

SPARSE METHODS FOR GRAVITATIONAL WAVE DETECTION

RA INTA,
THE AUSTRALIAN NATIONAL UNIVERSITY

BACKGROUND

SPARSE METHODS

The most useful data and signals generally have a degree of redundancy or compressibility and this information may be exploited in powerful ways.

For example, the highly successful JPEG-2000 image compression standard is based on the fact that most real images have a high degree of redundancy in a wavelet representation. The famous twelve ball problem (Figure 1) illustrates that you don't need twelve linear measurements (equations) to solve for twelve unknowns if you know the solution is sparse.

The amount of redundancy may be characterised by the number of non-zero (or non-negligible, in the case of compressible or noisy data) coefficients in the basis of interest, known as the sparsity, S , of the signal. The basis used to represent these signals is known as a sparse basis. For JPEG-2000, the sparse basis is a wavelet pixel space. For the twelve ball problem, the ball mass representation has a sparsity of $S=1$.

You have twelve 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?



Figure 1: The twelve-ball problem. How many measurements does it take to find the odd ball? (Hint: less than twelve)

GRAVITATIONAL WAVES ARE SPARSE

Much of the effort involved in sparse methods is in identifying a sparse basis. In a way, much of the description of physical phenomena involves finding a strongly sparse signal in a particular representation. In the case of gravitational wave (GW) data analysis, there are four main sparse representations, depending on the GW source (Table 1).

Phenomenon (GW)	Signal type	Sparse basis/frame
Rotating non-axisymmetric neutron stars	Periodic	Fourier domain
Binary compact object coalescence	Chirp/Gabor	Chirplet/wavelet
Supernovae, other transients	Impulse	Time domain
Stochastic GW background	Correlation	Correlation space

Table 1: Gravitational waves are sparse. Gravitational wave signals are analysed in four main sparse representations.

WHAT THIS POSTER IS ABOUT
There is a growing interest in the potential of sparse methods. While this work is in a preliminary stage, this poster is an attempt to illustrate two, amongst potentially many, possible applications of sparse methods in the analysis of GW data. Some of the pitfalls and potential issues are also discussed.

REDUCED COMPUTATION WITH THE SPARSE FFT

Recent developments [1] have produced a Fast Fourier Transform (FFT) implementation that exploits the sparse nature of underlying signals. This new implementation has very low relative computational complexity, and has been shown to be an optimum. The computational complexity for an N -point FFT, with known sparsity S , is $O(S \log(N))$, hence potentially giving a speed-up over the FFT ($O(N \log(N))$) proportional to the sparsity of the signal. This speed-up is very generic, applying for all ranges of sparsity.

The algorithm implements a non-recursive Orthogonal Matching Pursuit, similar to the OFDM (Orthogonal Frequency Division Multiplexing) schemes used in telecommunications.

Unlike the FFT, the assumption of sparse signals allows the use of very wide bin sizes: in the initial pass the number of bins used is \sqrt{SN} . The identified coefficients are removed from these bins rather than from the signal, so that the signal recovery is based on whole blocks, rather than recovery of each coefficient. The windowing functions used to determine the bins are leakage-free constructions (e.g. a Gaussian convolved with a box-car filter; Figure 2).

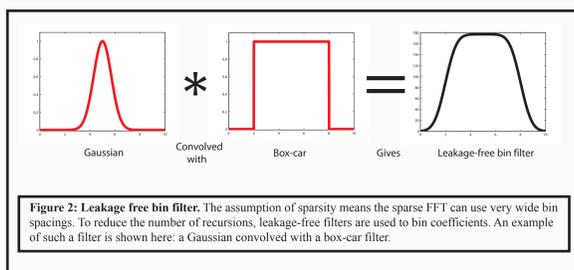


Figure 2: Leakage free bin filter. The assumption of sparsity means the sparse FFT can use very wide bin spacings. To reduce the number of recursions, leakage-free filters are used to bin coefficients. An example of such a filter is shown here: a Gaussian convolved with a box-car filter.

