Reconstructing the Gravitational Wave Signal of Supernovae

Deborah Good

August 14, 2013

Abstract

One candidate source for detectable gravitational waves is the core-collapse supernova. Unfortunately, no definitive model of a supernova gravitational waveforms has yet been found through either numerical or analytical methods. It is therefore desirable to have methods of detecting gravitational waves from these sources without a priori knowledge of the signal's appearance. Previous work has developed these codes, but has not thoroughly assessed their efficacy or examined methods of improving their performance. This project endeavored to do that, and met with limited success. We examined briefly the effects of retaining greater numbers of clusters of time-frequency pixels, but focused our efforts predominantly on establishing optimal percentages of retained time-frequency pixels, finding approximate optima for percentile in both the single and halo-core cases.

1 Introduction

Advanced LIGO and Advanced VIRGO detectors are expected to come on line as early as next year. Therefore, it is important that programs exist to robustly analyze data from a number of sources. One such source is core-collapse supernovae. This source is unusual in that no complete, reliable model for the waveforms exists. As a result, we must establish effective reconstruction programs which can find these signals without a priori knowledge of their appearance. In this work, we will present analysis of one such reconstruction script and the manners in which it can be improved.

2 Gravitational Wave Bursts & Supernovae

Einstein's theory of general relativity posits the existence of gravitational waves arising from sources with time-dependent mass quadrupole moments. These are transverse waves with two polarizations, plus (+) and cross (×). Additionally, in a linearized theory based on flat spacetime, these waves follow an ordinary wave equation [1].

Some sources of gravitational waves are expected to emit continuous signals, but others are expected to emit GW bursts. These burst sources include supernovae, binary black hole or neutron star mergers, and gamma ray bursts, which are also examples of gravitational collapse. A deeper understanding of these sources could lead to much advancement in relativistic astrophysics, making the study of burst sources a highly desirable undertaking. Additionally, these source are believed to occur at a rate of about 30 per year, increasing the probability of their detection [3], [2].

A supernova occurs when the electron degenerate iron core of a star with a zero age main sequence mass between eight and one hundred times the mass of the sun exceeds its effective Chandrasekhar mass and becomes gravitationally unstable. At this time, the core collapses, meaning the inner core material is compressed to nuclear densities. Then, the nuclear equation of state induces a core bounce, which sends a shock wave out from the core. This shock wave loses energy quickly, though. If it is revived, it will finish leaving the stellar envelope, explode the star in a supernova, and ultimately create a neutron star. If it is not revived, the star will just collapse again and become a black hole. (This process is incredibly energetic, but only 1 % of that energy is emitted as light; the rest is neutrino emission) [1].

The largest question in understanding supernovae is how gravitational energy is transferred to revive the shock wave. At present, there are three main theories (the neutrino mechanism, the magneto-rotational mechanism, and the acoustic mechanism). Of course, observational data is desirable to distinguish between the theories. This cannot easily be obtained, because traditional visual astronomy cannot be used. Therefore, the big questions in supernovae research require gravitational wave observation [1]. Gravitational wave emission is expected from rotating collapse and bounce, non-axisymmetric rotational instabilities, post-bounce convective overturn/SASI and PNS pulsations [1].

For a process as complicated as the core-collapse of a supernovae solving the Einstein equations and the radiation-magnetohydrodynamics equations require complex numerical modeling. A full model of the stellar core collapse and postbounce supernova core evolution would account for all length scales, be three-dimensional, and include full proper treat-

ment of general relativity, magneto-hydrodynamics, neutrino transport, and EOS nuclear and neutrino interactions. Unfortunately, the numerical modeling is so complicated that it is impossible to create a full model with present computational technology. A single full simulation would require thousands of cores on computers better than the ones now existing [1].

There are a number of simplifications which can be introduced to allow the creation of partial models. However, we cannot assume that a real gravitational wave signal would look just like these approximate models. Detection of core-collapse signals is possible (though not especially likely) during advanced LIGO operation, so we need to be able to reconstruct gravitational wave signals from core-collapse supernovae in the absence of a detailed model of the waveform.

3 Reconstruction Codes

3.1 X-Pipeline Analysis Processes

In order to understand possible improvements to the waveform reconstruction process, we must first understand the mathematics of waveform reconstruction. Therefore, we begin from provided antenna response functions and data streams and derive the reconstruction process.

The output ($\tilde{\mathbf{d}}$) of a gravitational wave detector (α) is given by an equation like Equation 1, where F_α^+ and $F_\alpha^\times$ are antenna response functions and n_α is the noise in the signal.

$$d_\alpha(t + \Delta t_\alpha(\hat{\Omega})) = F_\alpha^+(\hat{\Omega}) + F_\alpha^\times(\hat{\Omega})h_\times(t) + n_\alpha(t + \Delta t_\alpha(\hat{\Omega})) \tag{1}$$

Each individual detector takes little notice of the sky position of the signals it observes. However, the detectors are spread out over the earth to allow for accurate triangulation of sky positions. This separation is visible in the data collected by the detectors as a time delay between each detector. When conducting analysis, we want to be sure we see the same thing at the same time on each detector. Therefore, we introduce a time shift relative to a fixed point $\mathbf{r}_0$, prior to analysis. This shift is given by Equation 2, and it is assumed into further discussion.

$$\Delta t_\alpha(\hat{\Omega}) = \frac{1}{c}(\mathbf{r}_0 - \mathbf{r}_\alpha) \cdot \hat{\Omega} \tag{2}$$

Our actual data is discrete, so we use discrete notation in our analyses. The frequency and time series are related by a Fourier transform.

$$\begin{aligned} \tilde{x}[k] &= \sum_{j=0}^{N-1} x[j] e^{-i2\pi jk/N} \\ x[j] &= \frac{1}{N} \sum_{k=0}^{N-1} \tilde{x}[k] e^{i2\pi jk/N} \end{aligned} \tag{3}$$

For our own convenience, we also define our signal, our noise, and our antenna response figures as noise-weighted quantities

$$\begin{aligned}\tilde{d}_{w\alpha}[k] &= \frac{\tilde{d}_{\alpha}[k]}{\sqrt{\frac{N}{2}S_{\alpha}[k]}} \\ \tilde{n}_{w\alpha}[k] &= \frac{\tilde{n}_{\alpha}[k]}{\sqrt{\frac{N}{2}S_{\alpha}[k]}} \\ \tilde{F}^{+,\times}_{w\alpha} &= \frac{F^{+,\times}_{\alpha}(\hat{\Omega})}{\sqrt{\frac{N}{2}S_{\alpha}[k]}}\end{aligned} \tag{4}$$

These re-definitions allow us to re-write Equation 1 in a much more manageable format as

$$\tilde{\mathbf{d}} = \mathbf{F}\tilde{\mathbf{h}} + \tilde{\mathbf{n}} \tag{5}$$

where the data, antenna function, and noise are noise-weighted as in Equation 4 with an element for each detector in use and where $\tilde{\mathbf{h}}$ is the gravitational wave and is a vector with one element for the + polarization and one element for the × polarization.

Once we are in possession of equations for detector output, we need to establish likelihood ratios.

