Draft v0.4

Gravitational Wave Bursts: Characterization of Transients in LIGO Interferometer Data

Ed Brambley
Mentor: Dr. John Zweizig
Co-Mentor: Dr. Szabolcs Márka

May 2, 2004

This report details methods of detecting and categorizing transients in small sections of data. Transients are short lived, dramatic variations in a signal. Typical signal lengths are 10 to 60 seconds of data, sampled at 16384Hz.

A spectrographic transform was used to allow consideration of the time evolution of power in individual frequency bands. Various statistical procedures were then considered for detecting transients within these bands, and methods involving Likelihood Ratio testing based on Gamma and Normal distributions was selected. Finally, post processing of the results from each frequency band yielded a concise list of possible transients, together with their start and end times, total power, and frequency composition.

The final algorithms demonstrated their ability to successfully detect and classify transients in LIGO interferometer data. Problems with badness of fit were evident between 1–3kHz, and a large number of false signals were generated at isolated frequencies 4320Hz and 7168Hz (both ±16Hz). The latter two frequencies coincide with a dramatic increase in power compared to neighboring frequencies, and did not correspond to any previously known noise sources.

1 Introduction

Previous algorithms implemented to detect transients in LIGO data have been based on finding outliers in filtered time series, assuming a Normal distribution. Such methods include the `glitchMon` and `glitch` trigger generation programs. However, it was desired to further characterize the transients indicated by these triggers.

This report details the progress of a SURF investigation into methods of searching for and characterizing transients, in both frequency and time. It bears similarities to the work of Stuver [1] on BlockNormal analysis, and to the work of Sylvestre [2] on spectrographic processing of LIGO data.

2 Implementation

2.1 Pre-Processing

It was considered that each transient could be localized to a small frequency range, which implied far better results could be expected from considering several frequency ranges separately. This was implemented by a spectrographic transform, which involved Discrete Fourier Transforms (DFTs) the data in blocks which overlapped by 50%, using a Hanning window. Typically these blocks were 512 samples in length, which for a 16384Hz sampling frequency, gave a time resolution of 16ms and a frequency resolution of 32Hz. Statistical

analyses were performed independently on the time series of power in each frequency, or in some cases on the time series of real and imaginary parts of the DFT coefficients.

The power time series for each frequency is necessarily positive, but was not in general distributed according to a Normal distribution. To apply Normal analysis to these data, a Box-Cox transform was considered. This transform is a non-linear, smooth, monotonically increasing scaling of positive data — see A.1 for details.

2.2 Statistical Analysis

Previous analyses used by LIGO fitted a distribution to the data, and then looked for outliers on this distribution. It was decided for this project to investigate a different method, motivated partly by the BlockNormal analysis.

Given a section of data, the idea was to search for the interval (the possible transient) who's distribution differed the most from the rest of the data (the background noise). Unlike the BlockNormal, where sections at the beginning and end of the data were allowed to have different distributions, this method assumed such sections to have the same background distribution, and therefore did not have problems with edge effects. This search was then repeated on both "background noise" sections either side of the proposed transient, to search for other transients in the data.

For parametric distributions, it was decided to use classical statistics rather then the Bayesian statistics used for the BlockNormal analysis, as this would avoid inaccuracies introduced by specifying prior distributions. An interval test based on the Log Profile Likelihood Ratio statistic was used for this analysis (see A.2 for details). Tests based on this were termed Likelihood Ratio Interval Analysis tests.

Various distributions were considered for use with such a test. If the original time-series data were to have a Normal distribution, then the real and imaginary parts of the DFT coefficients would be expected to have a Normal distribution. This would suggest that the power (proportional to the sum of the squares of the real and imaginary components) would have a scaled χ^2_2 distribution. The Gamma distribution is a generalization of a scaled χ^2 distribution, and the corresponding shape parameter for a χ^2_2 is $\alpha = 1$. A Weibull distribution with shape parameter $\alpha = 1$ also has identical form.

The distributions considered for modeling the time series of power in each frequency band were the Normal distribution (both with and without a Box-Cox transform), the Gamma distribution, and the Weibull distribution. In addition, the Di-Normal (two independent Normal distributions) and Bivariate Normal distributions were considered for modeling the time series of real and imaginary parts of the DFT for each frequency band. See A.3 to A.6 for details. In addition to these distributions, two others were investigated; the Shifted Gamma and the Gamma/Normal distributions (the details of which are given in A.7 and A.8 respectively). However, these distributions or their approximations gave identical results to the Gamma distribution, and so are not considered further.

As well as the parametric tests above, two non-parametric tests were considered. These compared the distribution of data within an interval to the distribution of data outside it, and gave a statistic indicating how different the two samples are. The major advantage of these tests over the parametric tests was that they did not assume any kind of distribution for any of the data. The two tests considered were the Kolmogorov-Smirnov and Anderson-Darling tests. See A.9 for details.

