2018 26th European Signal Processing Conference (EUSIPCO)

Sparse Time-Frequency Representation of Gravitational-Wave Signals in Unions of Wilson Bases

Quentin Bammey*, Philippe Bacon†, Eric Chassande-Mottin†, Aurelia Fraysse‡, Stéphane Jaffard§
*École Normale Supérieure Paris-Saclay, CMLA
Cachan, France

†Université Paris Diderot
Astroparticule et Cosmologie APC
CNRS/IN2P3, CEA/Irfu, Obs. de Paris
Sorbonne Paris Cité, Paris, France

‡Laboratoire des signaux et systèmes
CNRS UMR8506
Centrale Supelec Université Paris-Sud
Paris, France

§LAMA
CNRS UMR8050
Université Paris-Est
Paris, France

***Abstract*—We investigate the question of obtaining a reduced time-frequency description of a chirp type signal that can be used as reference pattern in time-frequency searches. This is particularly relevant for searches of transient gravitational waves from astrophysical sources such as the mergers of neutron stars and/or black holes, the main area of this study. Sparse approximation algorithms that allow constraints on the approximation error do not perform well when the decomposition bases are redundant. This study puts in evidence some of the shortcomings of sparse approximation algorithms when dealing with unions of highly correlated bases, a case that currently lacks of a comprehensive mathematical analysis, and proposes solutions to mitigate them. We propose a variation of the matching pursuit algorithm that improves its robustness in the context of gravitational waves patterns construction. We also compare this algorithm to standard sparse approximation methods.**

I. Introduction

Gravitational waves are ripples in the metric of spacetime that propagate at the speed of light [1]. Their existence is a long-standing prediction of Einstein's theory of general relativity. Recently, the LIGO and Virgo detectors detected gravitational waves from distant astrophysical sources including binary black holes (BBH) [2] and binary neutron star [3]. For the first time dark sources that emit little or no photons at all are observed directly through a different radiation than electromagnetic waves. These major discoveries herald a new era for astronomy.

Finding rare and weak gravitational wave signals in non-stationary and non-Gaussian instrument noise is a particularly challenging problem. A range of data analysis approaches has been applied (see e.g., [4] for a review) to detect the gravitational-wave signature from transient sources such as binary mergers of neutron stars and/or black holes.

Transient searches, our focus here, identify bright, time-coincident and phase-coherent pixels in time-frequency representations of the data from multiple detectors. The data analysis pipeline Coherent WaveBurst [5] has been successfully applied in this context.

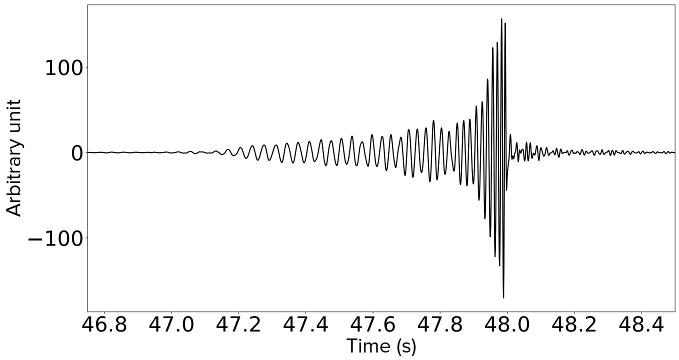
When searching for a particular astrophysical source, the gravitational waveform model can be used to enhance the search sensitivity. An approach is to search specifically for the time-frequency pattern(s) associated with the waveform model [6], [7]. Such an approach thus aims at selecting components that are likely to describe the gravitational wave signal and prevent the search algorithm from selecting those due to transient noise. The expected improvement is larger when the signal model can be completely characterized by a small number of time-frequency components.

We consider a range of sparse approximation algorithms to obtain the "template" time-frequency pattern associated with the expected gravitational-wave signals from binary mergers. Figure 1 provide few examples of such chirp signals. Here, we are interested in the sparse approximation of the *(noise-free) signal model*. The use of sparsity to estimate the gravitational wave signal from the noisy observations is explored elsewhere (see, e.g., [8]).