We begin this process by coming up with an expression for the probability of finding some gravitational wave $\tilde{\mathbf{h}}$ in some stream of data $\tilde{\mathbf{h}}$. To determine this probability, we assume (somewhat erroneously) a Gaussian noise distribution. Therefore, for one detector, the probability is given by Equation 6 and for a matrix $\mathbf{n}$ of D detectors by Equation 7.

$$P(n) = \frac{1}{\sqrt{2\pi}} exp[\frac{-|n|^2}{2}] \tag{6}$$

$$P(n) = \frac{1}{\sqrt{2\pi}^D} exp[\frac{-n^T n}{2}] \tag{7}$$

Solving Equation 5 for $\mathbf{n}$ yields $\mathbf{n} = \tilde{\mathbf{d}} - \mathbf{F}\tilde{\mathbf{h}}$. Substituting this into Equation 7 gives us the probability of finding some signal $\tilde{\mathbf{h}}$ in some data stream $\tilde{\mathbf{d}}$, as shown in Equation 8.

$$P(\tilde{\mathbf{d}}|\tilde{\mathbf{h}}) = \frac{1}{\sqrt{2\pi}^D}\ exp\left[\frac{-(\tilde{\mathbf{d}} - \mathbf{F}\tilde{\mathbf{h}})^T(\tilde{\mathbf{d}} - \mathbf{F}\tilde{\mathbf{h}})}{2}\right] \tag{8}$$

Equation 8 becomes

$$P(\{\tilde{\mathbf{d}}\}|\{\tilde{\mathbf{h}}\}) = \frac{1}{\sqrt{2\pi}^{Dn_p}} exp\left[\frac{-1}{2}\sum_k \left(\tilde{\mathbf{d}}[k] - \mathbf{F}[k]\tilde{\mathbf{h}})[k])^T(\tilde{\mathbf{d}}[k] - \mathbf{F}[k]\tilde{\mathbf{h}}[k])\right)\right] \tag{9}$$

for a set of N_p $\tilde{\mathbf{d}}$ values, $\{\tilde{\mathbf{d}}\}$.

In the case where there is no signal, this reduces to

$$P(n) = \frac{1}{\sqrt{2\pi}^D} exp\left[\frac{-\tilde{\mathbf{d}}^T\tilde{\mathbf{d}}}{2}\right] \tag{10}$$

which is identical to the noise only case considered previously.

The likelihood ratio is defined as the log-ratio of the probability of finding something and the probability of finding nothing, as given below by Equation 11:

$$L \equiv ln\frac{P(\{\tilde{\mathbf{d}}\}|\{\tilde{\mathbf{h}}\})}{P(\{\tilde{\mathbf{d}}\}|\{0\})} = \frac{1}{2}\sum_k \left[|\tilde{\mathbf{d}}|^2 - |\tilde{\mathbf{d}} - \mathbf{F}\tilde{\mathbf{h}}|^2\right]. \tag{11}$$

If waveform $\tilde{\mathbf{h}}$ is known in advance, finding L is straightforward. However, the entire problem lies in that we do not know what $\tilde{\mathbf{h}}$ is. Therefore, we find the best fit waveform to maximize likelihood ratio by solving

$$0 = \frac{\partial L}{\partial \tilde{\mathbf{h}}}|_{\tilde{\mathbf{h}}=\tilde{\mathbf{h}}_{\mathbf{max}}} \tag{12}$$

for $\tilde{\mathbf{h}}_{max}$. The result is a linear equation,

$$\tilde{\mathbf{h}}_{\mathbf{max}} = (\mathbf{F}^\dagger\mathbf{F})^{-1}\,\mathbf{F}^\dagger\,\tilde{\mathbf{d}} \tag{13}$$

where $\mathbf{F}$ is still the antenna response function, $\tilde{\mathbf{d}}$ is still the data stream from the detector, and $\dagger$ is the complex conjugate transpose.

Then, we can use a projection operator $\mathbf{P}^{GW} \equiv \mathbf{F}\left(\mathbf{F}^\dagger\mathbf{F}\right)^{-1}\mathbf{F}^\dagger$ and find the standard likelihood:

$$E_{SL} \equiv 2L\tilde{\mathbf{h}}_{\mathbf{max}} = \sum_k \tilde{\mathbf{d}}^\dagger\mathbf{P}^{GW}\tilde{\mathbf{d}}. \tag{14}$$

Projection operator $\mathbf{P}^{GW}$ projects the data from the detector into the + and $\times$ plane.

There is another projection operator known as the null projection operator, orthogonal to $\mathbf{P}^{GW}$, and given by $\mathbf{P}^{null} \equiv \mathbf{I} - \mathbf{P}^{GW}$ which can cancel out the gravitational wave signal in the data.

Projection operator $\mathbf{P}^{null}$ gives rise to the null energy, which is the minimum amount of energy in the whitened data that does not allow for a gravitational wave at that point. Null energy is

$$E_{null} \equiv E_{tot} - E_{SL} = \sum_k \tilde{\mathbf{d}}^\dagger\,\mathbf{P}^{null}\,\tilde{\mathbf{d}} \tag{15}$$

where E_{SL} is from Equation 14, E_{tot} is $E_{tot} = \sum_k |\tilde{\mathbf{d}}|^2$, and $\tilde{\mathbf{d}}$ is the data stream.

This projection operator analysis can be useful in that E_{null} allows us to remove some portion of the noise without affecting the actual signal.

The default basis for the detector data space (that formed by the single-detector strains) is not the optimal basis for detection statistics. Therefore, we choose to convert to a dominant polarization frame based in the two-dimensional subspace defined by $\mathbf{F}^+$ and $\mathbf{F}^\times$ which allows us to have a direction in which the detectors have maximum antenna response orthogonal to a direction with minimum antenna response.

We begin constructing this dominant polarization frame (DPF) by writing the antenna response functions related by a polarization angle ψ as shown in Equations 16 and 17:

$$\mathbf{F}^+(\psi) = \cos 2\psi\mathbf{F}^+(0) + \sin 2\psi\mathbf{F}^\times(0) \tag{16}$$

$$\mathbf{F}^{\times}(\psi) = -\sin 2\psi \mathbf{F}^{+}(0) + \cos 2\psi \mathbf{F}^{\times}(0). \tag{17}$$

We then rotate the polarization frame with the angle

$$\psi_{DP}(\hat{\Omega}, k) = \frac{1}{4} atan2(2\mathbf{F}^{+}(0) \cdot \mathbf{F}^{\times}(0), |\mathbf{F}^{+}(0)|^2 - |\mathbf{F}^{\times}(0)|^2 \tag{18}$$

where atan2(y,x) is the arctangent over $(\pi, \pi]$.

We now can write unit vectors $\mathbf{e}^{+} \equiv \frac{\mathbf{f}^{+}}{|\mathbf{f}^{+}|}$ and $\mathbf{e}^{\times} \equiv \frac{\mathbf{f}^{\times}}{|\mathbf{f}^{\times}|}$, and from there we can re-write all of our other quantities.

The projection operator is now

$$\mathbf{P}^{GW} = \mathbf{e}^{+}\mathbf{e}^{+\dagger} + \mathbf{e}^{\times}\mathbf{e}^{\times\dagger} \tag{19}$$

and the standard likelihood is now

$$E_{SL} = \sum_k \left[|\mathbf{e}^{+} \cdot \tilde{\mathbf{d}}|^2 + |\mathbf{e}^{\times} \cdot \tilde{\mathbf{d}}|^2 \right]. \tag{20}$$

Additionally, we can find the plus and cross energies, which are

$$E_{+} \equiv \sum_k |\mathbf{e}^{+} \cdot \tilde{\mathbf{d}}|^2 \tag{21}$$

and

$$E_{\times} \equiv \sum_k |\mathbf{e}^{\times} \cdot \tilde{\mathbf{d}}|^2 \tag{22}$$

respectively.

Once we have these quantities, we can determine their statistical properties. All these energies follow a non-central χ^2 distribution of the form

$$2E\ \chi^2_{2N_p D_{proj}}(\lambda) \tag{23}$$

where E is energy, N_p is number of pixels, D_{proj} is the number of projection dimensions, and λ is a non-centrality parameter.

This non-centrality parameter λ is given by

$$\lambda_{+} = 2 \sum_k |\mathbf{f}^{+}|^2 |h_{+}|^2, \tag{24}$$

$$\lambda_{\times} = 2 \sum_k |\mathbf{f}^{\times}|^2 |h_{\times}|^2, \tag{25}$$

and

$$\lambda_{SL} = 2 \sum_k \left[|\mathbf{f}^{+}|^2 |h_{+}|^2 + |\mathbf{f}^{\times}|^2 |h_{\times}|^2 \right] \tag{26}$$

for the plus, cross, and standard likelihoods.

