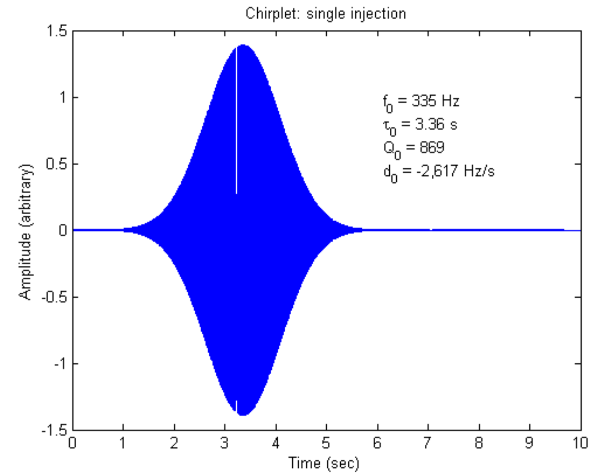
- Chirplet approach: add chirp rate parameter, d:

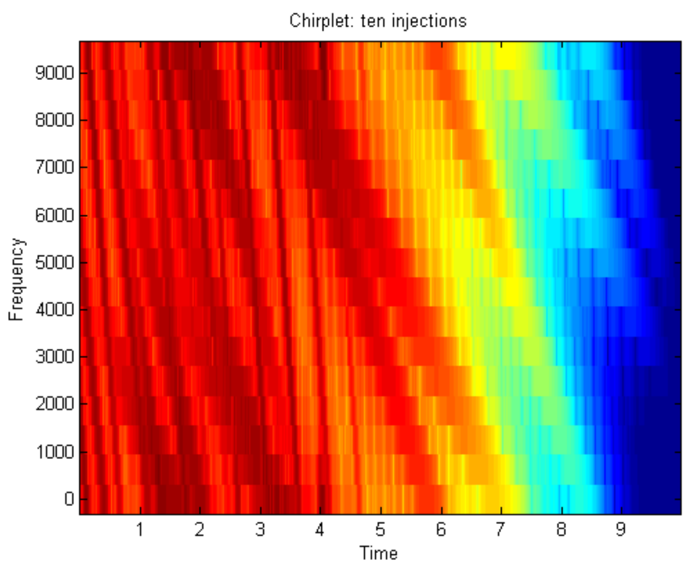
$$\psi(\tau) = A\ \exp\left(\frac{-(2\pi f)^2}{Q^2}(\tau - t)^2\right)\ \exp\left(2\pi i\left[f(\tau - t) + \frac{d}{2}(\tau - t)^2\right]\right)$$

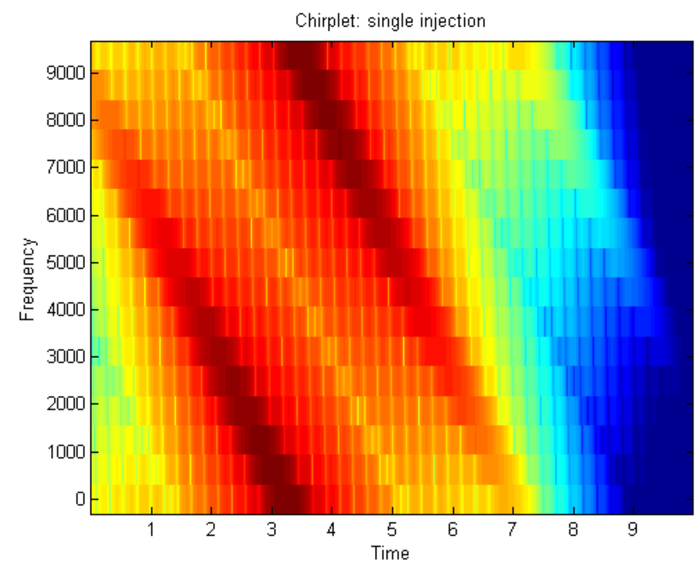
- Covers 10 times more parameter space than Omega
- SNR enhancement of 45% (range increase of ~40%)
- High SNR signal is sparse in the Chirplet domain