HOW THE sFFT MIGHT HELP GRAVITATIONAL WAVE DATA ANALYSIS

Continuous wave (CW) GW data analysis looks for quasi-monochromatic signals (spectral lines) potentially given off by certain rotating neutron stars. However CW searches are often so computationally intensive that computing resources comprise a serious constraint on targets to search for. These CW searches are based on spectral methods that heavily rely on the FFT, so an improvement in run-time would allow more candidates to be targeted.

A simplified example of such a search might be a neutron star with known position but unknown CW frequency in a 200 Hz search band; if the data is efficiently heterodyned, the effective Nyquist sampling rate is 400 Hz. In this band, we expect at most a single spectral line from the CW source, and perhaps ten spectral lines as may be found in LIGO or Virgo GW laser interferometer data, so we can expect the sparsity to be at most $S=20$. Finally, LIGO and Virgo CW data is often analysed in 30 minute (1,800 sec) chunks to optimise noise stationarity. Hence, in this case, analysis would require FFT sizes of $N=1,800 \times 400 = 720,000$ points.

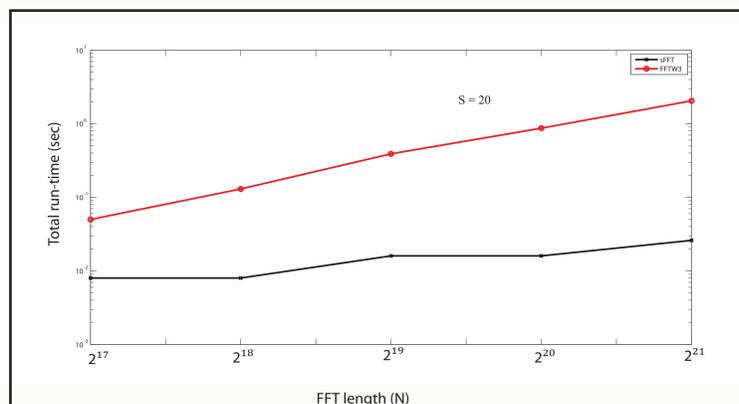


Figure 3: sFFT performance for varying FFT lengths. Compared to the currently most efficient implementation of the FFT, FFTW3, the sFFT performs as accurately but over 2.5 times the speed for a 20-sparse, ~500,000 point FFT.

PERFORMANCE

Sparse FFT simulations were applied to vectors having 20 spectral lines (i.e. $S=20$) of random frequencies and amplitudes, using a publicly available sFFT implementation [2]. The run-time performance of signal reconstruction was compared to a FFTW3 implementation for a range of FFT sizes (Figure 3).

For the example given above, a conservative comparison would be for $N=2^{19} \sim 500,000$ points. In this case, the run-time for the FFTW3 implementation was 40.2 ms per FFT, while that for the sFFT was 14.8 ms.

In other words, the sFFT showed a speed-up of 250% over FFTW3! For $N=10^6 \sim 2^{20}$, this speed-up increases to over 450%!

These bounds are conservative estimates; many searches cover higher bandwidths (requiring higher N) while the performance of the sFFT is likely to improve in the near future, making use of processor look-up tables, similar to how the FFTW3 obtains high performance.

ISSUES AND FUTURE WORK

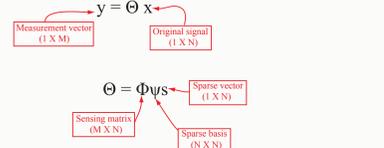
Although there have been apparently promising studies on the performance of the sFFT with noise [1], the comparison with another sparse FFT algorithm. Hence detailed studies of how robust the sFFT performs on noisy data compared to existing FFT implementations will have to be performed.

Also, the assumption of sparsity of spectral lines from CW sources in GW data sets is valid only for Gaussian noise; in real detectors the noise curve are heavily frequency dependent. In order for this implementation to be useful for CW searches, a frequency localisation (band-passing) or other compromising scheme may have to be adopted.

IMPROVED TIMING VIA COMPRESSIVE SAMPLING

Compressive sampling (CS; also referred to as 'compressed sensing') is a powerful mathematical framework that guarantees the reconstruction of signals that would normally be impossible because of the Shannon-Nyquist sampling theorem: in order to sufficiently sample a signal to reconstruct it without aliasing, you are required to sample at a rate that is at least twice the signal bandwidth [e.g. 3].