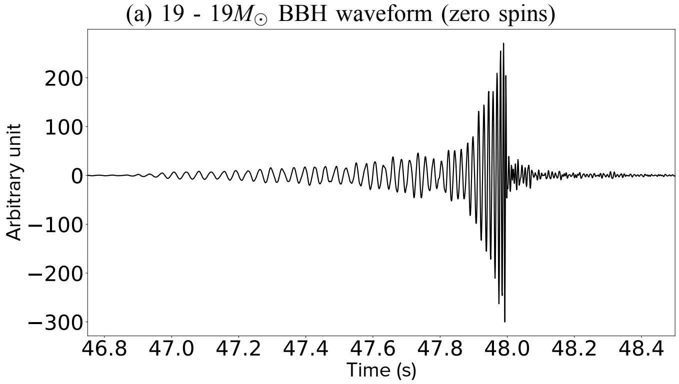
In order to determine a compact representation of model signals, a classical method is provided by the Matching Pursuit (MP) algorithm [9]. Here the approximation of a signal is constructed by iteratively determining the largest coefficient in the decomposition. However, in the present context of highly correlated atoms in the decomposition dictionary this method is known to be inefficient at some point. In this paper we determine a greedy approach, based on the MP algorithm and adapt it to the time-frequency decomposition of chirp-like signals associated with black hole binary mergers. We also compare our method to different classical sparse approximation algorithms.

This paper is structured as follows. In Section II we present the state of the art decomposition method. Section III is devoted to the description of classical sparse approximation methods. In Section IV, we construct a new method. Finally Section V presents some results whereas Section VI concludes the paper.

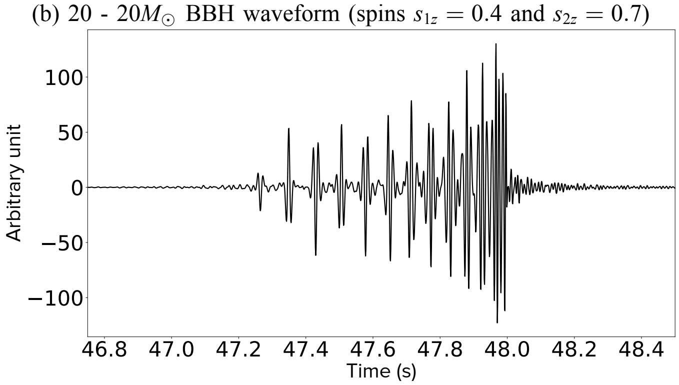
ISBN 978-90-827970-1-5



(a) 19 - 19$M_\odot$ BBH waveform (zero spins)



(b) 20 - 20$M_\odot$ BBH waveform (spins $s_{1z} = 0.4$ and $s_{2z} = 0.7$)



(c) 15 - 15$M_\odot$ BBH waveform (zero spins, eccentricity e = 0.3)

Fig. 1: Theoretical models of gravitational-wave signals emitted during the merger of two black holes. The waveform is a chirp signal with a time increasing (power-law) instantaneous frequency. Several examples are shown where the astrophysical parameters are varied, such as the component masses and spins s_{1z} and s_{2z} or the eccentricity e of the binary orbital motion. Those signals are processed through a whitening filter obtained from the detector noise power-spectral distribution. This filtering discards the part of the original signal where the instrumental noise is large (low and high frequencies, below $\sim$ 30 Hz and above few kHz) and retains the frequency band where the noise is low.

II. Sparse time-frequency approximation of chirp signals

Coherent Waveburst maps time series data to the time-frequency plane by projecting onto Wilson bases, through the Wilson-Daubechie-Meyer (WDM) transform [10], [11]. Wilson bases are variation of the well-known Gabor decomposition, with the additional advantage of being orthonormal bases with a good time-frequency localization. They are composed of functions distributed on a regular time-frequency lattice and obtained by the linear phase cosine modulation of a window. Coherent Waveburst uses the Meyer scaling function (which has a compactly supported Fourier transform) as the window [11] and considers a collection of Wilson bases (typically seven to nine) based on different window durations and bandwidths. These bandwidths are distributed in power of two, ranging from 1 to 64 Hz.