É Chassande-Mottin *et al.*, *CQG* **27**:194017 (2010)


Chirplet: ten injections
Amplitude (arbitrary)
Time (sec)
10
8
6
4
2
0
-2
-4
-6
-8
-10
0 1 2 3 4 5 6 7 8 9 10

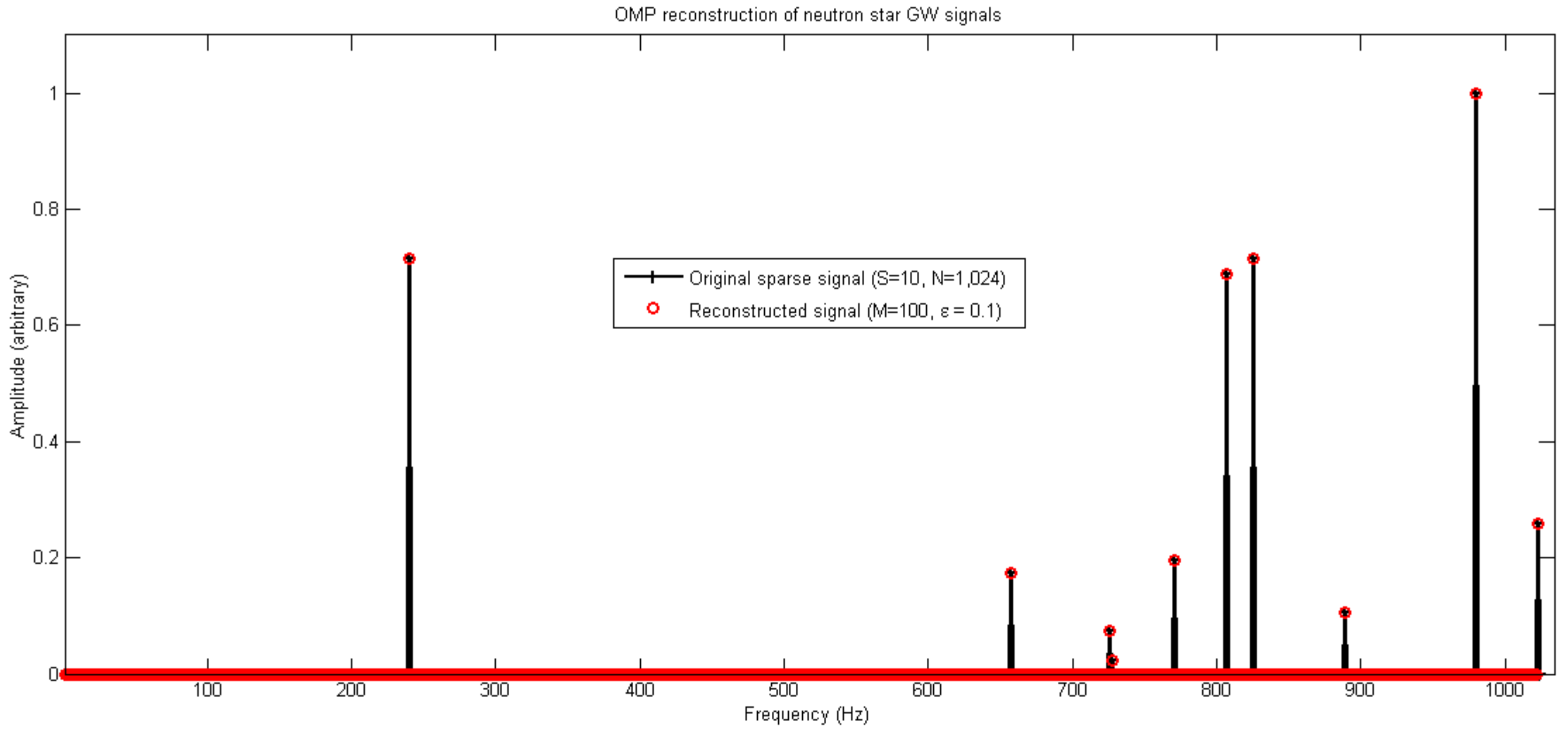

Chirplet: single injection
Amplitude (arbitrary)
Time (sec)
f_0 = 335 Hz
τ_0 = 3.36 s
Q_0 = 869
d_0 = -2,617 Hz/s
1.5
1
0.5
0
-0.5
-1
-1.5
0 1 2 3 4 5 6 7 8 9 10


Chirplet: ten injections
Frequency
Time
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
1 2 3 4 5 6 7 8 9


Chirplet: single injection
Frequency
Time
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
1 2 3 4 5 6 7 8 9

Nyquist limitations

- Nyquist limit in time-frequency plane (aliasing of chirplet templates), so reduced timing precision
- Nyquist limit: equivalent to saying that for a linear system with N unknown coefficients, require N equations to determine system.