Consider a length N signal vector x (i.e. $x \in \mathbb{R}^N$) that is sparse in some basis ψ , so that $x = \psi s$ where s is an S -sparse vector. Now consider a linear measurement system (Equation 1), with a length M output vector y (i.e. $y \in \mathbb{R}^M$), characterised by a sensing matrix Φ , so that $y = \Phi x = \Phi \psi s$. In this framework, the Shannon-Nyquist sampling theorem is just an expression of a 'well-known' property of linear systems: to solve an equation of N unknowns, you require at least N equations. Hence for the system we consider here, $M \geq N$. However, Figure 1 (the twelve ball problem) has hopefully convinced you that this is not necessarily the case, and you can still reconstruct signals accurately for $M < N$ if you know the underlying signal to be sparse. This is the essence of compressive sampling; a more formal description is in [3] and Equations 1 and 2.



Equation 1: A linear measurement system. Here an underlying signal x is transformed by a linear measurement system to give a set of measurements y . In a GW context, $\Phi = D \times A \times W \times \dots$ i.e. Doppler shift parameters, antenna functions, whitening matrices and any other linear transform necessary to render the data 'useable'.

To solve for s , impose sparsity:

$$\min \|s\|_0 \quad \text{s.t.} \quad \|y - \Phi x\|_2 = 0$$

...but this is a combinatorial problem! (Need to guess up to $\binom{N}{S}$ coefficients)

Instead, almost as good (esp. for large N):

$$\min \|s\|_1 \quad \text{s.t.} \quad \|y - \Phi x\|_2 = 0$$

In the case of noise, set a residual threshold:

$$\min \|s\|_1 \quad \text{s.t.} \quad \|y - \Phi x\|_2 \leq \epsilon$$

Equation 2: Compressive sampling. If we know a signal is sparse, we can impose this constraint in an optimisation framework. Formally, this involves minimising the ℓ_0 -norm of the sparse vector s . However, this is an NP-hard problem. It has been shown that this condition is effectively guaranteed for most signals by minimising the ℓ_1 -norm [3].

RECONSTRUCTION SCHEMES: ORTHOGONAL MATCHING PURSUIT (OMP)

One of the most important aspects of compressive sampling is the choice of optimisation method to reconstruct the underlying signal from the measurements. Here we choose to illustrate this problem using Orthogonal Matching Pursuit (OMP; Figure 4). Based on a recursive greedy algorithm, this is one of the most robust and accurate CS reconstruction methods. In addition, it is straight-forward to implement a noise based stopping criterion by parameterising the magnitude of the residual noise, ϵ , giving the possibility of tuning the algorithm for a given noise level. For example, given Gaussian noise, a chi-squared statistic may be constructed so that $p(\chi^2 \leq \epsilon^2) = \alpha$, for a confidence level α .

Put simply OMP automatically constructs a basis of dimensionality (of order S) of the underlying vector by taking the S most significant coefficients. In practise it is very similar to the CLEAN algorithm used in radio astronomy [4] and shares elements of Principal Component Analysis.

HOW OMP MIGHT HELP GRAVITATIONAL WAVE DATA ANALYSIS

One of the most promising GW sources detectable by ground based detectors are Compact Binary Coalescence (CBC) events. In the stationary phase approximation, they may be seen to 'chirp,' i.e. modulate in frequency, over a short period of time. A rapid detection algorithm, known as Omega, uses a sine-Gaussian wavelet based analysis of the time-frequency plane to identify chirp-like signals. However the Omega pipeline has a limited localisation of signals in the time-frequency plane which can contribute to large position reconstruction errors. This can be problematic in a multi-messenger astronomy context [5]. Hence it would be desirable to localise the signal further in the time-frequency plane.

A proposed solution to give a similarly rapid, but more localised, signal reconstruction in the time-frequency plane is known as Chirpletised Omega [6]. This adds a slope parameter, d , in the time-frequency plane to the Omega (sine-Gaussian) wave-form (Equation 3). For a CBC event with total mass $< 20 M_{\odot}$, Chirpletised Omega templates cover ten times more parameter space than Omega, resulting in better signal localisation, and giving a corresponding SNR enhancement of ~40%. If we ignored constraints with aliasing, template coverage would be ten times higher again. For ~100 M_{\odot} , aliasing restricts the number of templates to one hundredth the theoretical limit.