By using the mean and standard deviation of the noncentral χ^2distribution, we can state that a gravitational wave is expected when

$$\frac{\lambda}{\sqrt{2N_p D_{proj}}} >> 1 \tag{27}$$

.

This ratio can be very helpful in establishing why a program is failing to detect injected gravitational waves.

This lengthy definition underlies the X-Pipeline set of functions, which were already made prior to this project and many of which we used in our reconstructions. For more information, see the appropriate X-Pipeline documentation [3].

3.2 Pre-made Functions

During this project, we drew heavily on pre-made functions in X-Pipeline. These functions will not be directly discussed here, but discussion of them can be found in the appropriate X-Pipeline documentation [3].

3.3 script_wfr

The primary code we used to conduct our reconstruction analysis was a MATLAB script called script_wfr.m. It is based in X-Pipeline and uses X-Pipeline functions extensively.

Before running the script, the user sets a series of parameters. Users select data parameters and noise models; we used LIGO Hanford 1, LIGO Livingston, and VIRGO but assumed a Gaussian version of the LIGO noise model for all three detectors. Users can also have dynamic noise creation, but we did not use that option. Users set a data duration and sample rate and minimum and maximum frequencies of interest. We use a data duration of 64 seconds and a minimum frequency of interest 40 Hz for all analyses. We predominantly used a sample rate of 1024 Hz and maximum frequencies of 450 Hz for most of our analyses. The notable exception is the halo-core grid created near the end. There we used a data sample rate of 2048 Hz and a maximum frequency of interest of 1000 Hz. The user may also choose to turn noise on and off and, in our edited version, how many time-frequency clusters to keep.

The next parameters are the wavelet decomposition parameters: the level of wavelet decomposition, the wavelet family, and the entropy type. We used a level of 6, wavelet 'dmey,' and entropy 'shannon' for the whole project. After them, the user sets the wavelet reconstruction parameters: an energy for pixel thresholding, the percentile of wavelet pixels to zero out, and the expected gravitational wave strain amplitude. The script defaulted to standard energy, an amplitude of 1×10^{-21}, and a percentile of 97.

Percentile can be either a single value (known as the single case), in which that percent of time-frequency pixels are zeroed out during the reconstruction, or it can be a pair of values (known as the halo-core case). In the halo-core case, some lower (halo) percentile value is used to eliminate an initial set of values. Then, these values are refined farther using the higher core value. Only those values which are adjacent to pixels greater than the core values are retained. While it is not the default operation of the program, we also tested the halo-core variation extensively.

The user then sets injection parameters. Users must choose the deterministic or random case. The deterministic case injects a waveform with fixed parameters and allows the user to examine it. The random case injects a specified number of signals with randomized parameters and allows the user to examine more the overall than individual effects. We

used only the deterministic case, so we will not further discuss the random case. In the deterministic case, the user must specify a signal, its the root sum square amplitude for the plus and cross polarizations (hrssp and hrssc), and a waveform model. The user must also specify polarization angles and sky position.

Once all the parameters are set, the program autonomously generates a waveform and a noise spectrum, calculates the antenna response functions, calculates the time delay due to sky position, and injects and applies these quantities. Then, the program time shifts the data and calls the function hreconstruction, using parameters provided by the user, and reconstructs the waveform.

After the waveform is reconstructed, script_wfr evaluates the reconstruction. It calculates the inner product of the reconstructed waveform and the injected waveform and divides by the magnitude of the injected waveform, for both the plus and cross polarizations as shown in Equation 28:

$$match = \frac{\langle \mathbf{h_{rec}} | \mathbf{h}_{inj} \rangle}{\langle \mathbf{h}_{inj} | \mathbf{h}_{inj} \rangle} \tag{28}$$

This value is known as the match in the plus and cross and is the most reliable indicator of the effectiveness of the reconstruction program.

Finally, script_wfr plots the injected and reconstructed waveform to allow for visual comparison.

3.4 hreconstruction

hreconstruction is a MATLAB function for a Tikhonov-regulated gravitational waveform estimator. It takes in an array of detector data, a data structure with minimum, maximum and Nyquist frequencies, a wavelet decomposition structure, a noise spectrum name, antenna response functions, a gravitational wave strain amplitude, an energy threshold, a one or two element percentile vector, and a number of clusters to keep (in our revised variation).

It first performs a wave packet decomposition. Then it constructs a vector to hold the central frequency of each wavelet coefficient and re-orders this vector into ascending regions. It next eliminates areas of time-frequency maps which are likely to be highly un-reliable, such as the first and last 10% of time bins.

The function then computes the one-sided power noise spectrum based on the given theoretical noise curve. Next, the function double-whitens the data by dividing each sample by its noise amplitude and noise power.

The function constructs a map of time-frequency pixels based on a chosen detection statistic, zeros a number of pixels appropriate to the given percentile value, and forms clusters of the remaining values. Clusters are formed from locations where the nearest neighbor of a valid pixel is also a valid pixel. Then, the function chooses which clusters to keep in reconstructing the waveform. During this project, we asked the function to keep only the loudest cluster of pixels, only the loudest N clusters of pixels, and all clusters with values above some fraction of the loudest cluster.

Finally, the function turns these retained clusters back into a waveform and outputs it.

4 Method

4.1 Preliminary Analysis

There are a myriad of proposed supernova waveforms. Therefore, the first task in improving reconstruction for supernova waveforms was to choose a set of waveforms to work with. For this project, it was important to generate a diverse but not overwhelmingly large set of data, which focused on more likely waveforms. Therefore, the data set is composed of ten waveforms selected from the optical supernova search catalog. These waveforms were generated by numerical analysis, all of them had an adjustable root sum square amplitude for the plus polarization, and half had a root sum square amplitude equal to zero for the cross polarization. The waveforms and their most important properties are given in Table 1 .

Table 1: Waveform List

Waveform No.	Waveform Name	hrssp	hrssc
1	processed_signal_s15a2o05_ls-plus	value	0
2	processed_signal_s15a2o09_ls-plus	value	0
3	processed_signal_s15a3o15_ls-plus	value	0
4	processed_s15-time-rhplus_matter-plus	value	0
5	processed_s15.0.h-plus	value	0
6	L15-2_theta0.000_phi0.000-plus and -cross	value	value
7	L15-3_theta0.000_phi0.000-plus and -cross	value	value
8	N20-2_theta0.000_phi0.000-plus and -cross	value	value
9	W15-2_theta0.000_phi0.000-plus and -cross	value	value
10	W15-4_theta0.000_phi5.585-plus and -cross	value	value

Once all the waveforms were selected, the next step was to develop an intuition about the performance of script_wfr. Though the script was completed and functional, its performance on supernovae was only vaguely understood. To improve our understanding, we generated a detailed spreadsheet, recording the signal-to-noise ratios and plus and cross match values for each waveform at root sum square amplitudes of 1×10^{-19}, 1×10^{-20}, and 1×10^{-21}. This process established expectations for level of match and allowed the selection of realistic root sum square amplitudes based on known LIGO and VIRGO signal-to-noise ratios for further analysis.

4.2 Altered Cluster Analysis

After developing an intuition for the performance of the waveform reconstruction script without modifications, we began modifying the script to improve its matching ability.

The first step we took was altering the number of clusters retained by the reconstruction script. By default, only the loudest cluster was kept. We modified this section of code so that rather than finding the maximum cluster and retaining it, the code sorted the clusters by size and then retained a number of clusters set by the user. In the case where the number of clusters to keep was equal to one, the new version functioned identically to the old version.

Additionally, we also wrote a section of code where the script found the maximum cluster value and then retained all clusters greater than some fraction of the loudest cluster. This fixed-level version shared the concept of a tunable cluster keeping parameter with the other alteration. As in the basic multiple cluster case, setting the keep parameter equal to one returned the original version of the script.

4.3 Altered Percentile Analysis

Altered percentile analysis was the heart of this summer's project. As in the altered cluster analysis, the goal was to maximally improve the script's matching ability and find optimal values for percentiles to use in reconstructing waveforms. Also like before, altered percentile analysis encompassed a number of smaller procedures.

First, we turned off the noise so that the only thing injected to the reconstruction program was the signal. We then ran the script with percentile set at intervals of 0.1 between 90 and 99 and large intervals between 0 and 90 for the single percentile case. For the halo-core case, we took strips of data with constant core and varying halo. In this manner, we determined the base effectiveness of the reconstruction script.

Then, we turned noise back on and repeated our single percentile and halo-core strip analysis. This allowed us to look for optimal percentiles in the single case and develop a first intuition regarding optimal percentiles in the halo-core case. To find these general optimal values, we averaged results from all 10 waveforms and plotted the average.

Finally, we developed a program which allowed us to alter both the halo and core values in the halo-core case and thereby plot a grid of halo and core percentiles as a color-coded map. We first plotted a low resolution grid with intervals of 1 to test the efficacy of our script then plotted a high resolution grid with intervals of 0.25 to determine optimal halo-core values.

5 Analysis

5.1 Preliminaries

Our preliminary analysis sought to give us a general understanding of the waveform reconstruction script's performance by using it on a variety of waveforms at a variety of root sum square amplitudes, all at the same default percentile value and with only the loudest cluster maintained. The results are presented below in Table 2. Additionally, example waveform reconstructions showing the variation in the preliminary analysis are shown in Figures 1 to 3.

This process had three important conclusions. First, a mathematically high match can still sometimes lead to a visually inconsistent waveform. Second, changing the root sum square amplitude can have a very large, very rapid effect on the ability of the program to accurately reconstruct the signal. Third, waveforms with zero cross root sum square amplitudes reach a realistic signal-to-noise ratio when plus root sum square amplitude on the order of 1.00×10^{-20} while waveforms with a non-zero cross root sum square amplitude do not reach a realistic signal-to-noise ratio until the order of 1.00×10^{-21}. This third conclusion dictated the values used for hrssp and hrssc in all future analysis. As the expected amplitude values of the signal are around 6×10^{-21} it is not shocking that some waveforms perform more realistically at 10^{-20} and some at 10^{-21} **??**.

5.2 Altered Cluster Analysis

Initially in the altered cluster analysis, we were unsure whether multiple clusters would be an effective strategy in the absence of noise and therefore we tested the approach. This

Table 2: Appendix: Preliminary Waveform Results

Waveform No.	hrssp	hrssc	SNR (H1)	SNR (L1)	SNR(V1)	matchhp	matchhc
1	1.00×10^{-19}	0	3.25×10^{2}	1.16×10^{3}	3.97×10^{2}	0.9641	NaN
1	1.00×10^{-20}	0	32.5	1.16×10^{2}	39.7	0.9388	NaN
1	1.00×10^{-21}	0	0.248	11.6	3.97	3.52×10^{-11}	NaN
2	1.00×10^{-19}	0	3.26×10^{2}	1.61×10^{3}	3.99×10^{2}	0.9307	NaN
2	1.00×10^{-20}	0	32.6	1.16×10^{2}	39.9	0.9011	NaN
2	1.00×10^{-21}	0	3.26	11.6	3.99	2.28×10^{-11}	NaN
3	1.00×10^{-19}	0	2.34×10^{2}	83.5	2.86×10^{2}	0.9977	NaN
3	1.00×10^{-20}	0	23.4	83.3	28.7	0.9233	NaN
3	1.00×10^{-21}	0	2.34	8.33	2.86	0.1877	NaN
4	1.00×10^{-19}	0	1.64×10^{2}	5.85×10^{2}	2.01×10^{2}	0.9569	NaN
4	1.00×10^{-20}	0	16.4	58.5	20.1	0.7829	NaN
4	1.00×10^{-21}	0	1.64	5.85	2.01	1.23×10^{-10}	NaN
5	1.00×10^{-19}	0	3.15×10^{2}	1.07×10^{3}	3.69×10^{2}	0.9511	NaN
5	1.00×10^{-20}	0	30.2	1.07×10^{2}	36.9	0.7409	NaN
5	1.00×10^{-21}	0	3.02	10.7	3.69	2.63×10^{-11}	NaN
6	1.00×10^{-19}	1.00×10^{-19}	2.07×10^{3}	2.26×10^{3}	2.44×10^{3}	0.9990	0.9998
6	1.00×10^{-20}	1.00×10^{-20}	2.07×10^{2}	2.26×10^{2}	2.44×10^{2}	0.9551	0.9886
6	1.00×10^{-21}	1.00×10^{-21}	20.7	2.23×10^{2}	24.4	0.2616	0.6548
7	1.00×10^{-19}	1.00×10^{-19}	2.67×10^{3}	3.18×10^{3}	3.16×10^{3}	0.9993	0.9998
7	1.00×10^{-20}	1.00×10^{-20}	2.67×10^{2}	3.17×10^{2}	3.16×10^{2}	0.9764	0.9915
7	1.00×10^{-21}	1.00×10^{-21}	26.7	31.7	31.6	0.6348	0.8819
8	1.00×10^{-19}	1.00×10^{-19}	1.99×10^{3}	2.08×10^{3}	2.34×10^{3}	0.9943	0.9969
8	1.00×10^{-20}	1.00×10^{-20}	1.99×10^{2}	2.08×10^{2}	2.34×10^{2}	0.9353	0.9835
8	1.00×10^{-21}	1.00×10^{-21}	19.9	20.8	23.4	0.5269	0.8062
9	1.00×10^{-19}	1.00×10^{-19}	1.96×10^{3}	2.15×10^{3}	2.31×10^{3}	0.9799	0.9661
9	1.00×10^{-20}	1.00×10^{-20}	1.96×10^{2}	2.15×10^{2}	2.31×10^{2}	0.9072	0.9528
9	1.00×10^{-21}	1.00×10^{-21}	19.6	21.5	23.1	0.2967	0.5701
10	1.00×10^{-19}	1.00×10^{-19}	2.38×10^{3}	2.62×10^{3}	2.81×10^{3}	0.9896	0.9958
10	1.00×10^{-20}	1.00×10^{-20}	2.38×10^{2}	2.62×10^{2}	2.81×10^{2}	0.9443	9.8840
10	1.00×10^{-21}	1.00×10^{-21}	23.8	26.2	28.1	0.4863	0.8123

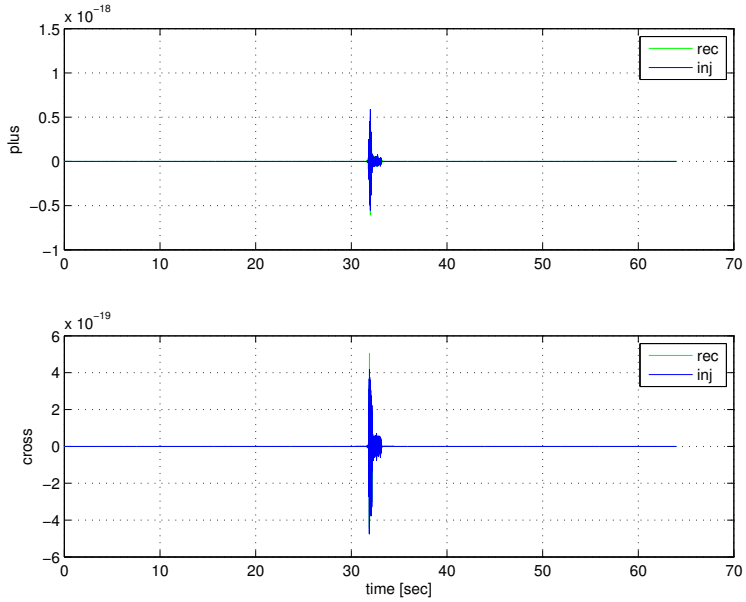

Figure 1: Injected and reconstructed signal for Waveform (6) at hrssp = hrssc = 10^{-19}.

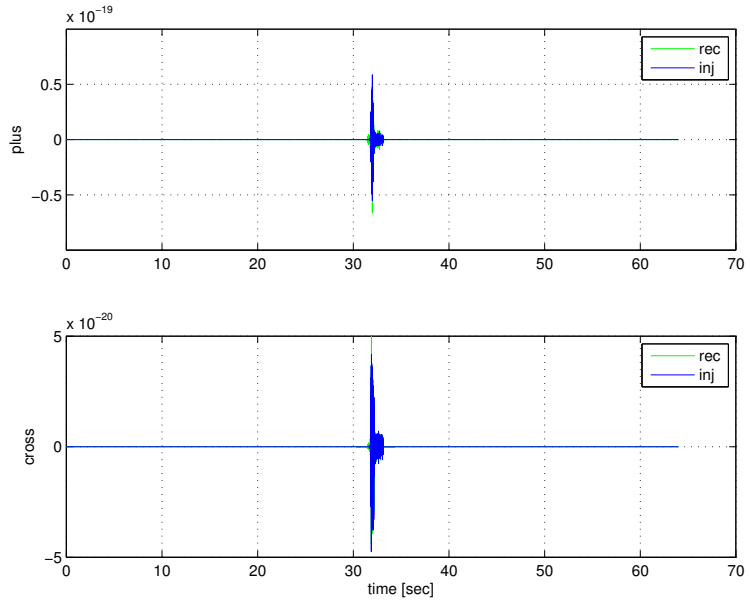

Figure 2: Injected and reconstructed signal for Waveform (6) at hrssp = hrssc = 10^{-20}.

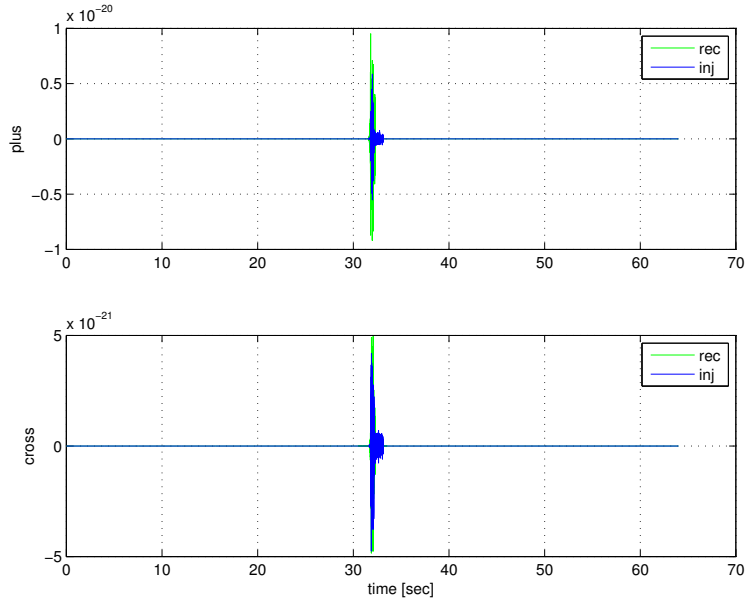

Figure 3: Injected and reconstructed signal for Waveform (6) at hrssp = hrssc = 10^{-21}.

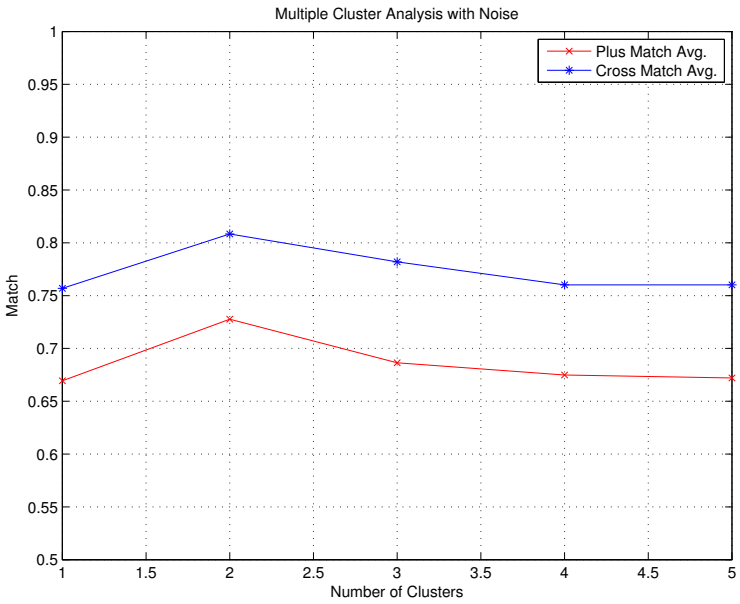

Figure 4: Results for multiple cluster analysis with noise, averaged over all 10 waveforms.

was an ineffective undertaking, as some of the signals did not have any secondary clusters in the absence of noise. For those that could form multiple clusters, the alteration had no effect on the reconstruction. Since increasing the number of clusters had no effect on the match values, we concluded that noise was necessary to make any difference in the quality of reconstruction. This result is not entirely unexpected, as there is little danger of the reconstruction script being unable to distinguish which pixels are signals when all the pixels are signals.

Then, we tested the multiple clusters strategy with noise. As shown in Figure 4 , increasing the number of clusters to two increased the match, but the effect dropped off for three, four, and five clusters maintained.

We also conducted brief analyses with fixed threshold variations and found that maintaining a threshold of about 0.3 times the value of the maximum cluster was very good for the cross polarization. However, the threshold version did little to help the plus version.

There are a number of reasons to be cautious with these results. For many gravitational waveforms, maintaining multiple clusters is helpful and prevents the reconstruction program from missing elements of the waves. However, this is only helpful if there are in fact multiple segments of the time vs. frequency plot that contain signal. We were unsure if the waveforms we used would be helped by having multiple clusters because not all met that qualification. Additionally, the distribution of noise is not something we are directly controlling, nor is it constant between different waveforms or different runs. This means that results may be skewed in ways we cannot directly show. Further work would be necessary to speak definitively on this topic.

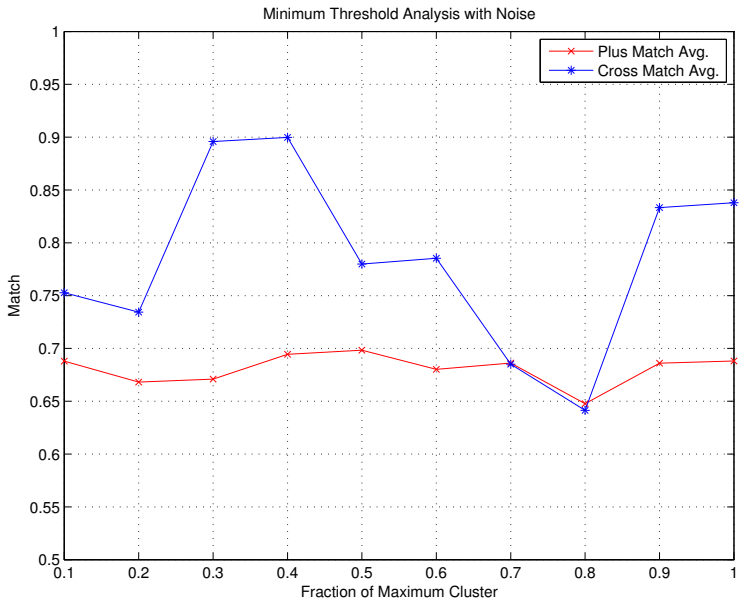

Figure 5: Results for minimum threshold analysis with noise, averaged over all 10 waveforms.

5.3 Altered Percentile Analysis: Noise-free

In conducting altered percentile analysis, we began by removing noise from the system and altering the single and halo-core percentiles. We sought to establish what the maximum reconstruction ability of the program was when maintaining only the loudest cluster. In other words, we wanted to understand if the program had an obvious maximum point.

Figure 6 demonstrates a plot at single percentile values between 90 and 99 of the average plus match for all ten waveforms and the average cross match for the five waveforms with non-zero cross amplitudes. For each waveform, the match value was constant before reaching a percentile value where the signal began to be excluded and dropping off precipitously. Each waveform had a slightly different drop-off point, but all ranged between about 98.5 and 99.5. The average drop-off point was 98.7.

Figure 7 is a plot of halo-core values where the core percentile is held constant at 99.9 and the halo value is varied between 90 and 99.9 by steps of 0.1. The cross match appears almost identical to the cross match in the single case, though with a drop-off point of 99. The plus match appears slightly differently than the cross match, as the signal is no longer smooth. However, the trend of the plus match curve is the same as the other noise-free curves. It also shares a drop-off point with the cross match curve.

Altered percentile analysis without noise proved a plausible upper limit on useful percentile values. It allowed us to establish that it was entirely possible for the program to be too discriminating and exclude valid signals. As it established the highest percentile values we could use without excluding signals, it provided a tentative target point for analysis with noise included.

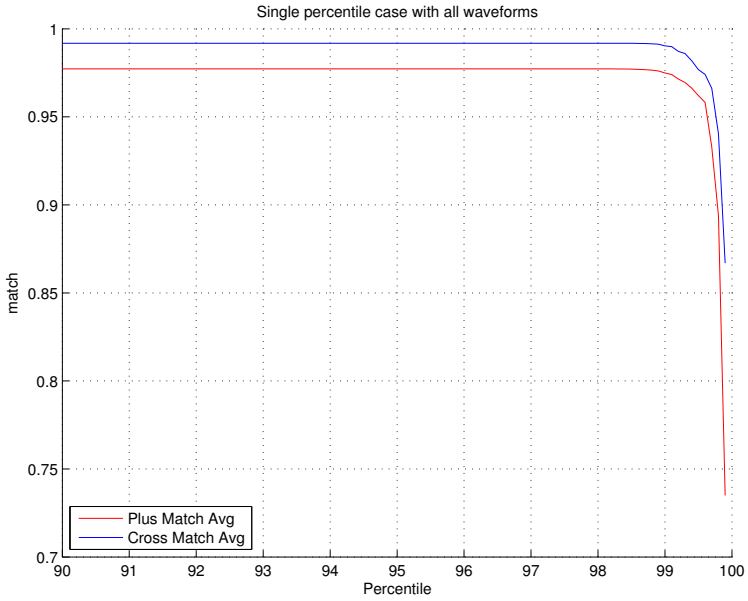

Figure 6: Results for varied single percentile values averaged over all 10 waveforms without noise.

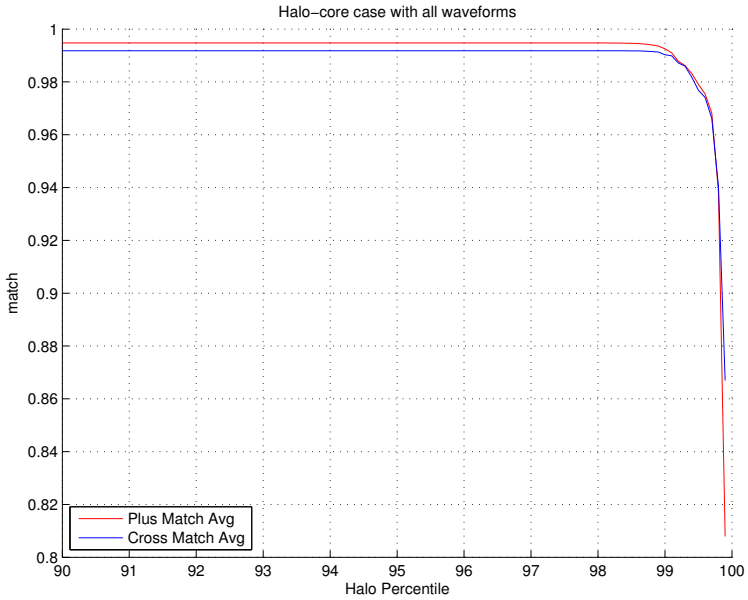

Figure 7: Results for strips of halo-core percentile values averaged over all 10 waveforms without noise.

5.4 Altered Percentile Analysis: Noise Included

After developing a basic understanding of the effects of altering the percent of pixels retained without noise, we switched to analyzing the effects with noise. In many ways, this was the heart of the process, as it was here we were able to make realistic statements about performance of the match program in non-ideal situations.

We first repeated the single percentile analysis with the noise turned on. As before, we altered the single percentile value between 90 and 99 by steps of 0.1 for each waveform. We then averaged the results from all ten waveforms and plotted these results in Figure 8. This plot also includes the results for one particular waveform, designated waveform 6, as a comparison. Those comparative curves are designated Plus Match Baseline and Cross Match Baseline. Determining an optimal point for the case including noise was more difficult than in the no-noise case, as the optimum is less obvious. There appear to be two locations of approximately equal quality, one between 97 and 98 and one between 95 and 96. This is illustrated by Tables 3 and 4, recording the top five match values for plus and cross.

Table 3: Top 5 Plus Matches - Single percentile with noise

Rank	Percentile	Plus Match	Cross Match
1	98.7	0.6191	0.8089
2	97.0	0.6096	0.7621
3	97.5	0.6069	0.7884
4	95.7	0.6042	0.8092
5	95.8	0.6006	0.7695

Table 4: Top 5 Cross Matches - Single percentile with noise

Rank	Percentile	Plus Match	Cross Match
1	65	0.5810	0.8146
2	83	0.5818	0.8110
3	58	0.6042	0.8092
4	88	0.6191	0.8089
5	55	0.5625	0.8060

We then repeated the halo-core percentile analysis with noise. In this case, we also refined our search area slightly and examined only the area with percentile values between 94 and 99.9, based on the location of our perceived optimum in the single percentile case. We held the core percentile constant at 99.9 and increased the halo percentile from 94 to 99.9 in steps of 0.05. As in the other noise case, we determined the match values at each percentile combination for each waveform then averaged all ten waveforms for plotting and analysis, as seen in Figure 9. Note that the exceptionally low curve in that figure is the Plus Match Baseline, indicating that our example waveform had a particularly poor plus match. The average plus match over all ten curves is noticeably higher. Here again, as in the single percentile case, the two best areas seem to be for halo values between 95 and 96 and again

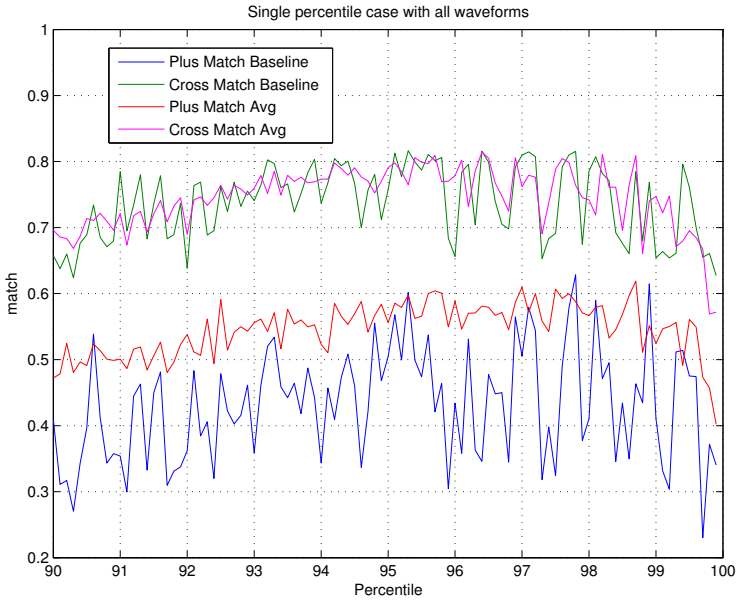

Figure 8: Results for varied single percentile values averaged over all 10 waveforms with noise.

between 97 and 98. Specifically, the highest value of plus match is 0.7130, observed at a halo value of 96.10 and a core value of 99.9. The highest value of cross match is 0.8112, observed at a halo value of 96.65 and a core value of 99.9.