OMP: Perfect reconstruction at 10% of Nyquist frequency

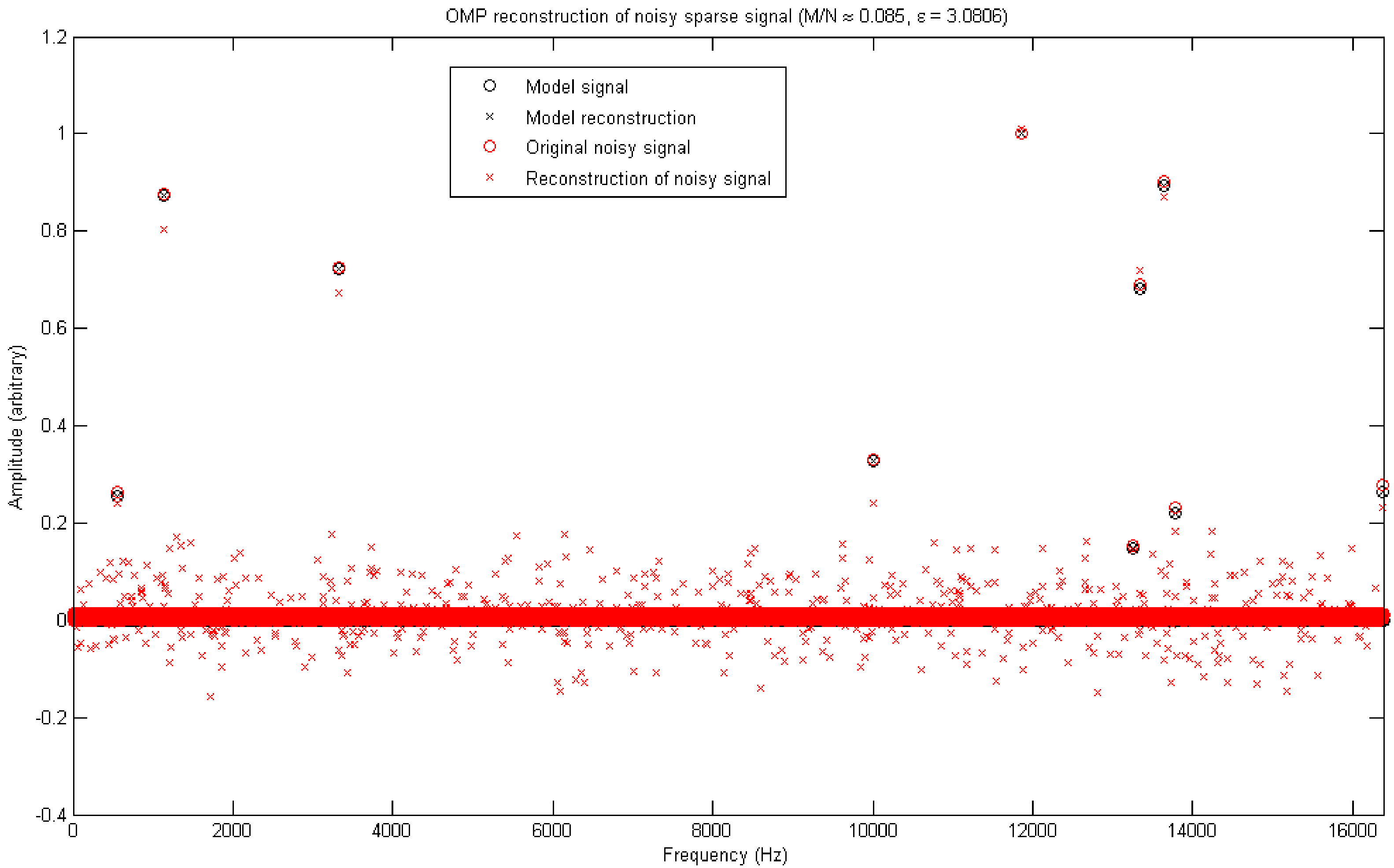
How to deal with noise

Assuming Gaussian noise, Chi-squared with 4 d.f.

$$y = \phi x + n \qquad \sum_{k=1}^{N} x^2 \sim \chi_4^2(\lambda)$$

$$\chi_4^2(x;\ \phi, y) = \sum_{k=1}^{4} \frac{n_k^2}{\sigma_k^2} \leq \epsilon^2$$

noting that p($\chi_4^2 \leq 9.49) = 0.95$, can get 95% C.I. from ϵ^2 = 9.49



OMP reconstruction of noisy sparse signal (M/N ≈ 0.085, ε = 3.0806)
Model signal
Model reconstruction
Original noisy signal
Reconstruction of noisy signal
Amplitude (arbitrary)
Frequency (Hz)
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
2000
4000
6000
8000
10000
12000
14000
16000

Issues

- Pure delta functions in time give maximum μ… Can't use CS
- LIGO/Virgo GW detectors most sensitive ~O(100) Hz---confined to compact region of Gabor plane
- Increased SNR---improve localisation of transient events in Gabor plane