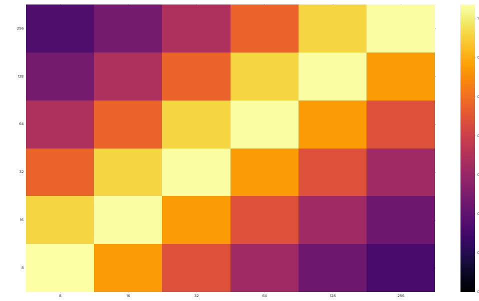


Fig. 2: Redundancy measurement between two Wilson bases with windows of different durations/bandwidth. The redundancy is characterized by the largest correlation (scalar product) between all pairs of elements from the two bases. This measurement is performed for different window bandwidths, ranging from 2 Hz (which corresponds to 256) to 64 Hz (i.e., 8).

The union of these orthonormal bases forms a redundant dictionary, which constitutes a tight frame, see Fig. 2. We denote $y \mapsto x = Wy$, the vector composed by the coordinates of the orthogonal projections of x on each vector of the dictionary W.

Given the astrophysical model y (preconditioned by a whitening that selects the signal content in the detector bandwidth), the goal is to obtain a sparse approximation x that satisfies

$$\min_x \|x\|_0 \quad \text{s. t.} \quad \left\|y - W^T x\right\|_2 \leq \delta \tag{1}$$

where $r \triangleq y - W^T x$ denotes the approximation error or residual and $\|\cdot\|_p$ denotes the L_p norm. Our problem is fairly generic and common to many applications: seek the simplest linear approximation from a dictionary up to a given error.

The minimization problem in (1) is NP-hard, hence the need for reasonable approximations.

To cover the astrophysical parameter space, the set of models y typically include about 10,000 waveforms. A key astronomical issue is to fasten the already existing algorithms in order to lower the time needed to precisely identify the characteristics of the gravitational wave recorded.

III. State-of-the-art sparse approximation algorithms

A. Standard Matching Pursuit

Matching Pursuit [9] is one of the simplest sparse approximation algorithms. The approximation is constructed iteratively by selecting the pixel with the largest coefficient in the

ISBN 978-90-827970-1-5 © EURASIP 2018

residual transform and adding it to the current approximation. The iteration can be formalized as follows:

$$x^{n+1} = x^n + \mathcal{H}_1(Wr^n) \tag{2}$$

where the residual is defined as $r^n \equiv y - W^T x^n$ and $\mathcal{H}_s(x)$ is the hard-thresholding operator, that sets all components of x but the s largest ones to zero.

B. Orthogonal Matching Pursuit

The Orthogonal Matching Pursuit (OMP) [12], [13] makes use of a Gram-Schmidt procedure to orthogonalise the pixels found by the MP.

If $p^1, \ldots, p^n$ are the pixels selected up to iteration n, the pixel $\tilde{p}^{n+1}$ is selected as the pixel whose energy in the residual is highest, as in the MP algorithm:

$$\tilde{p}^{n+1} = \mathcal{H}_1(Wr^n)$$

where $r^n = y - \sum_{i=1}^n p^i$ is the residual at the previous step. This pixel is then orthogonalised with regard to the previously selected pixels:

$$p^{n+1} = \tilde{p}^{n+1} - \sum_{i=1}^{n} \frac{\langle \tilde{p}^{n+1} \mid p^i \rangle}{\|p^i\|^2} p^i.$$

Although the selected pixels are not modified in future iterations, the orthogonalisation makes each new pixel change the projection of the global approximation on previously selected pixels. By linearity, each iteration can thus be understood as selecting a pixel from the residual, then updating each of the previously selected pixel coefficients from the newly selected pixel.

C. Iterative hard thresholding

Iterative hard thresholding (IHT) [14] algorithms make repeated use of the hard thresholding operator to converge towards a locally optimal solution of the following problem:

$$\min_x \|r\|_2 \quad \text{s. t.} \quad \|x\|_0 \leq s. \tag{3}$$

where s is the number of pixels used.

