Efficient gravitational wave searches from compact binaries using a random projection based template-factorization

Anand S. Sengupta, Amit Reza and Anirban Dasgupta

Indian Institute of Technology, Gandhinagar, Gujarat, India 382355.

MG15 registration ID - - { **937SE1441AR** }

Abstract

Compact-binary coalescences are an important class of sources for the advanced-LIGO detectors. Improved sensitivity at low frequencies and increased bandwidth of the advanced-LIGO and Virgo detectors has resulted in a vast increase in the number of templates to cover the deemed parameter space covering target mass and spin ranges. Further, it has also led to longer template waveforms used for matched-filtering based searches. These in-turns amplify the computational cost of the gravitational wave searches by several orders of magnitude. Random matrix factorization is an efficient technique for the low-rank approximation of a given large template matrix by projecting the template waveforms into a much lower dimensional space. We report the application of such a technique to reduce the high computational costs of adv-LIGO CBC searches. We demonstrate that one can efficiently factorize large template matrices over a distributed memory computer architecture and overcome time and space complexity problem associated with the standard techniques. We also show how this method can potentially benefit existing low-latency searches and also lead to a hierarchical search paradigm using multi-processor and multi-GPU systems. Numerical experiments indicate that the reconstructed detection statistic in the proposed new algorithm is as accurate as the standard truncated-SVD searches within a fixed probabilistic error bound, and is afforded at a fraction of the current computational cost.

General relativity (GR), proposed by Albert Einstein is a geometric theory of gravity, where, the effects of gravitational force is manifested through the curvature of space-time. Not only has it been feted for the beauty and deep symmetries of its mathematical structure but also, has been vindicated by numerous observations e.g. perihelion advance of Mercury's orbit around the sun, the deflection of light near strong fields, etc. More recently, the twin LIGO gravitational-wave observatories recorded the gravitational wave signals from a distant pair of stellar-mass blackholes at a distance of 130 crore light years from Earth, as they inspiraled and ultimately plunged into each other to form a final blackhole. The direct detection of gravitational waves (a key prediction of GR) [7], has opened up a new observational window to the Physical Universe and we are now in the era of Gravitational Wave Astronomy.

One of the great ambitions of this nascent area of astronomy is to follow-up a gravitational wave trigger by an optical telescope (in the high-energy X-Ray or γ-ray band of the electromagnetic spectrum). A coincidental optical follow-up will not only provide the ultimate test for some of the detailed physical models of the scenarios under which the final burst of gravitational waves from the merger of neutron-stars leads to a flash of light, but a detailed measurement of the energies will further help to unravel the mysteries of ultra-compact object in the Universe. We are witness to a unique epoch in the history of astronomy where a brand new branch is taking shape, and where the scientific community has been recharged to seek results cutting across multiple disciplines in science and engineering in the pursuit of suitable solutions.

The astrophysical gravitational wave signals embedded in the noisy data streams are very faint, and special data analysis techniques are used to infer their presence using the broad framework of statistical signal analysis. Making general but reasonable assumptions about the statistical properties of the noise in the detector, one of the techniques (known as templated matched filtering) involve projecting the data over the signal manifold and recording the "triggers" which cross a threshold, corresponding to a fixed pre-set false-alarm rate (FAR). The art of signal processing lies in tuning data analysis pipeline parameters which optimize the detection probability at this FAR.

One can induce a metric on this manifold using which, it is possible to place a discrete grid over it at some pre-set minimal match. The construction of the grid over the range of search parameters itself is an active field of research within the community, and is more commonly referred to as the problem of "template-placement". The aim is to place the fewest templates at a fixed granularity with a view of directly reducing the overall computational cost of the search. Ideas from the well known sphere-packing problems can be used to optimize the grid layout but non-vanishing curvature effects, coupled with boundary effects often reduce the efficacy of such techniques. In a recent series of work, Cannon et. al. [1 - 4] have shown that the vectors in the signal manifold (corresponding to each of the grid points) do not form a linearly independent set. As such, they could be represented effectively as a linear combination of basis waveforms using the well studied singular value decomposition (SVD) method. The authors have also shown that using a truncated SVD, one could effectively speed up the analysis by an order of magnitude with high reconstruction accuracy where one filters the data using the most important basis vectors of the signal matrix, and reconstructs the filter output for the templates using the SVD components. This was perhaps the first work in GW signal analysis where ideas of dimensionality reduction from big-data analytics were used to meet the computational challenges of this upcoming area of astronomy.

One of the major challenges in the proposed truncated SVD approach mentioned above is that it is not a scalable solution. In the advanced detector era, improvements in detector sensitivity at low frequencies and availability of fully precessing spinning theoretical template waveforms from compact binary coalescences have resulted in an explosion of the volume and number of templates over which the search is carried out. In the O1 run, some 250,000 templates were used. One expects this number to increase several times more as advanced detectors are commissioned. Considering each template to nominally have 106 time samples, one is now faced with the gigantic task of (a) performing SVD on a 105 x 106 sized matrix of double-precision numbers and (b) logistics of using a large number of basis vectors to actually filter the data and (c) reconstruct the filter output time series. A possible solution proposed by the authors [3, 6] is to break up the parameter space into smaller sub-regions and perform an independent search on each of these smaller parts. However, in this approach there are two potential problems: (a) one loses the computational advantage afforded by the linear independence of basis vector over the entire space and (b) it may also lead to sub-optimal searches due to inability to perform signal consistency vetoes on the filter output by looking over a small subdomain of the search space.

In this work, we revisit the problem of dimensionality reduction using efficient low-rank factorizations of the entire template matrix using recently developed randomized linear-algebra algorithms. Random Projections (RP), conceived by the pioneering work of Johnson and Lindenstrauss [9], is a computationally efficient technique for dimension reduction and finds applications in many areas of data science. Improvements in low-frequency response and bandwidth due to detector hardware upgrades pose a data analysis challenge in the advanced LIGO era as they result in increased redundancy in template databases and longer templates due to a higher number of signal cycles in the sensitive band. We apply this technique to address both these above-mentioned data analysis challenges in future CBC searches, namely, handling redundancies in large template databases, and efficiently correlating noisy data against long templates.

The primary impact of this work is multi-fold: (1) Efficient template matrix factorization can be used to address the redundancy problem. This is similar in spirit to the SVD factorization that is at the heart of the "GstLAL"-based inspiral pipeline [10, 3, 6]. The RP based method presented here scale well for a very large number of templates embedded in high dimensional Euclidean space. In many scenarios, such factorizations can be done off-line, in advance of a CBC search. Nonetheless, there can be situations when the factors need to be updated on-line owing to the non-stationarity of data. Our adaptation of the method for CBC searches will benefit both scenarios. (2) We show the explicit connection between the new factorization scheme and the extant SVD method. This would not only bridge the two approaches but also make the new scheme readily usable. (3) The computational challenges arising from correlating noisy data against long templates (also known as the curse of dimensionality) is addressed by casting the matched-filtering operation itself in a lower-dimensional space. For certain types of template banks (e.g., for CBCs with precessing spins), the template matrix may be less amenable to an SVD-like factorization. In such cases, the total computational cost can be significantly reduced by using the RP-based correlation alone. (4) Finally, we show that RP-based template matrix factorization, on the one hand, and matched-filtering computation in reduced dimension, on the other, can be combined effectively for efficient CBC searches.

We demonstrate that the new techniques from the area of computer science and engineering and big-data analytics provide a scalable and computationally efficient solution to the signal analysis problem in the domain of gravitational wave astronomy. We not only try to theoretically expound the ramifications of using these new ideas in LIGO's marquee searches for gravitational waves from compact binaries but also indicate their deployment of these techniques within the low-latency search analysis pipeline [8] on large compute clusters. The use of randomized numerical linear algebra to the domain of gravitational wave data analysis would lead to significant computational advantages (order of magnitude) over existing analysis techniques - opening up further opportunities of discovery in this new branch of astronomy.

References

1. *Random projections in gravitational wave searches of compact binaries*, Sumeet Kulkarni, Khun Sang Phukon, Amit Reza, Sukanta Bose, Anirban Dasgupta, Dilip Krishnaswamy, Anand S. Sengupta; arXiv:1801.04506

2. Kipp Cannon, Adrian Chapman, Chad Hanna, Drew Keppel, Antony C. Searle, and Alan J. Weinstein, *Singular value decomposition applied to compact binary coalescence gravitational-wave signals*, Phys.Rev.D82:044025, 2010.

3. K Cannon et al., May. 2014, *Toward early-warning detection of gravitational waves from compact binary coalescence*, Astrophys. J. 748, 136 (2012), arXiv:1107.2665.

4. Kipp Cannon, Chad Hanna, Drew Keppel, *Efficiently enclosing the compact binary parameter space by singular-value decomposition*, Phys.Rev.D84:084003, 2011.

5. R. J. E. Smith, K. Cannon, C. Hanna, D. Keppel and I. Mandel, *Towards rapid parameter estimation on gravitational waves from compact binaries using interpolated waveforms*, Phys.Rev.D87: 122002, 2013

6. C Messick et al., Jan 2017, *Analysis Framework for the prompt discovery of compact binary mergers in gravitational wave data*, Phys. Rev. D95, 042001 (2017), arXiv:1604.04324v3.

7. *Observation of gravitational waves from a binary black hole merger*, BP Abbott, R Abbott, TD Abbott, MR Abernathy, F Acernese, K Ackley, Physical Review Letters 116 (6), 061102

8. `gstLAL` *library (software suite for GStreamer signal-processing pipelines)*: https://wiki.ligo.org/DASWG/GstLAL [Project librarians: Kip Cannon and Chad Hanna]

9. W. B. Johnson and J. Lindenstrauss, in Conference in modern analysis and probability (New Haven, Conn., 1982), Contemp. Math., Vol. 26 (Amer. Math. Soc., Providence, RI, 1984) pp. 189–206.

10. S. Privitera, S. R. P. Mohapatra, P. Ajith, K. Cannon, N. Fotopoulos, M. A. Frei, C. Hanna, A. J. Weinstein, and J. T. Whelan, Phys. Rev. D89, 024003 (2014), arXiv:1310.5633 [gr-qc].