localisation of signals in the time-frequency plane

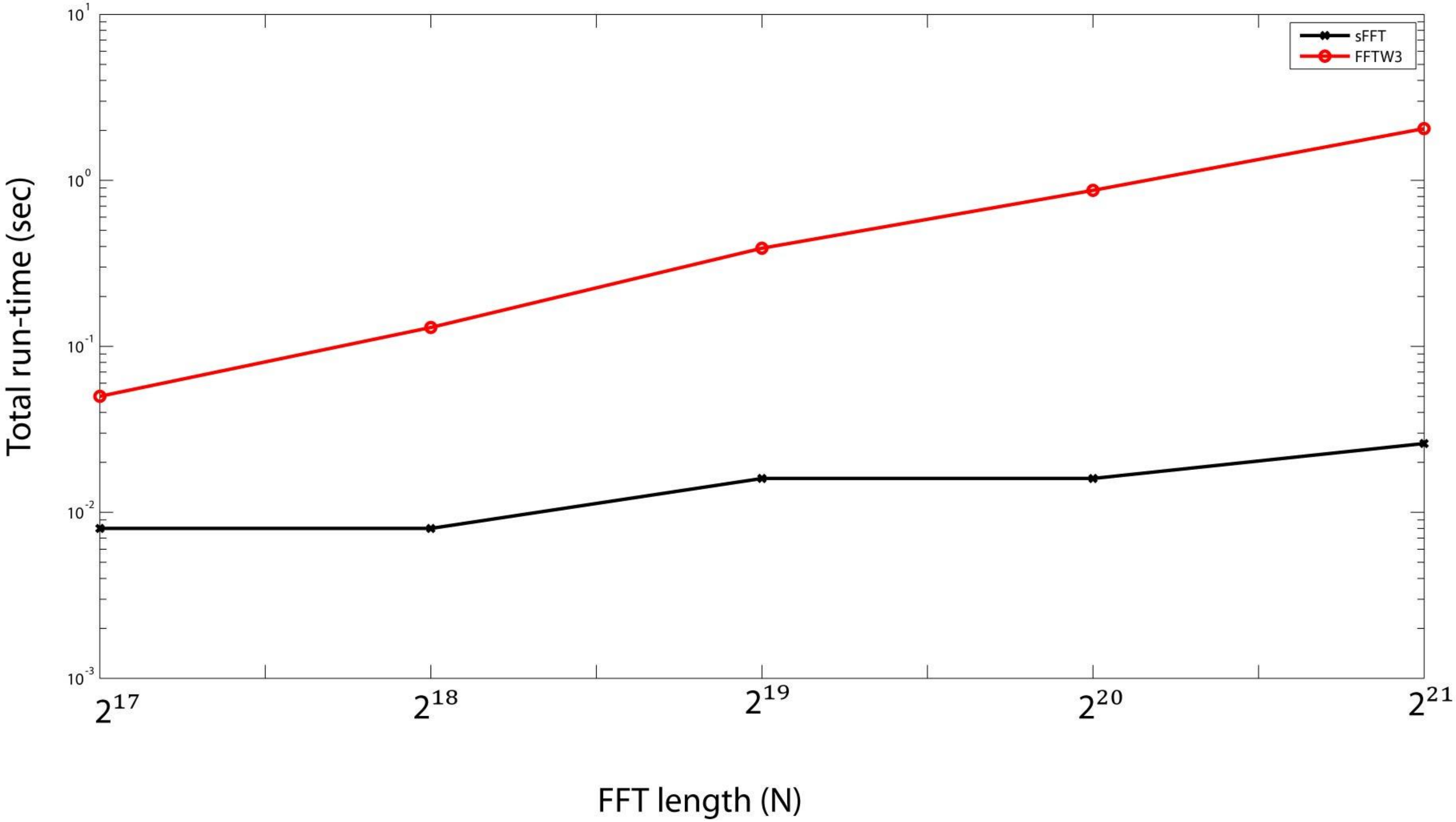
- Superior reconstruction of signals over Wigner-Ville
- Applies very generally to signals in the time-frequency plane that are not purely impulses in time
- Noise is an issue