Table 5: Top 5 Plus Matches - Halo-core percentile with noise

Rank	Halo	Core	Plus Match	Cross Match
1	96.10	99.90	0.7130	0.7814
2	95.90	99.90	0.6999	0.7839
3	97.10	99.90	0.6965	0.7362
4	95.35	99.90	0.6942	0.7882
5	98.25	99.90	0.6938	0.7852

While a 1-dimensional alteration in percentile value represents a fairly robust analysis of the single percentile case, this approach is not as useful for the halo-core case because it does not provide a complete set of possible percentile pairs. Therefore, we created a grid which allowed us to alter both the halo and core values for percentile.

We began with a low-resolution grid, which had halo and core values ranging between 90 and 99 in steps of 1. The results are show in Figure 10. This test served primarily as a diagnostic, ensuring that the scripts to generate and plot the grids were working properly. Additionally, we hoped it would allow us to see gather some impression of where optimal values might reside. In the plus case, we ascertained a general idea that the optimal core existed between 95 and 97. In the cross case we ascertained a idea that the optimal core was

Table 6: Top 5 Cross Matches - Halo-core percentile with noise

Rank	Halo	Core	Plus Match	Cross Match
1	96.65	99.90	0.6722	0.8112
2	97.00	99.90	0.6738	0.8079
3	95.00	99.90	0.6893	0.8066
4	97.20	99.90	0.6823	0.8059
5	95.75	99.90	0.6696	0.8058

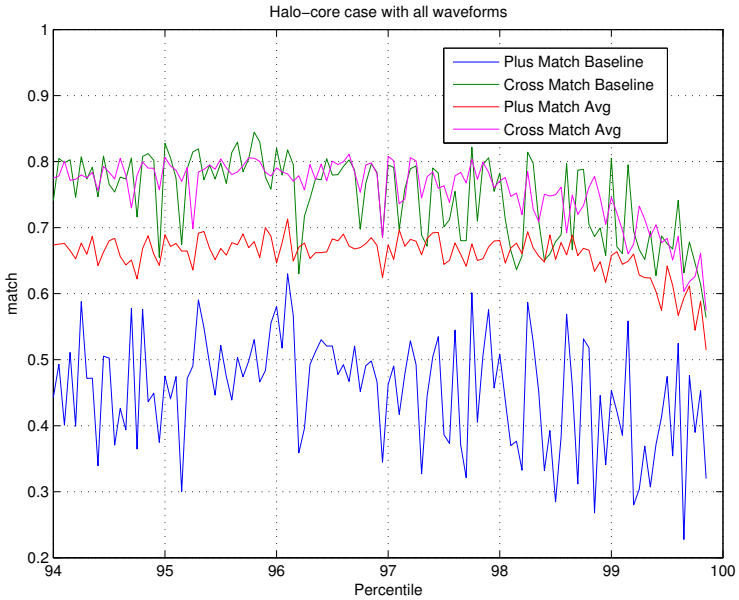

Figure 9: Results for strips of halo-core combinations percentile values averaged over all 10 waveforms with noise.

also between 95 and 97 and the optimal halo was between 94 and 98.

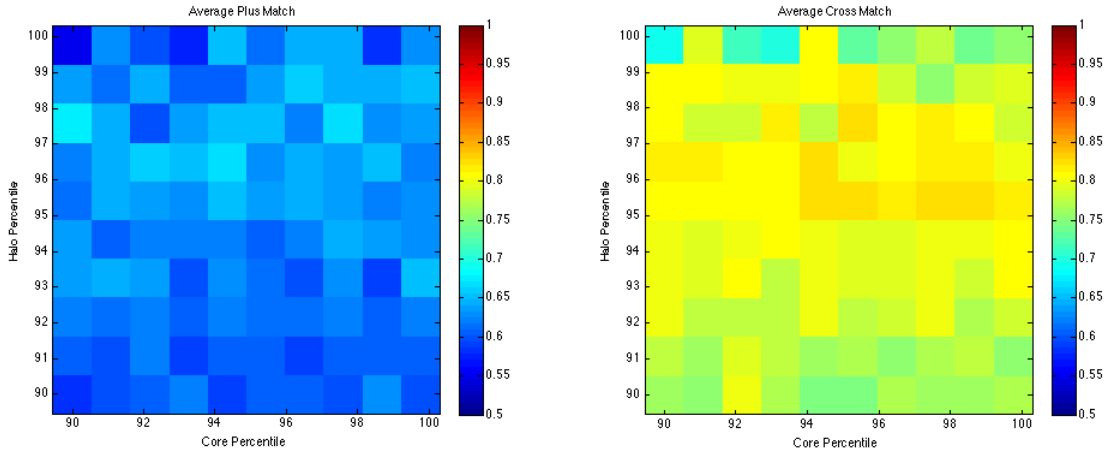

Figure 10: Low-resolution (1% per step) halo-core grid. Maximum match points are marked with a *.

After generating the low-resolution grid, we switched to a high-resolution grid, with halo and core values ranging between 90 and 99.75 by steps of 0.25. The decision to use step sizes of 0.25 rather than 0.1 or 0.05 (as we had used previously) was a pragmatic one; we were concerned about the amount of time necessary to run the program and felt that the additional resolution was not worth the additional time.

We determined the top five match values for both plus and cross, as shown in Tables 7 and 8. We also created maps (Figures 11 with halo percentile on the x-axis, core percentile on the y-axis, and match value on the color bar to illustrate the quality of match. The best plus polarization match was 0.6811 at a halo value of 96.75 and a core value of 92.75. The best core match was 0.8369 at a halo value of 95.75 and a core value of 91.00. The single best point overall was probably at halo of 96.5 and core of 92.75, where the plus match was the highest in the grid and the cross match was second highest.

Table 7: Top 5 Plus Matches - High-resolution Grid

Rank	Halo	Core	Plus Match	Cross Match
1	96.00	92.75	0.6811	0.8156
2	96.50	92.75	0.6788	0.8341
3	98.25	99.75	0.6746	07832
4	97.25	98.75	0.6739	0.7876
5	93.50	96.50	0.6721	0.8101

Another thing which is important to consider in analyzing these grids is that anytime the halo percentile value is greater than the core value, the analysis is actually the same as the single percentile analysis. It does not matter what the core value is because the halo is more

Table 8: Top 5 Cross Matches - High-resolution Grid

Rank	Halo	Core	Plus Match	Cross Match
1	95.75	91.00	0.6490	0.8369
2	96.50	92.75	0.6788	0.8341
3	98.00	90.50	0.6303	0.8331
4	96.00	90.50	0.6303	0.8331
5	95.75	91.25	0.6241	0.8327