The algorithm repeats the following iteration

$$x^{n+1} = \mathcal{H}_s(x^n + \mu W r^n) \tag{4}$$

where the step size μ should be smaller than the W's operator norm to ensure convergence.

The Normalised Iterative Hard Thresholding (NIHT) [15] proposes to optimise the step size μ at each iteration to improve the convergence speed.

Assuming the approximation support Γ_x does not change between two iterations, the iteration amounts to a gradient descent step, and the optimal step size is

$$\mu^{n+1} = \mu(x^n, g^n) = \frac{(\Gamma_{x^n}(g^n))^T \Gamma_{x^n}(g^n)}{(\Gamma_{x^n}(g^n))^T \left(W_{\Gamma_{x^n}} W_{\Gamma_{x^n}}^T W_{\Gamma_{x^n}} \Gamma_{x^n}(g^n) \right)} \tag{5}$$

where $g^n = Wr^n$ is the WDM transform of the residual. A temporary iteration $\tilde{x}^{n+1} = \mathcal{H}_s(x^n + \mu^{n+1} W r^n)$ is computed. If the support did not change, that is, if

$$\Gamma_{\tilde{x}^{n+1}} = \Gamma_{x^n}, \tag{6}$$

then this iteration is optimal.

If the support did actually change, then the iteration necessarily goes toward convergence if

$$\mu^{n+1} \leq (1-c) \frac{\|\tilde{x}^{n+1} - x^n\|_2^2}{\|W^T(\tilde{x}^{n+1} - x^n)\|_2^2} \tag{7}$$

for a small fixed constant c. If this is not the case, then μ^{n+1} is shrunk and a new iteration is computed until either (6) or (7) is verified.

It must be noted that the solved problem is slightly different from (1), in that instead of minimising the number of pixels with a constraint on the approximation error, it minimises the approximation error while fixing the number of pixels. To solve our original problem, we need to scan the sparsity levels s and find the smallest that allows an approximation error lower than δ. The IHT thus has to be repeated several times, introducing a significant computational overhead to the algorithm.

D. Other sparse approximation methods

Several methods expand on the IHT algorithm to improve its speed or results. A survey of them can be found in [16]. A widely used approach is to relax the L_0 norm in (1) by another L_p norm for $p \leq 1$, sacrificing some sparsity in order to improve robustness to noise.

In particular, using the L_1 norm yields a convex problem known as the Basis Pursuit, to which a globally optimal solution can be found using the Iterative Shrinkage/Thresholding Algorithm (ISTA) [17], which replaces hard-thresholding with soft-thresholding in the IHT algorithm to optimise on the L_1 norm.

One of the main advantages of these methods is their ability to provide an approximation that is robust to noise, whereas MP and IHT-based algorithms may have trouble finding the right pixels in a noisy environment. Such an advantage is lost in our case, where we are working with templates – and thus are in a noiseless environment –, while the compromise on sparsity stays real, and is even increased in highly redundant bases[18] such as ours, as seen in Fig. 2.

Although powerful in the right environment, these algorithms are consequently irrelevant for our work, and shall not be discussed further in this article.

IV. Updating Matching Pursuit

While keeping the greedy approach of the MP for the choice of pixels, we can improve its sparsity by simply updating the values of already selected pixels, as inspired by viewing the OMP as an update on the coefficients. However, if we drop the orthogonality constraint of the basis, it makes more sense to update the previously selected pixels directly from the residual.

We define our iteration as

$$\begin{aligned} \tilde{x}^{n+1} &= x^n + \mathcal{H}_1(Wr^n) \\ \tilde{r}^{n+1} &= y - W^T \tilde{x}^{n+1} \\ x^{n+1} &= \tilde{x}^{n+1} + \mu^{n+1} \Gamma_{x^n}(\tilde{r}^{n+1}) \end{aligned} \tag{8}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

```
function UPDATING_MP(y, δ)
    x ← 0                   ▷ Initialize time-frequency approximation
    r ← y − W^T x                                  ▷ Compute residual
    repeat
        g ← Wr              ▷ Compute residual's Wilson transform
        p ← argmax_n ||g_n||_2                      ▷ Select best pixel
        x ← x + p                           ▷ Add it to approximation
        r ← y − W^T x                               ▷ Update residual
        g ← Wr
        μ ← μ(x, g)                        ▷ Update step size, see (5)
        t ← Γ_x(g)                 ▷ Get non-zero pixels in residual
        x ← x + μt                       ▷ Update with these values
        r ← y − W^T x                               ▷ Update residual
    until ||r||_2 < δ                ▷ Loop until precision is small
end function
```

Algorithm 1: Pseudo-code for the Updating Matching Pursuit

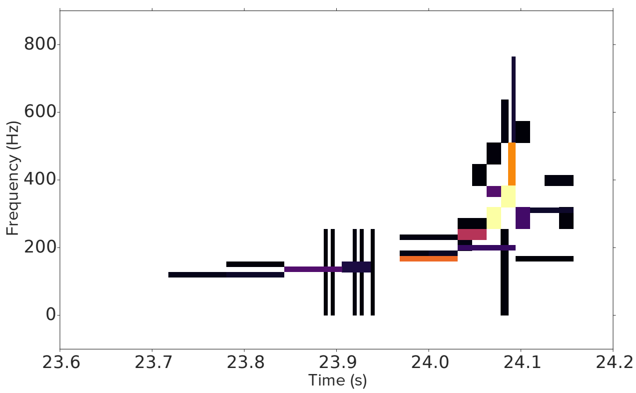


Fig. 3: Example of a sparse time-frequency approximation of gravitational-wave signal, obtained with UMP for an approximation error of 10%. The amplitudes are expressed in arbitrary units.

where μ^{n+1} is the step size at the next iteration. It amounts to projecting the residual on the selected pixels to modify their values between each step of the MP algorithm.

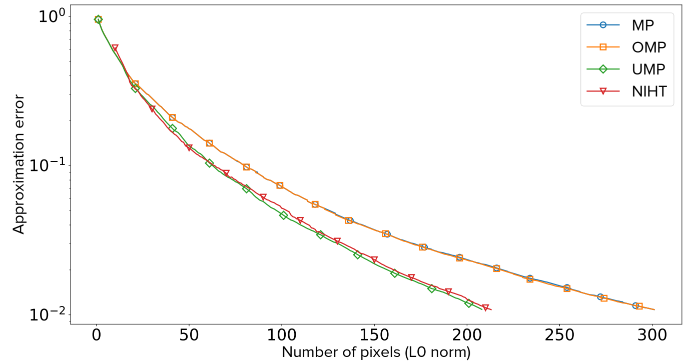
The update necessarily improves the approximation if $\mu^{n+1} \leq \frac{1}{|W|}$. However, we can optimize the step size in a similar way as in the NIHT. As the approximation support does not change during the update step (8), we can simply compute μ^{n+1} as in (5), see Algorithm 1.

In the following we refer to this original algorithm as *Updating Matching Pursuit* (UMP). As the original MP, it is linear in the number of selected pixels, although its cost is slightly higher since the step size computation requires additional transforms.

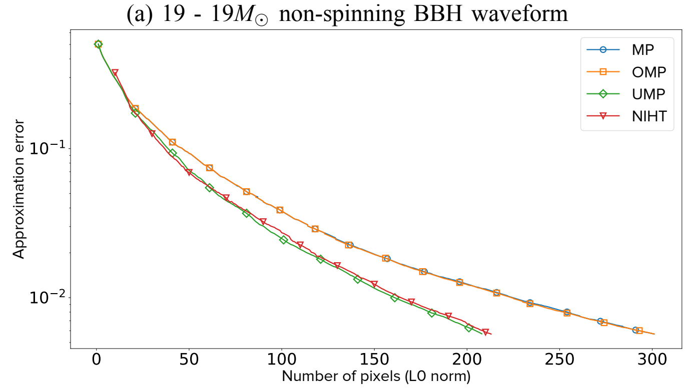
V. Results and discussion

We apply the methods of Sec. III to the gravitational-wave signals shown in Fig. 1. We measure the number of pixels needed to reach a given approximation error.