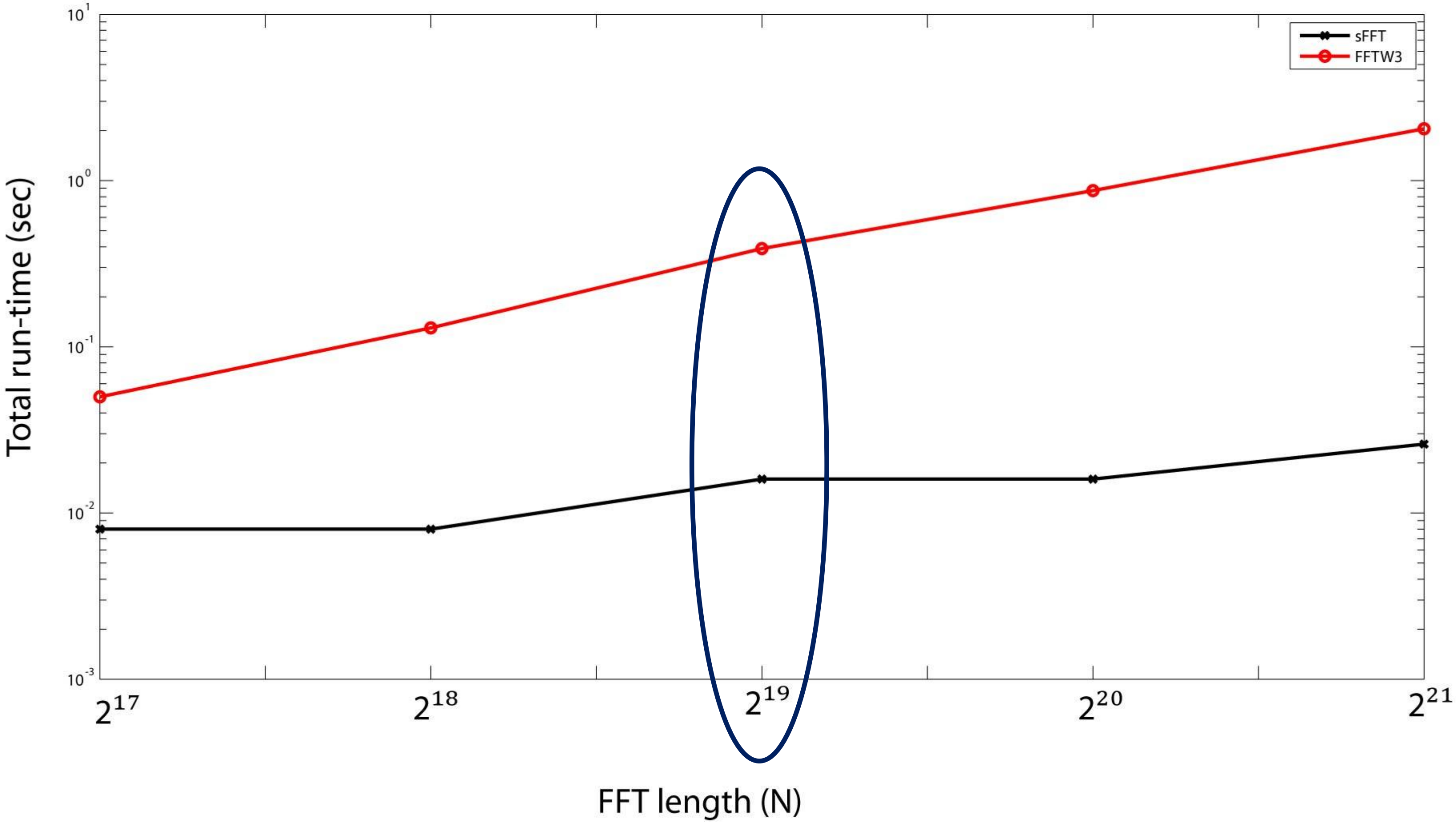
Application 2: Reducing computational bounds with the sFFT

- Many tasks relying on spectral methods searching for signal sparse in Fourier domain (metrology, pulsar discovery, radar, gravitational waves, imaging)
- Interesting problems have a serious computational bound

Scenario: continuous gravitational waves

- Also applies to pulsar discovery


sFFT
FFTW3
Total run-time (sec)
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
2^{17}
2^{18}
2^{19}
2^{20}
2^{21}
FFT length (N)


sFFT
FFTW3
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
Total run-time (sec)
2^{17}
2^{18}
2^{19}
2^{20}
2^{21}
FFT length (N)

Performance

- For a single 30 minute data stretch, searching between 100-300 Hz, gives N=720,000 (> 2^19)
- Randomised trails, for S=20, give run-times:

  sFFT: 14.8 ms

  FFTW3: 40.2 ms

  → Speed-up of over 250% !!!

- For $N \sim 10^6$, this becomes > 467%

Reducing computational bounds with the sFFT

Potentially make many (~ 1 dozen) currently borderline targets feasible; many more for aLIGO

E.g.

