Low-latency search for gravitational-waves: prospects for near real time detection of compact binary coalescence

Kipp Cannon, Adrian Chapman, Chad Hanna, Drew Keppel, Antony C. Searle, and Alan J. Weinstein
The LIGO Laboratory - California Institute of Technology, Pasadena, CA 91125
(Dated: October 5, 2010)

PACS numbers:
Keywords: Gravitational Waves, Laser Interferometry

I. INTRODUCTION

Modern gravitational-wave observatories are likely to deliver on the promise that their name implies by observing multiple gravitational-wave signals from compact binary coalescence within the coming decade [1]. There is now a world-wide network of gravitational-wave observatories including the laser interferometers LIGO, VIRGO, GEO, and TAMA, which should assure that coincident detections will be possible among at least three separate sites lending confidence to any detections. [2–6]. However, having an electromagnetic or neutrino counterpart to a gravitational-wave detection would not only increase the confidence in the detection but would also greatly improve the astrophysical information available from the event.

Work has already commenced to trigger gravitational-wave searches from electromagnetic observations [7]. LIGO has recently ruled out the possibility of GRB070201 originating from the merger of a neutron star / neutron star or neutron star / black hole system in the Andromeda galaxy [8]. The still open question is, will gravitational-wave observations be able to trigger electromagnetic followup?

This question has been investigated and it seems at least possible that gravitational-wave detections could provide a region of the sky to prompt electromagnetic observation followups [9]. Much work is currently underway in providing the best possible source localization from gravitational-wave detector networks [10–12]. Colloborations are currently forming to provide infrastructure for the gravitational-wave Astronomers to provide targets of opportunity for electromagnetic Astronomers [13]. Most models that predict simultaneous gravitational-wave and electromagnetic observations also predict that the peak amplitude of electromagetic radiation will occur soon after gravitational-wave emmission [9]. Thus in order to maximize the chance of a successful electromagnetic followup the latencey of gravitational-wave signal analysis must be made minimal.

The parameter space of compact binary coalescence signals is large [14, 15] . It is a computationally burdensome task to analyze these signals with even moderate latency [16]. This work will describe how to exploit degeneracy in the signal parameter space to answer more quickly whether or not a gravitational-wave is present, with what confidence, and where it is coming from. Specifically we will explore using the singular value decomposition (SVD) to reduce the effective number of filters required to search the data. We note that others have applied the use of SVD to gravitational wave data analysis to analyze optimal gravitational-wave burst detection [17, 18] and coherent networks of detectors [19]

The paper is organized as follows. First we provide an overview of the standard method for detecting compact binary coalescence signals and describe how it can be modified to accomodate low latency analysis. We then describe the pipeline we have constructed to implement these changes. To validate the approach we present results of simulations and finish with some concluding remarks.

II. METHOD

A. The standard approach: matched filtering

Compact binary coalescence searches typically employ matched filtering as the first step to locating a gravitational wave signal [20]. A template waveform is convolved with the whitened data, and if properly normalized it produces a signal-to-noise ratio (SNR) that indicates the probability of a signal being present [21]. Since the parameters of the signal are not known apriori, multiple templates must be filtered for a wide range of parameters. Let us assume that the waveforms are simply parameterized by component mass and denote the ith template waveform of N templates necessary to search a given parameter space as $h_i(t) = f(m1, m2, t)$. The SNR, ρ_i for that template is

$$\rho_i(t) = \int h_i(t) s(t - \tau)\, d\tau, \tag{1}$$

where $s(t)$ is the detector data stream suitably whitened having mean zero and unity variance. Time domain convolutions don't have any inherent latency. The template waveform can be "slid onto" the detector data as it is produced giving the possibility of measuring a peak SNR at the time of inspiral coalescence. Typically it is not tractable to compute SNR via a time domain convolution as indicated in (1) because it involves an $\mathcal{O}(N^2)$ operation per template. Instead the convolution theorem is applied and the SNR is obtained via an inverse Fourier transform

$$\rho_i(t) = \int \tilde{h}_i(f)^* \, \tilde{s}(f) \, e^{-2\pi i f t} \, df, \qquad (2)$$

where the $\tilde{h}_i(f)$ and $\tilde{s}(f)$ denote the Fourier transforms of the template and detector data respectively. This operation is far faster to compute with $\mathcal{O}(N \log N)$ operations per template. However, the acausal nature of the Fourier transform introduces an inherent latency. One must wait until they have collected data as long as their signal and for minimally overlapping computations that means having a latency about as long as the template waveform. The latency can be reduced by having redudant overlapping transforms, but at an additional cost.

Thus we have that low latency searches are inherently more computationally costly than moderate latency searches that employ frequency domain matched filtering. The next section discusses a method to mitigate the cost of low latency searches.

B. Reducing the number of filters with singular value decomposition

It is clear that in order to reduce the cost of a low latency search for compact binaries of unknown mass we need to reduce the number of convolutions necessary to find the gravitational-wave signal. Rather than filter the i templates corresponding to binaries of different component masses it is possible to find a basis set of fewer filter u_j that can be linearly combined to reproduce ρ_i to some accuracy. For the remainder of this section we will drop the explicit time dependence of the convolution and recognize that the snr for any point in time is simply the innerproduct between a template waveform vector $\mathbf{\hat{h}_i}$ and a vector of the detector data $\mathbf{\hat{s}}$, $\rho_i = (\mathbf{\hat{h}_i}|\mathbf{\hat{s}})$. The goal is to have

$$\rho_i = \sum_{j=1}^{N'} M_{ij} (\mathbf{\hat{u}_j}|\mathbf{\hat{s}}) \qquad (3)$$

where the number of inner products is reduced from N to N'. In order to find the basis set $\mathbf{u_j}$ we use the singular value decomposition of the matrix $H_{ik} = \{\mathbf{\hat{h}_1}, \mathbf{\hat{h}_2}, ...\mathbf{\hat{h}_N}\}$. Where i is the index over the template number and k is the index over sample points. The singular value decomposition factors a matrix such that

$$H_{ik} = A_{ij} \, B_{jl} \, U_{lk} \qquad (4)$$

The matrix B_{jl} is a diagonal matrix of singular values ranked in order of importance of reconstructing the original matrix H_{ik}. However since a search for binary inspirals only needs waveform accuracy to a few percent to be successful it is possible to make an approximate reconstruction of H_{ik}

$$H_{ik} \approx A_{ij} \, B'_{jl} \, U_{lk} \qquad (5)$$

where the rank of B'_{jl} is less than the rank of B_{jl}. This reduces the number of rows in U_{lk}. Now we have a new basis of template waveforms derived from $U_{lk} = \{\mathbf{\hat{u}_1}, \mathbf{\hat{u}_2}, ...\mathbf{\hat{u}_{N'}}\}$ and can write (3)

$$\begin{aligned} \rho_i &= \sum_{j=1}^{N'} A_{il} \, B'_{lj} (\mathbf{\hat{u}_j}|\mathbf{\hat{s}}) \\ &= \sum_{j=1}^{N'} M_{ij} (\mathbf{\hat{u}_j}|\mathbf{\hat{s}}) \end{aligned} \qquad (6)$$

This result differs from other work that models gravitational-wave chirp signals in approximate ways [22–24] by starting with an exact representation of the desired template family and producing a rigorous approximation with a tunable accuracy.

C. Detection statistic

It is not necessary to reconstruct the SNR of all templates in order to determine if a signal is present. By thresholding on the sum of squares of the orthoganal decomposition described above it is possible to answer whether or not an inspiral like event is in the data. Then only potentially interesting times can be reconstructed on a sample by sample basis. Following the ideas in [25] we employ a detection statistic as

$$\Gamma^2 = \sum_l \sum_j {B'_{lj}}^2 (\mathbf{\hat{u}_j}|\mathbf{\hat{s}})^2 \qquad (7)$$

D. Signal consistency check

After reconstructing the SNR of an individual template it is possible to to check that the components along the SVD basis were consistent with expectation by performing a χ^2 fit to the decomposition.

$$\chi_i^2 = \sum_j \left[(\mathbf{\hat{u}_j}|\mathbf{\hat{s}}) - \rho_i / M_{ij} \right]^2 \qquad (8)$$

E. Selectively reducing the sample rate of the data and template waveforms

The computational cost of filtering can be further reduced by selectively downsampling certain intervals of the template waveform. Note that the frequency of gravitational radiation resulting from the inspiral of two compact object is given by the quadrupole aproximation as [26]

$$f_{\mathrm{GW}} = \frac{1}{\pi \mathcal{M}} \left[\frac{5}{256} \frac{\mathcal{M}}{t_c} \right]^{3/8} \qquad (9)$$

where $\mathcal{M}$ is the chirpmass in units of $G = c = 1$ and t_c is the time before the coalescence of the binary. Typically the filter is integrated to a finite frequency, f_e which corresponds to a time, t_e before the coalescence [20, 27–29]. Given the relationship $f/f_e = (t_e/t)^{3/8}$ it is possible to reduce a filter waveform into critically sampled bands by dividing them up in time. This is possible regardless of the time-frequency relationship as long as it can be computed rigorously and thus the method doesn't depend on (9). As a concrete example consider a waveform obeying (9) with a chirp mass of 1 $M_\odot$ for which we decide to cut the filter at $f_e = 2048$Hz (corresponding to $t_e = 0.00093$s). It is then possible to compute sub bands which are each critically sampled according to table I where it is assumed that the filter is cut to some finite length, 387s for this example. This idea has been demonstrated by the Virgo Collaborations MBTA pipeline [3, 4].

TABLE I: Example critically sampled sub bands for a 1 $M_\odot$ filter with a time frequency structure given by (9). Most of the filter is low frequency and it is possible to save computation by breaking up the filter into bands each of which has restricted frequency content.

Sub band	Sample rate (Hz)	Interval (s)
1	2048	[0.00093, 0.0059]
2	1024	[0.0059, 0.037]
3	512	[0.037, 0.24]
4	256	[0.24, 1.51]
5	128	[1.51, 9.60]
6	64	[9.60, 60.95]
7	32	[60.95, 387.0]

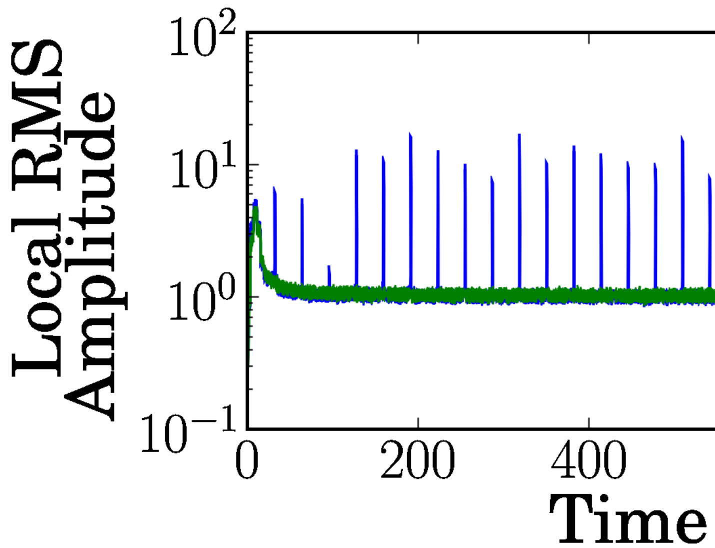
F. Comparison of computational costs and latency

Convolutions are best done with the use of the Fast Fourier Transform (FFT). However there is an inherent latency in performing convolutions this way. This section describes the computational cost and latency of FFT convolutions, time domain (TD) convolutions and TD convolutions with SVD.

Convolving one filter with data using FFTs involves $\mathcal{O}[10 N_s \log_2(N_s)]$ floating point operations where N_s is the number of sample points in the data being convolved and results in $N_s - N_f$ samples of usable output. In comparison time domain convolutions require $\mathcal{O}[N_s N_f]$ where N_f is the number of sample points in the filter. The FFT is faster to compute when $N_f > 10 \log_2(N_s)$. However the FFT has an inherent latency of $\mathcal{O}[N_s/f_s]$ seconds, where f_s is the sample rate when minimally overlapping FFTs are computed. It is possible to reduce the latency of the FFT by performing vastly overlapping FFTs. The cost of such a procedure would be $\mathcal{O}[10 N_o N_s \log_2(N_s)]$. Thus for comparable computational cost where $N_f \approx 10 N_o \log_2(N_s)$ the latency of the FFT method would be $\mathcal{O}[N_s/f_s/N_o]$ seconds. Depending on the task at hand, latency requirements and computational costs can dictate what procedure to use.

G. Data Whitening

Matched filtering the SVD basis vectors has motivated a decoupling of the whitening routine from the matched filtering engine. This was necessary in order weight templates appropriately by whitening them before the SVD calculation, and, as such, we have also moved the data whitening outside the match filter engine. We have chosen to whiten the data using a running geometric average of the PSD computed by Hann windowing 8 second buffers of data and using 50% overlapping buffers. The running geometric average is updated once per buffer from a median of the recent PSDs. We find that this algorithm is extremely fast to converge to an accurate PSD and also very robust against glitches. This is demonstrated below where we have injected band-limited white noise burst with a central frequency of 150 Hz, a bandwidth of 100 Hz, a time duration of 0.2 seconds, and a injection density of one injection every $100/\pi \approx 32$ seconds. At this density we find there should only ever be one injection affecting the median history so the glitches should minimally affect PSD estimation as can be seen by comparing the local RMS amplitude between injections to that without injections.



An inspiral, to leading order, has gravitational-wave frequency evolution given by

$$f(t) = \frac{1}{8\pi\mathcal{M}} \left(\frac{t_c - t}{5\mathcal{M}} \right)^{-3/8} , \qquad (10)$$

whose time derivative is given by

$$\frac{df}{dt} = \frac{3}{320\pi\mathcal{M}^2} \left(\frac{t_c - t}{5\mathcal{M}} \right)^{-11/8} . \qquad (11)$$

Combining these, we find the frequency evolution as a function of frequency to be

$$\frac{df}{dt}(f_0) = \frac{3}{320\pi\mathcal{M}^2} \left(8\pi\mathcal{M} f_0\right)^{3/11} , \qquad (12)$$

which can be inverted to get the minimum frequency which has a given frequency derivative,

$$f_0 = \frac{1}{8\pi\mathcal{M}} \left(\frac{320\pi\mathcal{M}^2}{3} \frac{df}{dt} \right)^{11/3} . \qquad (13)$$

If we allow frequency bins to be affected by an injection only once within the median history, we need to calculate the corresponding df/dt for our PSD calculation. An FFT of a buffer with length T will have a frequency resolution $df = 1/(2T)$. Hann windowing the data and overlapping buffers by 50% introduces correlations in neighboring frequency bins of the FFT. This means that we actually want the frequency to change by $df = 3/(2T)$ before the next PSD calculation, which happens a $dt = T/2$ later, resulting in a minimum $df/dt = 3/T^2$.

Combining these results we find the lower frequency bound for our injections, assuming a chirp mass for a binary with $m_1 = m_2 = 1M_\odot$ and $T = 8$ s, is 23 Hz.

We want to compute the equivalent injection density which would be biased as much as lalapps_inspiral's PSD estimator, which allows ~ 3 injections per 2048 seconds. It computes the PSD by breaking the segment up into 16 256 second chunks. These chunks are then combined into to sets of 8 chunks each. The median of each set is then averaged across the sets for each frequency bin to produce the PSD estimate for that 2048 second segment. This procedure results in an average injection density of 1.5 per 8 PSDs in the median history, which is comparable to LLOID's 1 per 7 PSDs in the median history. Since LLOID has an injection history of 7 PSDs which span 32 seconds, this means LLOID can perform one injection every 32 seconds, or 64 injections per 2048 second segment, more than an order of magnitude increase in density over lalapps_inspiral.

III. IMPLEMENTATION AND ANALYSIS

IV. RESULTS

V. CONCLUSIONS

Acknowledgments

LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0107417. This paper has LIGO Document Number LIGO-P0900004-v1

[1] R. K. Kopparapu, C. Hanna, V. Kalogera, R. O'Shaughnessy, G. Gonzalez, P. R. Brady, and S. Fairhurst, The Astrophysical Journal **675**, 1459 (2008), http://www.journals.uchicago.edu/doi/pdf/10.1086/527348, URL `http://www.journals.uchicago.edu/doi/abs/10.1086/527348`.

[2] B. Abbott et al. (LIGO Scientific), Phys. Rev. **D73**, 102002 (2006), URL `http://arxiv.org/abs/gr-qc/0512078`.

[3] F. Beauville, M.-A. Bizouard, L. Blackburn, L. Bosi, L. Brocco, D. Brown, D. Buskulic, F. Cavalier, S. Chatterji, N. Christensen, et al., Class. Quant. Grav. **25**, 045001 (2008), http://www.iop.org/EJ/abstract/0264-9381/25/4/045001/, URL `arXiv:gr-qc/0701027v1`.

[4] F.Beauville, M.-A.Bizouard, L.Blackburn, L.Bosi, P.Brady, L.Brocco, D.Brown, D.Buskulic, F.Cavalier, S.Chatterji, et al., J. Phys. Conf. Ser. **32**, 212 (2006), URL `arXiv:gr-qc/0509041v1`.

[5] L.Blackburn, F.Beauville, M.-A.Bizouard, L.Bosi, P.Brady, L.Brocco, D.Brown, D.Buskulic, S.Chatterji, N.Christensen, et al., Class. Quant. Grav. **22**, S1149 (2005), URL `arXiv:gr-qc/0504050v1`.

[6] B. Abbott et al. (LIGO Scientific Collaboration), Nucl. Instrum. Methods **A517**, 154 (2004).

[7] L. S. Collaboration and V. Collaboration, Class. Quant. Grav. **25**, 114051 (2008), http://www.iop.org/EJ/abstract/0264-9381/25/11/114051/, URL `arXiv:0802.4320v1[gr-qc]`.

[8] L. S. Collaboration and K. Hurley, The Astrophysical Journal **681**, 1419 (2008), URL `http://arxiv.org/abs/0711.1163`.

[9] J. Sylvestre, Astrophys. J. **591**, 1152 (2003), URL `arXiv:astro-ph/0303512v1`.

[10] J. Markowitz, M. Zanolin, L. Cadonati, and E. Katsavounidis (2008), URL `arXiv:0810.2264v1[gr-qc]`.

[11] V. Raymond, M. van der Sluys, I. Mandel, V. Kalogera, C. Roever, and N. Christensen (2008), URL `arXiv:0812.4302v1[gr-qc]`.

[12] F. C. (LAL), M. B. (LAL), M.-A. B. (LAL), V. B. (LAL),

A.-C. C. (LAL), M. D. (LAL), P. H. (LAL), S. K. (LAL), N. L. (LAL), and M. V. (LAL), Phys. Rev. D **25**, 082004 (2006), URL arXiv:gr-qc/0609118v2.
[13] J. Kanner, T. L. Huard, S. Marka, D. C. Murphy, J. Piscionere, M. Reed, and P. Shawhan (2008), URL arXiv:0803.0312v4[astro-ph].
[14] B. J. Owen, Phys. Rev. D **53**, 6749 (1996).
[15] B. J. Owen and B. S. Sathyaprakash, Phys. Rev. D **60**, 022002 (1999).
[16] B. Abbott et al. (LIGO Scientific), Phys. Rev. **D77**, 062002 (2008), URL http://arxiv.org/abs/0704.3368.
[17] P. R. Brady and S. Ray-Majumder, Classical and Quantum Gravity **21**, S1839 (2004), URL arXiv:gr-qc/0405036.
[18] I. S. Heng (2008), URL arXiv:0810.5707v1[gr-qc].
[19] L. Wen, Int. J. Mod. Phys. D **17**, 1095 (2008), URL http://arxiv.org/abs/gr-qc/0702096v2.
[20] B. A. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton (2005), gr-qc/0509116.
[21] L. A. Wainstein and V. D. Zubakov, *Extraction of signals from noise* (Prentice-Hall, Englewood Cliffs, NJ, 1962).
[22] E. Chassande-Mottin and A. Pai, Phys. Rev. D **73**, 042003 (2006), URL http://arxiv.org/abs/gr-qc/0512137.
[23] E. J. Cands, P. R. Charlton, and H. Helgason, Class. Quant. Grav **25**, 184020 (2008), URL http://arxiv.org/abs/0806.4417.
[24] A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **67**, 024016 (2003), erratum-ibid. 74 (2006) 029903(E).
[25] W. G. Anderson, P. R. Brady, J. D. E. Creighton, and E. E. Flanagan, Phys. Rev. **D63**, 042003 (2001), gr-qc/0008066.
[26] L. Finn and D. Chernoff, Phys. Rev. D **47**, 2198 (1993).
[27] K. L. E, W. C. M, and W. A. G, Class. Quantum Grav. **9**, L125 (1992).
[28] L. Blanchet, Phys. Rev. D **65** (2002).
[29] C. Hanna, M. Megevand, E. Ochsner, and C. Palenzuela, Class. Quantum Grav. **26**, 015009 (2009).

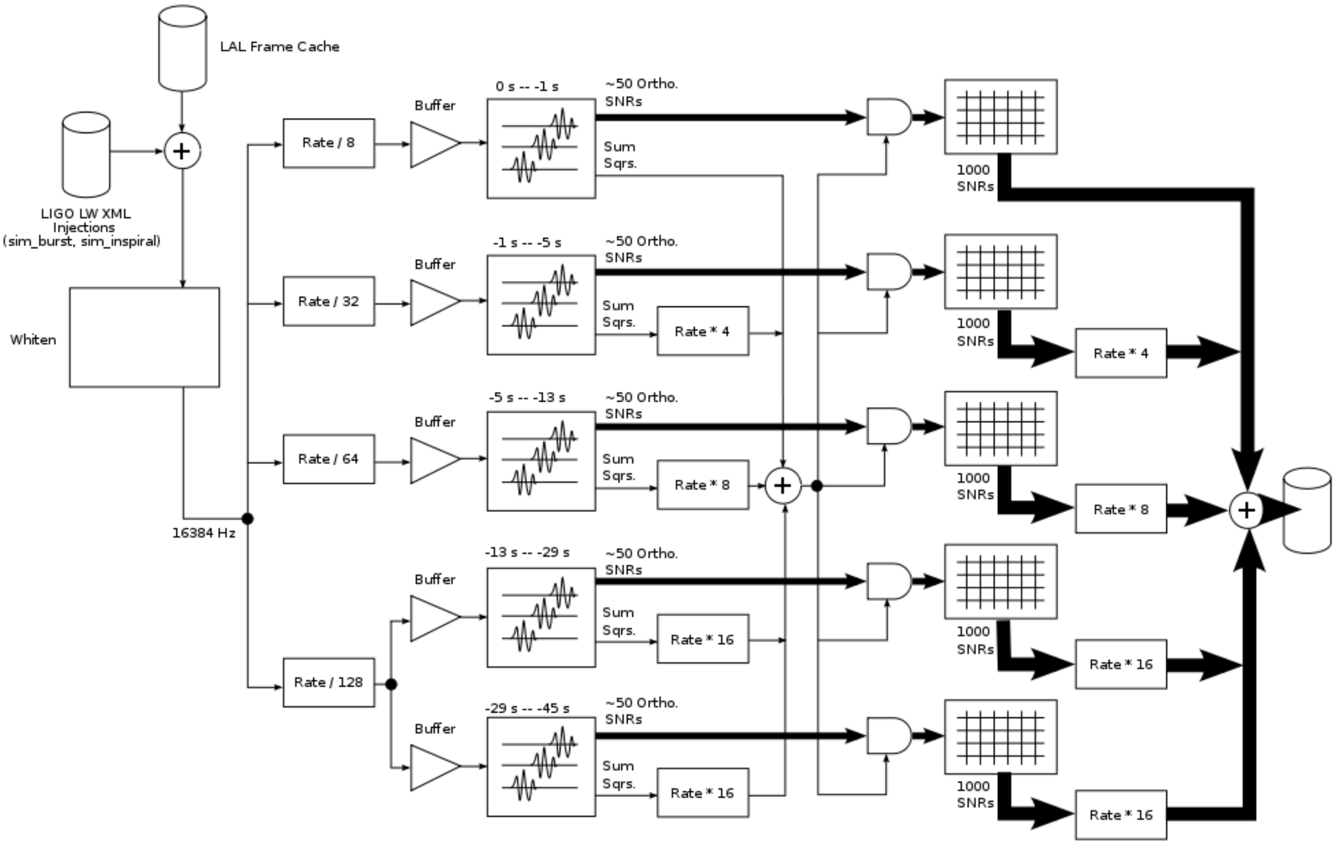

LAL Frame Cache
LIGO LW XML
Injections
(sim_burst, sim_inspiral)
Whiten
16384 Hz
Rate / 8
Rate / 32
Rate / 64
Rate / 128
Buffer
Buffer
Buffer
Buffer
Buffer
0 s -- -1 s
-1 s -- -5 s
-5 s -- -13 s
-13 s -- -29 s
-29 s -- -45 s
~50 Ortho.
SNRs
Sum
Sqrs.
Rate * 4
Rate * 8
Rate * 16
Rate * 16
1000
SNRs
Rate * 4
Rate * 8
Rate * 16
Rate * 16

FIG. 1