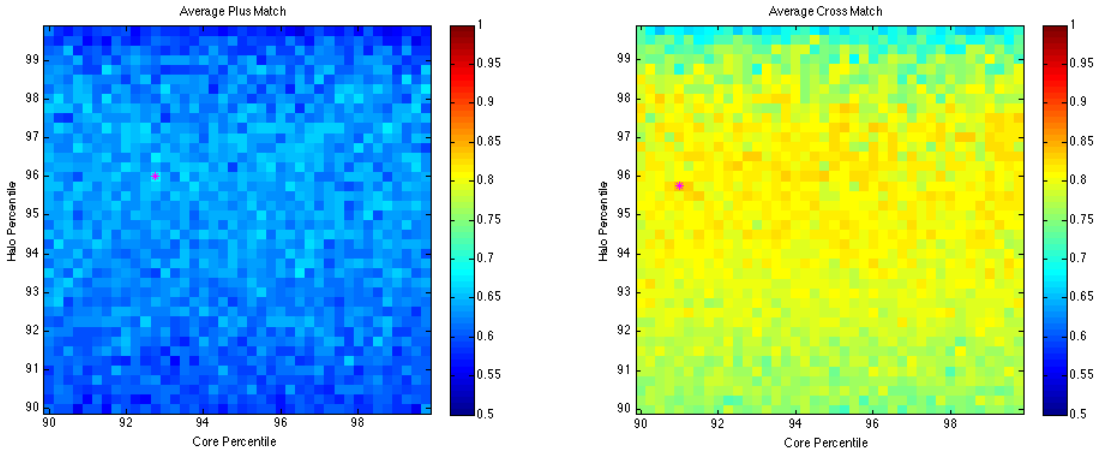

Figure 11: High-resolution (0.25 % per step) halo-core grid. Maximum match points are marked with a *.

discriminating. This means that in order to get a grid picture of just the halo-core cases, we divided the grid data into two triangular sections and removed the portion representing the single percentile case. This new triangle is plotted in Figure 12. When we did this, we found a different maximum values: 0.6746 at halo of 98.25 and core of 99.75 for the plus polarization and 0.8303 at halo of 97.00 and core of 98.50 for the cross polarization.

One reassuring conclusion we drew from the percentile grid is that our previous estimation that the best points were distributed around percentiles of 96 and 98 seems to be upheld by the grid. In fact, this set seems to imply that the best percentiles at single percentiles between 95 and 96. However, we do not think it wise to definitively state that this region is the best without repeating the analysis to modulate uncertainties, particularly those from varying noise.

Uncertainty in our simulation, which we hypothesize to be derived from variations in injected noise, is highlighted by the top five lists. If, as the lists suggest, single percentile values are the best values, then the top should be all different core combinations with the best halo. However, that is not what we observe. This indicates that match value does vary over different injections and reconstructions.

We examined the full percentile grid as well as the halo-core and single triangles, and

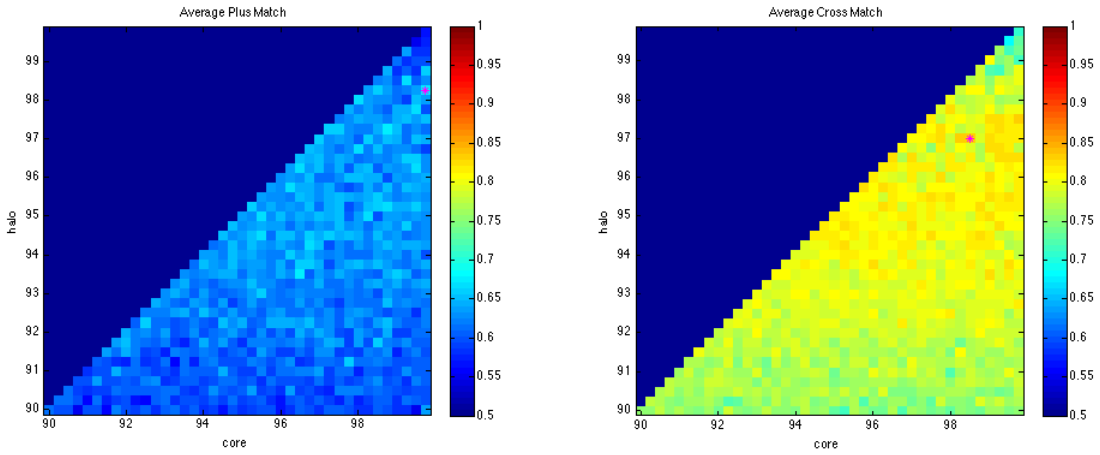

Figure 12: The halo-core percentile grid with all single points removed

concluded that the halo-core method is not significantly better at reconstructing signals than the single percentile case. In fact, the maximum values for both plus and cross in our a occurred in the single case. The mean cross match value is greater for the halo-core case, but the mean plus match value is greater for the single case. (See Table 9.) Additionally, the differences in match value are very small relative to the match value itself. Therefore, in the situations we have investigated, the choice to use single percentile or halo-core percentile seems to be a matter of taste.

Table 9: Maxima and Means: Single and Halo-Core from Percentile Grid

Rank	Value	Type
Maximum +	0.6811	Overall
Maximum +	0.6811	Single
Maximum +	0.6746	Halo-core
Maximum +	0.8369	Overall
Maximum ×	0.8369	Single
Maximum ×	0.8303	Halo-core
Mean +	0.6225	Overall
Mean +	0.6250	Single
Mean +	0.6202	Halo-core
Mean ×	0.7849	Overall
Mean ×	0.7826	Single
Mean ×	0.7872	Halo-Core

6 Conclusion & Directions for Future Research

Results thus far indicate that it is possible to optimize the reconstruction process, but that even in its optimal state, the reconstruction is unlikely to be anywhere near perfect. While the average match for the cross case hovers around 0.8, the average match for the plus case hovers around 0.6. This is a low enough value to introduce a visible difference between the injected and reconstructed waveform in plots of the reconstruction. Therefore, while what we have done so far is good, it can be made better.

There are several avenues for continued investigation into reconstruction of gravitational wave signals from supernovae.

First, we tested only a small selection of ten waveforms during this project. There are actually fourteen more distinct waveforms in the optical supernova catalog, plus 99 more orientations for each of the last five waveforms we tested. In total, then, there are about 500 other variations of waveforms which could be investigated. Applying this analysis to more variations may increase our confidence in our results.

Second, we could repeat our analyses and develop a range of match values for each set of percentiles. At the moment, we only have one value for each variation, meaning that variations in noise or other factors can completely skew our perceptions of which percentiles are best.

Third, the optimum for percentile values is still rather fuzzy. Values presented following this summer's work represent more of a range than a particular value. Higher resolution could be useful in establishing where the best places to set percentile values. We think that it would be particularly nice to build grids with a 0.1 resolution, so that the grids and single strips have the same resolution.

Fourth, most of the altered cluster analysis is open for further investigation. While we did work on it a little, we focused primarily on the altered percentiles analysis. Further analysis can determine which waveforms (if any) the multiple cluster analysis is actually effective for. For those waveforms, further analysis can also determine optimal numbers of clusters and appropriate base values for the fixed value case.

Fifth, we could improve the manner in which we deal with noise in testing our reconstruction script. At present, the manner in which the code is constructed causes it to re-inject the fake data (and therefore form a different noise spectrum) for every call to the reconstruction script. This means that even for " identical" tests, the match values are not always the same. It would be helpful (especially for constructing grids such as those in Figures 10 and 11) if noise curves were consistent for each point in a test.

References

[1] Christian D. Ott. The gravitational-wave signature of core-collapse supernovae. *Classical and Quantum Gravity*, 26, Februrary 2009.

[2] B.S. Sathyaprakash and Bernard F. Schutz. Physics, astrophysics and cosmology with gravitational waves. *Living Reviews in Relativity*, 12(2), 2009.

[3] Patrick Sutton, Gareth Jones, Shourov Chatterji, Peter Kalmus, Isabel Leonor, Stephen Poprockl, Jameson Rollins, Antony Searle, Leo Stein, Massimo Tinto, , and Michal Was. X-pipeline: an analysis package for autonomous gravitational-wave burst sources. *New Journal of Physics*, 12, May 2010.