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

Equation 3: Chirpletised Omega. Omega wave-forms are based on sine-Gaussians, i.e. a sinusoid of frequency f is localised in time by a Gaussian with a quality factor Q . These wave-forms are rapid to calculate and reasonably capture a signal in the time-frequency. Chirpletised Omega extends this concept by introducing a chirp parameter, d , giving a slope to the wave-form.

PERFORMANCE

A set of ten spectral lines ($S=10$) were selected with random frequencies and amplitudes within a $N=1024$ vector. The number of samples selected from a random measurement system was $M=100$, i.e. roughly one tenth the Shannon-Nyquist limit. The stopping criterion for a one-dimensional OMP implementation was based on a residual noise level $\epsilon \leq 0.1$.

Figure 5 shows the result of an arbitrarily chosen reconstruction; the reconstruction agrees well with the original signal. In this example the choice of ϵ is arbitrary, and only affects the run-time. However, in a similar experiment with much lower signal to noise, the reconstruction begins to fail such that the reconstruction forms additional points around the threshold level.

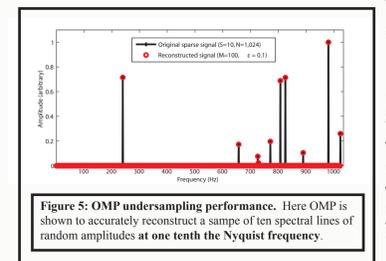


Figure 5: OMP undersampling performance. Here OMP is shown to accurately reconstruct a sample of ten spectral lines of random amplitudes at one tenth the Nyquist frequency.

ISSUES AND FUTURE WORK

Because orthogonal matching pursuit projects the most significant elements of a set into a new basis set of reduced dimension (of order S), contamination with noise can present a significant problem. The choice of residual noise parameter, ϵ , can be important in reducing the level of contamination. However, this choice of a noise-based stopping criterion means that the run-time of the OMP algorithm is exceedingly poor ($O(N^3)$). Non-CPU based hardware accelerators such as FPGAs have been shown to have a 3,000 times speed-up for OMP reconstruction. Finally, a CS technique with much promise in this direction, having a more reasonable computational complexity of $O(SN \log(N))$, and being deterministic, are CS-based 'chirp sensing codes' [7].

CONCLUSION

Many common signals found in nature are sparse (or at least compressible) in a particular representation. Likely candidate sources of gravitational waves are no exception, and associated data analysis techniques are effectively sparse in a number of representations. There are many ways this sparse information may be exploited. As examples, a recent sparse FFT implementation has been shown to have a superior run-time to that of FFTW3, possibly alleviating computational bounds in CW searches. A compressive sampling algorithm, such as orthogonal matching pursuit, may improve localisation of chirped burst signals in the time-frequency plane. Although powerful, sparse methods should be applied with caution. Perhaps the most serious issue relates to how gracefully these methods perform with noisy signals; a completely noisy signal is obviously not sparse.

REFERENCES

- [1] H. Hassaneh, P. Indyk, D. Katabi and E. Price: "Nearly Optimal Sparse Fourier Transform," *arXiv* 1201.2501v1 (2012)
- [2] sFFT code available from <http://groups.csail.mit.edu/netmit/sFFT/code.html>
- [3] E. Candès, J. Romberg and T. Tao: "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Information Theory* 52(2):489-509 (2006)
- [4] J. Högbom: "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines," *Astron. & Ap. Suppl.* 15, p.417 (1974)
- [5] J. Abadie et al.: "Implementation and testing of the first prompt search for gravitational wave transients with electromagnetic counterparts," *A&A* 539:A124 (2012)
- [6] É. Chassande-Mottin et al.: "Detection of GW bursts with chirplet-like template families," *CQG* 27:194017 (2010)
- [7] L. Applebaum et al.: "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Applied Comp. Harm. Analysis* 26(2):283-290 (2009)

CONTACT

For more information, please email:
Ra.Inta@anu.edu.au
or scan this QR code:

