arXiv:1201.1739v1 [physics.data-an] 9 Jan 2012

# Use of Singular-Value Decomposition in Gravitational-Wave Data Analysis

Drew Keppel

*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik*
*D-30167 Hannover, Germany*
*Leibniz Universität Hannover*
*D-30167 Hannover, Germany*

*drew.keppel@ligo.org*

Singular-value decomposition is a powerful technique that has been used in the analysis of matrices in many fields. In this paper, we summarize how it has been applied to the analysis of gravitational-wave data. These include producing basis waveforms for matched filtering, decreasing the computational cost of searching for many waveforms, improving parameter estimation, and providing a method of waveform interpolation.

## 1. Introduction

Within the field of gravitational-wave (GW) data analysis, there are several difficult problems that need to be addressed in order to extract information from the data that will be taken by advanced GW detectors. These include determining when a GW signal occurred, where this signal came from, and what the properties of the signal's source are. In this paper, we summarize how singular value decomposition (SVD) has been used to ease the computational burden of answering these questions.

This paper is outlined as follows: Sect. 2 gives a brief overview of SVD, Sect. 3 describes how SVD can be used to approximate different classes of waveforms, Sect. 4 illustrates how SVD has be used to construct composite detection statistics, Sect. 5 discusses a method of waveform interpolation based on SVD, and Sect. 6 summarizes the use of SVD in different aspects of parameter estimation.

## 2. Singular Value Decomposition

SVD decomposes an arbitrary matrix $\mathbf{H}$ into three components:

$$\mathbf{H} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}, \tag{1}$$

where $\mathbf{U}$ is a unitary matrix of basis vectors such that the columns of $\mathbf{U}^*$ are the right eigenvectors of $\mathbf{H}^*\mathbf{H}$, $\mathbf{V}$ is a unitary matrix of reconstruction coefficients such that the columns of $\mathbf{V}$ are the left eigenvectors of $\mathbf{H}\mathbf{H}^*$, and $\mathbf{\Sigma}$ is a diagonal matrix of singular values whose non-zero components are the square root of the eigenvalues associated with $\mathbf{H}^*\mathbf{H}$ or $\mathbf{H}\mathbf{H}^*$.[a] The eigenvectors of $\mathbf{U}$ and $\mathbf{V}$ are ordered such that

---

[a]Throughout this paper, $\mathbf{M}^*$ denotes the complex conjugate-transpose of matrix $\mathbf{M}$.

the entries of $\mathbf{\Sigma}$ are in descending order.

We will see in the following sections that the properties of $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{\Sigma}$ have been used in different contexts.

## 3. Classes of Waveforms

GW signal waveforms can be classified into two groups: those that are well modelled, and those that are not.[b] Collections of waveforms of both types can be called *waveform catalogs*, however here we restrict use of that term for non-well-modelled waveforms. We will refer to waveforms that are well modelled as *template banks*. Further discussion of temaplte banks and waveform catalogs can be found in Sect. 3.1 and 3.2, respectively.

### 3.1. *Waveform Template Banks*

An example of a type of waveform that can be used to construct a template bank is a GW signal from the inspiral of a compact binary coalescence (CBC). These waveforms are well described theoretically by the post-Newtonian (PN) approximation. This has allowed the construction of a metric on the signal parameter space that characterizes the distance between two points on parameter space based on the mismatch between the two waveforms.[1] Using this metric, a template bank can be constructed with these waveforms such that any point within a chosen region of parameter space will be no more than a specified distance from the nearest template, commonly chosen such that maximum mismatch between any template in the bank and any signal in the region of interest is 3%.

In order to meet this requirement with the fewest number of templates, for a two dimensional parameter space with a constant metric, adjacent templates will have a mismatch of 9%, which means they have an overlap of 91%. Cannon *et al.* investigated how the SVD could compress a matrix $\mathbf{H}$ whose rows are composed of the highly overlapping waveform time-series from such a template bank.[2] They derived the expected fractional signal-to-noise ratio (SNR) loss $\delta\rho/\rho$ of approximating the waveforms in the template bank by the truncation of the SVD of $\mathbf{H}$ to be

$$\frac{\delta\rho}{\rho} = \frac{1}{2N}\sum_{i=N'+1}^{N}\sigma_i^2, \tag{2}$$

where $\sigma_i$ is the $i$th entry of $\mathbf{\Sigma}$, $N$ is the number of rows of $\mathbf{H}$ or twice the number of templates in the template bank, and $N'$ is where the SVD is truncated. In addition, the maximum phase-averaged fractional SNR loss can be bounded by

$$\left(\frac{\delta\rho}{\rho}\right)_{\max} < \frac{1}{2}\sigma_{N'+1}^2, \tag{3}$$

---

[b]One reason a waveform may not be able to be well modelled is because of complex and possibly chaotic physical processes at work in the production of the GW signal (*e.g.*, the presence of matter in a GW generating system).

which can be derived from the unitarity of the reconstruction coefficients.

This approach to waveform compression has been implemented within a low-latency GW search for inspiral waveforms as one of the techniques to reduce the computational cost of data processing.[3] Similar techniques have also been applied to other types of GW signals using Gram-Schmidt orthogonalization.[4,5]

### 3.2. *Waveform Catalogs*

Waveform catalogs are necessary for waveforms that are not well modelled theoretically. Such waveforms can arise from numerical simulations that are either too computationally expensive to study the parameter space thoroughly (*e.g.*, numerical simulations of merging spinning binary black holes), or from numerical simulations that include matter (*e.g.*, numerical simulations of supernovae (SN) or coalescing binary neutron stars).

Ref. 6 first proposed decomposing waveform catalogs of SN waveforms using Gram-Schmidt orthogonalization. Ref. 7 expanded on this approach using the SVD and comparing it to Gram-Schmidt orthogonalization. Both approaches were found to perform similarly, being able to extract and prioritize waveform characteristics from the SN waveform catalogs. However, in the SVD only a single decomposition of the catalog was needed, and additional information was obtained. Namely, features that were present in many waveforms produced basis vectors with large singular values, whereas features present in only a few produce basis vectors with small singular values.

## 4. Detection Statistics

Constructing detection statistics when searching for GW signals is different depending on what type of search you are doing. Searches can be classified into two groups, those searching for known signal waveforms and those searching for unknown signal waveforms.

The standard approach to search for any known signal within GW data from a collection of GW waveforms is to matched filter the data, producing a time-series of SNR, with each of the template waveforms. GW signals are then identified by identifying which templates at certain times produce large values of SNR.[8]

If one is only interested in answering the time component of the question, a *composite detection statistic* can be used, which tells whether the data looks like anything from a collection of template waveforms.[8] Cannon *et al.* showed how the SVD can be used to construct such a composite detection statistic $\Gamma$ associated with a region of parameter space covered by a template bank of the form

$$\Gamma = \sum_{k=1}^{N'} \frac{\sigma_k^2}{\sigma_k^2 + 2N/\langle A^2 \rangle} \rho_k^2, \tag{4}$$

where $\langle A^2 \rangle$ is the expected signal amplitude squared and $\rho_k$ is the SNR associated with the $k$th basis vector.[9] This statistic was derived using the orthonormality of the

reconstruction coefficients and the basis vectors, as well as the ordered significance given by the singular values. It shows promise in reducing the computational cost of searching GW data for signals of known waveforms.

When searching for a signal of unknown waveform, the search can be performed by looking at other aspects of how the signal interacts with the detectors. In particular, when there is a network of GW detectors in operation, the data from these detectors can be combined in such a way to produce coherent data streams that are direction dependent.[10,11]

Wen outlined how the detector responses associated with a given direction can be coherently combined and condensed using the SVD.[12] The SVD of the network detector responses $\mathbf{A}$ provides the basis vectors $\mathbf{U}$, singular values $\mathbf{\Sigma}$, and reconstruction coefficients $\mathbf{V}$. The data from the different detectors $\mathbf{d}$ can be recombined into new data streams $\mathbf{d}'$ using the reconstruction coefficients

$$\mathbf{d}' = \mathbf{V}^{-1}\mathbf{d}. \tag{5}$$

The detection statistics for an unmodelled GW search could then be constructed using

$$\Gamma = \sum_k w_k \left| d_k' \right|^2, \tag{6}$$

where $w_k$ is a chosen weight associated with different basis vectors, and the sum can be over different combinations of the new data streams. Different choices for each of these parameters results in different optimizations of the detection statistic.

## 5. Waveform Interpolation

Producing the waveforms in the first place can be a computational challenge, which is the case for binary black hole coalescence waveforms produced by numerical relativity. To help solve this problem, there have been several efforts to parameterize these waveforms, allowing waveforms from arbitrary points in parameter space to be produced.[c] These efforts are based on models of the waveforms that need to be tuned. Cannon *et al.* have pursued a different approach by showing that the basis vectors from the SVD of waveforms from a template bank approximately enclose the space of waveforms from that region of parameter space.[13] Further, they have also laid out a procedure that could be used to interpolate waveforms to high accuracy from arbitrary points in parameter space using the components of the SVD.[14] This is done by interpolating the reconstruction coefficients. Once the interpolated reconstruction coefficients $\mathbf{V}'$ are in hand, the interpolated waveforms can be generated using

$$\mathbf{H}' = \mathbf{V}'\mathbf{\Sigma}\mathbf{U}, \tag{7}$$

---

[c]References for these efforts can be found in Ref. 14.

where $\mathbf{H}'$ is a matrix containing interpolated waveforms at the locations where $\mathbf{V}'$ was interpolated. With a sufficient density of numerical relativity waveforms, highly accurate SVD interpolated waveforms could be produced for arbitrary points in parameter space.

## 6. Parameter Estimation

Once it is known that there exists a GW signal in a data set, the next question that comes up is what are the parameters of the source of the signal. Several sets of authors have used the SVD toward this goal in different contexts.

Wen identified how the SVD could be used to construct *null data streams* and *semi-null data streams*.[12] Null data streams are linear combinations of data from a network of GW detectors that will contain no GW signals,[15] which are associated with zero-valued singular values from the SVD of the response of a network of GW detectors. Similarly, semi-null data streams are associated with small-valued singular values from the SVD of the response of a network of GW detectors. Wen *et al.* showed how analyzing these semi-null data streams could improve the accuracy with which one could locate the source of a GW signal on the sky.[16]

Wen also explained how the SVD could be used to extract a GW signal waveform from the data of a network of GW detectors.[12] This was done by truncating the SVD of the detector response matrix in order to regularize it. The extracted waveform $\mathbf{h}^{\mathrm{T}}$ is then the data from all detectors $\mathbf{d}$ combined using the inverse of the regularized detectors response matrix,

$$\mathbf{h}^{\mathrm{T}} = \left(\mathbf{A}^{\mathrm{T}}\right)^{-1}\mathbf{d}, \tag{8}$$

where $A_{ij}^{\mathrm{T}} = \sum_{k=1}^{N'} v_{ik}\sigma_k u_{kj}$ is the detectors' response matrix regularized by truncating the SVD after the $N'$th singular value.

Röver *et al.* applied the SVD of waveform catalogs to the parameter estimation problem.[17] Using the basis vectors associated with the SVD of a SN waveform catalog, a parameter estimation search based on Markov chain Monte Carlo (MCMC) methods extracted the amplitudes of the signal associated with each of the basis vectors. By associating the extracted amplitudes with those of the waveforms in the catalog, physical parameters of the simulated signal could be estimated.

Extending the previously mentioned work, Logue *et al.* applied the SVD to multiple waveform catalogs, each associated with different SN production mechanisms or different types of stellar cores.[18] By computing the Bayes odd ratios associated with these different models by nested sampling techniques, the correct model for the SN can be determined.

Finally, Cannon *et al.* discussed how the likelihood function, which is the quantity computed for parameter estimation, could be interpolated to decrease the computation cost of parameter estimation searches.[14] Using the interpolated reconstruction coefficients, the interpolated likelihood function would be given by

$$\ln\left(\mathbf{\Lambda}'\right) = \mathbf{V}'\rho^2, \tag{9}$$

6    *Drew Keppel*

where the locations the likelihood function is estimated are given by the locations where the reconstruction coefficients are interpolated.

## 7.  Conclusion

SVD has been applied with good results in a number of contexts within GW data analysis. From waveform compression and interpolation, to deriving new detection statistics, to improving parameter estimation, the powerful properties of the SVD have proved to be of great use.

## Acknowledgments

## References

1.  B. J. Owen, Phys. Rev. D **53**, 6749 (1996), arXiv:gr-qc/9511032
2.  K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, Phys. Rev. D **82**, 044025 (2010), arXiv:1005.0012 [gr-qc]
3.  K. Cannon, R. Cariou, A. Chapman, M. Crispín-Ortuzar, N. Fotopoulos, M. Frei, C. Hanna, E. Kara, D. Keppel, L. Liao, S. Privitera, A. Searle, L. Singer,  and A. Weinstein, Submitted to Astrophysical Journal, arXiv:1107.2664 [astro-ph.IM]
4.  S. E. Field, C. R. Galley, F. Herrmann, J. S. Hesthaven, E. Ochsner,  and M. Tiglio, Phys. Rev. Lett. **106**, 221102 (2011), arXiv:1101.3765 [gr-qc]
5.  S. Caudill, S. E. Field, C. R. Galley, F. Herrmann,  and M. Tiglio, ArXiv e-prints, arXiv:1109.5642 [gr-qc]
6.  P. R. Brady  and S. Ray-Majumder, Classical and Quantum Gravity, **21**, 1839 (2004), arXiv:gr-qc/0405036
7.  I. S. Heng, Classical and Quantum Gravity **26**, 105005 (2009), arXiv:0810.5707 [gr-qc]
8.  L. A. Wainstein, and V. D. Zubakov, *Extraction of signals from noise* (Prentice-Hall, Englewood Cliffs, NJ, 1962)
9.  K. Cannon, C. Hanna, D. Keppel,  and A. C. Searle, Phys. Rev. D **83**, 084053 (2011), arXiv:1101.0584 [physics.data-an]
10. Y. Gürsel  and M. Tinto, Phys. Rev. D **40**, 3884 (1989)
11. É. É. Flanagan  and S. A. Hughes, Phys. Rev. D, **57**, 4566 (1998), arXiv:gr-qc/9710129
12. L. Wen, Int. J. Mod. Phys. D **17**, 1095 (2008), arXiv:gr-qc/0702096
13. K. Cannon, C. Hanna,  and D. Keppel, Phys. Rev. D **84**, 084003 (2011), arXiv:1101.4939 [gr-qc]
14. K. Cannon, C. Hanna,  and D. Keppel, Submitted to Phys. Rev. Lett., arXiv:1108.5618 [gr-qc]
15. L. Wen and B. F. Schutz, Classical and Quantum Gravity **22**, 1321 (2005), arXiv:gr-qc/0508042
16. L. Wen, X. Fan,  and Y. Chen, Journal of Physics Conference Series **122**, 012038 (2008), arXiv:0711.0726 [gr-qc]

17.  C. Röver, M.-A. Bizouard, N. Christensen, H. Dimmelmeier, I. S. Heng,  and R. Meyer, Phys. Rev. D **80**, 102004 (2009), arXiv:0909.1093 [gr-qc]
18.  J. Logue, C. D. Ott, I. S. Heng, P. Kalmus,  and J. H. C. Scargill, Private communication.