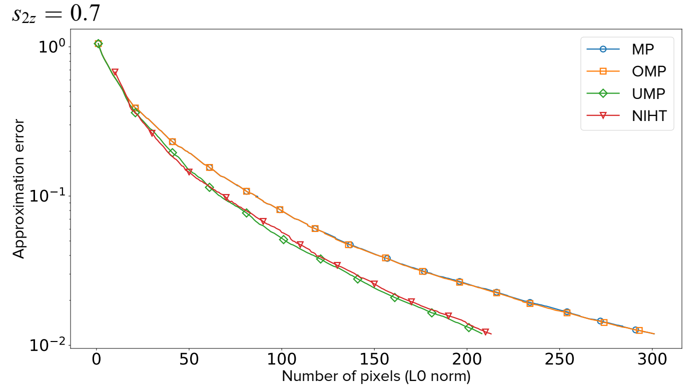
In the first iterations the standard matching pursuit algorithm is able to extract the main *coherent structures* [19] of the signal. Then the convergence slows down considerably due to the artefacts. Those artefacts make it impossible to go past



(a) 19 - 19$M_\odot$ non-spinning BBH waveform



(b) 20 - 20$M_\odot$ BBH spinning waveform with spins $s_{1z} = 0.4$ and $s_{2z} = 0.7$



(c) 15 - 15$M_\odot$ non-spinning BBH waveform with eccentricity $e = 0.3$

Fig. 4: Approximation error vs. sparsity level obtained with MP, OMP, normalised IHT and UMP. For the normalised IHT, the algorithm was first called with 10 pixels. When an iteration brought an improvement smaller than 0.1% of the signal energy, it was called again with one more pixel, initialised with the previous result.

a certain approximation level within a reasonable number of pixels.

In our highly redundant basis, orthogonalising the pixels does not correct their coefficients and the OMP does not yield better results than the MP.

The IHT provides the best overall results. As the sparsity level needs to be tuned, however, its computational cost is much larger, as seen in Fig. 5. Furthermore, the high operator norm of our basis means that the step size of this method must be shrinked, adding even more to its computational cost.

ISBN 978-90-827970-1-5 © EURASIP 2018

	Number of pixels	Computation time
UMP	54	31s
MP	69	14s
OMP	69	163s
NIHT	46	1808s

Fig. 5: Number of needed pixels and computation time to approximate a 20 - 20$M_\odot$ BBH spinning waveform with spins $s_{1z} = 0.4$ and $s_{2z} = 0.7$ with the UMP, MP, OMP and normalised IHT algorithms.

	Number of pixels	Computation time
UMP	72	12s
MP	91	6s
OMP	71	11s
NIHT	68	962s

Fig. 6: Number of needed pixels and computation time to approximate a 20 - 20$M_\odot$ BBH spinning waveform with spins $s_{1z} = 0.4$ and $s_{2z} = 0.7$ with the UMP, MP, OMP and normalised IHT algorithms, when we only use one Wilson basis instead of six to remove redundancy in the basis.

The UMP algorithm corrects the MP artefacts and provides results that are almost as good as those yieled by the IHT – and sometimes better – for a fraction of the computational cost. Needed computational effort per iteration is only triple that of the MP. In our case, the greedy pixel selection is good enough as long as their values are updated.

The updating of pixel coefficients is useful because of the redundancy of bases, which is seen in Fig. 2. When this is not the case, for instance in Fig. 6 where we only use one orthogonal basis instead of a union of multiple bases, the OMP becomes as good as our method, although our UMP algorithm should still be faster with larger instances because of its linear complexity with regards to the number of iterations.

VI. Conclusion

The present study highlights that the standard MP algorithm fails to provide a good sparse approximation in the case of highly-redundant dictionaries such as unions of Wilson bases. Although hard-thresholding-based method perform well in thoses cases, more iterations are needed to converge when dealing with highly-correlated atoms. They are moreover extremely slow since this is the approximation error and not the sparsity itself that is constrained. The UMP algorithm appears to be the best trade-off with an almost as good approximation as the IHT algorithm and for a much lower computational cost. By updating the coefficients, the Updating Matching Pursuit makes the original Matching Pursuit algorithm much better at dealing with highly correlated bases, as it natively allows constraints on the approximation error instead of the sparsity. The current study has been performed on noise-free signals and can thus be extended to the case where signals are buried in some simulated or real instrumental noise as measured by GW detectors. It would give an appreciation on how the approximation we propose changes with the signal-to-noise ratio.