2.3 Post-Processing

All the analysis techniques above were applied independently to each frequency from the spectrographic transform. Post-processing was therefore needed to amalgamate blocks found in each frequency into a concise list of candidate transients, together with their frequency components. In order to do this, a histogram was generated, with time along the x axis, and the number of frequency bands for which this time lay in a possible transient along the y axis. A candidate transient was then considered to be an interval of time for which this histogram was non-zero, and to consist of those frequencies contributing to the histogram anywhere in this interval.

The power of the possible transient in each frequency band was considered to be the average power during that time interval, minus the average power outside any signal in that frequency band. These powers were then summed to give the total power in the candidate transient, and its relative frequency power composition.

2.4 Trigger Monitoring

A trigger monitor called `BramMon` was written to apply the above to categorize transients indicated by triggers. It had a per frequency threshold that enabled "noisy" frequencies that gave lots of false signals to be ignored, without compromising sensitivity for other frequencies. The thresholds were estimated from LIGO data using a second program called `BramThreshold`, by analyzing the data in several small blocks and finding the mean and standard deviation of the significance in each frequency; the threshold for that frequency was then set to the mean plus some multiple of the standard deviation.

3 Results

The data samples used throughout this section are described in Appendix B. Samples H1 and H2 were taken from the 4km Hanford interferometer, and where chosen because they each contained a large trigger as detected by `glitch`, about 30 seconds into each sample. Both of these samples were part of LIGO 2 science run 2.

3.1 Goodness of Fit of Parametric Distributions

To test the goodness of fit of the various parametric distributions considered to the data, a generalized Kolmogorov-Smirnov goodness of fit test was used. This test is similar, but not identical, to the K-S two sample test discussed above. For details, see A.10.

3.1.1 Goodness of Fit for Samples H1, H2, and S1

Figures 1–8 show how good a fit the three distributions considered were to the time series of power in various frequency bands, for both LIGO and simulated data. As is clear from Figures 1 and 2, none are Normally distributed. The Gamma and Weibull distributions both provided a good fit for the majority of the LIGO data, and the entirety of the simulated data. The Gamma and Weibull distributions both had worse between 1–3kHz, although of the two, the Weibull performed much better.

Figures 5 and 6 show the values of α and β used to generate Figure 3. As can be seen from Figure 5, the optimal α value was very close to the expected value of $\alpha = 1$, which corresponds to a scaled χ^2_2 distribution. It deviated from this slightly at very low frequencies and between 1kHz and 3kHz, and dramatically at 544Hz and 4320Hz; this deviation was almost identical in both Samples H1 and H2.

Figures 9 and 10 show the results of pre-processing each frequency's power time series by a Box-Cox transform. They demonstrate that the transform successfully manipulated all frequencies of both LIGO and simulated data to be a good fit to a Normal distribution.

Figures 11–13 show the goodness of fit of a Normal distribution to both the real and imaginary parts of the DFT coefficients, for various frequencies. They show that the problematic range of 1–3kHz still remained, but was better fitted than the power time series fits above. The majority of the frequency range was fitted very well.

In all of the above except the Weibull and Box-Cox'ed Normal distributions, a particularly bad fit when compared to neighboring frequencies was seen in the 4320Hz power time series. Figure 21 shows this to have coincided with dramatically higher power compared to neighboring frequencies.

3.1.2 Goodness of Fit for Other Interferometers

Figures 14–17 show similar graphs to those commented on above, for the Hanford 2km and Livingston interferometers. Of these distributions, the Di-Normal seemed to show the best fit, and the Weibull distribution also showed a better fit than the Gamma distribution.

Figures 14 and 15 show that both interferometers gave a bad fit to both Weibull and Gamma distributions in the main sensitivity band up to about 1kHz. Figures 16 and 17 show that the Di-Normal distribution was a good fit in these regions, strongly suggesting its use over the other two algorithms for these interferometers.

Figures 14 and 15 also demonstrate that the problematic area for the Hanford 4km interferometer between 1kHz and 3kHz was not present in the Hanford 2km interferometer, and for the Livingston interferometer, the range was 4kHz–6kHz.

3.2 Transients Responsible for the Triggers

Figures 18 and 19 show the power time series at 288Hz for Sample H1; Figure 19 is a Box-Cox transform of Figure 18. The anomaly probably responsible for the trigger is clearly visible about 30 seconds into the sample, and a second transient is shown about 3 seconds later. Figure 20 shows a closeup of these transients, and demonstrates that each one consisted of three, closely grouped peaks.