- G350.1-0.3
- DA 495 (G65.7+1.2)
- Some estimates for Vela Jr. (G266.2-1.2)

Future work

- Extend OMP code to 4D
- Produce a receiver operating characteristic (ROC) curve for the sFFT

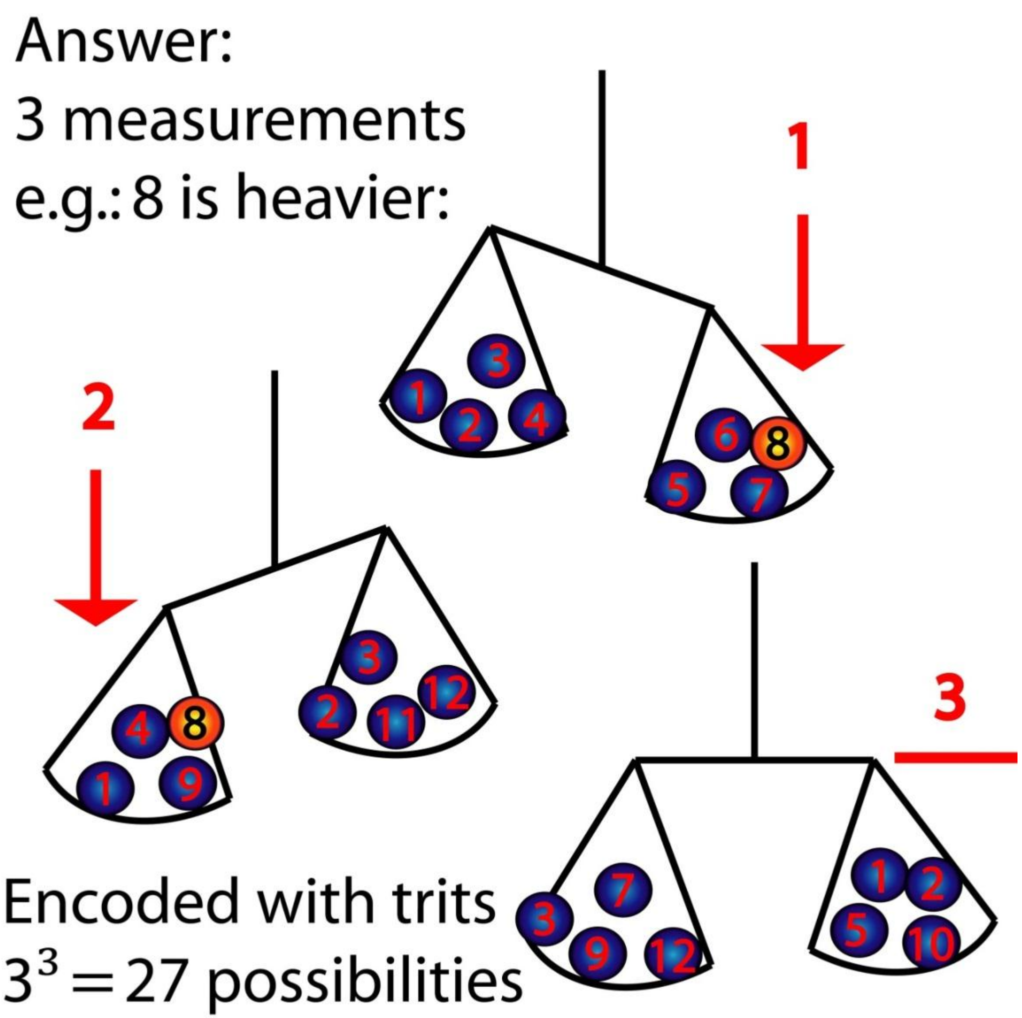
Conclusion

- Sparse methods are very powerful
- Two applications to illustrate the methods
- Widely applicable to many problems in physics
- Care needed with noise

Thanks for listening!

Ra.Inta@anu.edu.au

Solution to the twelve-ball problem (non-adaptive)



http://www.primepuzzle.com/leeslatest/12_ball_solution.html (2006)