Finally, some extensions of the MP and OMP algorithms, for instance [20] or [21], show that the approximation quality can be improved by searching for multiple candidates at each step through a combinatorial approach. Such extensions could also be applied to the UMP algorithm and would probably further improve its results.

Acknowledgements: This research was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477, by the French Agence Nationale de la Recherche (ANR) under reference ANR-15-CE23-0016 (Wavegraph project).

References

[1] K. Thorne, "Gravitational radiation," in *300 Years of Gravitation*, S. W. Hawking and W. Israel, Eds. Cambridge University Press, 1987.

[2] B. Abbott *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.*, vol. 116, p. 061102, 2016.

[3] ——, "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral," *Phys. Rev. Lett.*, vol. 119, p. 161101, 2017.

[4] K. Riles, "Gravitational waves: Sources, detectors and searches," *Prog. Part. Nucl. Phys.*, vol. 68, pp. 1–54, 2013.

[5] S. Klimenko *et al.*, "Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors," *Phys. Rev. D*, vol. 93, no. 4, p. 042004, 2016.

[6] E. Chassande-Mottin *et al.*, "Wavelet graphs for the direct detection of gravitational waves," in *XXVth Gretsi Colloquium*, Lyon, France, 2015, arXiv:1710.09256.

[7] P. Bacon, V. Gayathri, E. Chassande-Mottin, A. Pai, F. Salemi, and G. Vedovato, "Driving unmodelled gravitational-wave transient searches using astrophysical information," 2018, arXiv:1805.04023.

[8] F. Feng, P. Bacon, E. Chassande-Mottin, and A. Fraysse, "Structured sparsity regularization for gravitational-wave polarisation reconstruction," This volume, 2018.

[9] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Proc.*, vol. 41, no. 12, pp. 3397–3415, 1993.

[10] I. Daubechies, S. Jaffard, and J.-L. Journé, "A simple wilson orthonormal basis with exponential decay," *SIAM J. Math. Anal.*, vol. 22, no. 2, pp. 554–573, 1991.

[11] V. Necula, S. Klimenko, and G. Mitselmakher, "Transient analysis with fast Wilson-Daubechies time-frequency transform," *J. Phys. Conf. Ser.*, vol. 363, p. 012032, 2012.

[12] Z. Z. Geoffrey M. Davis, Stephane G. Mallat, "Adaptive time-frequency decompositions," *Optical Engineering*, vol. 33, pp. 33 – 33 – 9, 1994. [Online]. Available: https://doi.org/10.1117/12.173207

[13] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, Nov 1993, pp. 40–44 vol.1.

[14] T. Blumensath and M. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 629–654, 2008.

[15] T. Blumensath and M. E. Davies, "Normalized iterative hard thresholding: Guaranteed stability and performance," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 298–309, 2010.

[16] Y. Wang, J. Zeng, Z. Peng, X. Chang, and Z. Xu, "Linear convergence of adaptively iterative thresholding algorithms for compressed sensing," *IEEE Transactions on Signal Processing*, vol. 63, no. 11, pp. 2957–2971, June 2015.

[17] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging. Sciences.*, vol. 2, no. 1, pp. 183–202, 2009.

[18] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Info Theory.*, vol. 49, no. 12, pp. 3320 – 3325, 2003.

[19] S. Mallat, *A wavelet tour of signal processing: the Sparse way*, 3rd ed. Burlington, MA: Academic Press, 2009.

[20] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6202–6216, Dec 2012.

[21] S. Kwon, J. Wang, and B. Shim, "Multipath matching pursuit," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2986–3001, May 2014.

ISBN 978-90-827970-1-5 © EURASIP 2018