From Figure 19, the first transient can be seen to lie about 7 standard deviations off the mean (at 0) in the Box-Cox transformed power time series; this is clearly significant. However, the amount by which the transient is larger than the surrounding noise was far more dramatic in the non-Box-Cox transformed time series, which suggested that accurately modeling the power distribution directly would lead to better results than using a Box-Cox transform and then assuming a Normal distribution.

These candidate transients are the only ones discovered in Sample H1 that may have caused the trigger, and the fact that the largest occurred at exactly the time suggested by the trigger adds weight to this being the cause.

No candidate transient was found in Sample H2 that could have caused a trigger at the appropriate time.

3.3 Likelihood Ratio Interval Analysis Results using a Normal Distribution

Figures 22–27 show the results of a Likelihood Ratio Interval Analysis based on a Normal distribution. They plot the Log Profile Likelihood Ratio statistic against each frequency band considered. In these and following analyses of the LIGO data samples, false signals were more prevalent between 1–3kHz, and at isolated frequencies 4320Hz and 7168Hz. The 1–3kHz and 4320Hz problems were also apparent as a bad fit to several distributions, but the fact that they were present in Figures 23 and 25 despite the Box-Cox transformed data being well fitted by a Normal distribution for *all* frequencies indicated that a bad fit was not the only cause of these false signals. Figure 21 shows both the 4320Hz and 7168Hz frequencies to contain a significantly higher power than their neighboring frequencies. A possible cause for problems in these frequencies would be non-stationarity, where the parameters of the underlying distribution for the noise vary with time.

Figure 22 shows a hugely significant peak centered on 288Hz, which corresponds to the first of the anomalies shown in Figure 18. Figure 23 also picks out this frequency as being significant, but much less so, and in addition shows false signals at 4320Hz and 7168Hz.

Figure 24 and 25 show similar results for Sample H2. None of the peaks shown in either of these figures correspond to an anomaly at the right time to have caused the trigger.

Figures 26 and 27 show the same graphs for the simulated data. Figure 26 demonstrates how well this algorithm can work in an ideal environment, despite the Normal distribution nowhere being a good fit. Figure 27 also shows the same effect of a Box-Cox transform as Figure 23 — that the transform reduces the significance of dramatic outliers, and hence that just being a better fit to the distribution assumed does not guarantee better results.

3.4 Likelihood Ratio Interval Analysis Results for Other Distributions

Figures 28–33 show similar results to the previous section for three different distributions; the Gamma, the Di-Normal, and the Bivariate Normal. Of all the distributions considered, the Gamma and Di-Normal Likelihood Ratio Interval Analyses seemed to perform the best. While the peaks shown in these figures were lower than for the Normal distribution, so was the background level of noise. The 4320Hz and 7168Hz frequencies also stood out far more using these distributions, which suggested that their sensitivity was greater.

From Figures 29, 31, and 33, it can be seen that the Di-Normal distribution gave a lower level of background noise than the Bivariate Normal distribution, while still showing the same sensitivity. This was because considering covariance between real and imaginary parts effectively equates to taking note of the phase of the signal in addition to its power, and a phase change with no corresponding amplitude change easily triggered a false alarm, while it was considered unexpected for a real transient to have had this effect.

Figures 28 and 29 show that both the Gamma and Di-Normal analyses successfully detected the 288Hz transient shown in Figure 18. Of the two distributions, the Di-Normal seemed less susceptible to the 4320Hz and 7168Hz frequencies, although the Gamma distribution showed less background noise by about a factor of 2.

Figures 30 and 31 show similar results for Sample H2. Although the Di-Normal algorithm successfully suppressed problems at 4320Hz, both methods gave significant false signals at 7168Hz. Once again, the Di-Normal's sensitivity seemed better, while the Gamma distribution's background noise was lower.

Figures 32 and 33 show how the two methods performed in an ideal environment. Both methods successfully picked out the transients to the same significance, but again the Gamma distribution's background noise was about half that of the Di-Normal.

The Weibull distribution was not considered for this analysis, as it would have taken approximately two weeks of computer processing to generate results for the three samples considered.

3.5 Non-Parametric Interval Analysis Results

Figures 34–39 show the results of the K-S and A-D nonparametric tests. They plot the corresponding maximum statistic against frequency. Figures 34–37 still show large false signals were generated at 1–3kHz, 4320Hz, and 7168Hz. Since neither test assumed any distribution, this demonstrated that problems for these frequencies must have been due to non-stationarity.

For Figures 34 and 35, a local maximum can be seen at 288Hz. However, these peak are by no means the largest, which occur at either 4320Hz or 7168Hz — results based on these tests would have therefore ignored the "obvious" anomalies shown in Figure 18 in favor of false signals at 4320Hz and 7168Hz.

Figures 36 and 37 show similar graphs for Sample H2. Again, both are dominated by false signals at 4320Hz and 7168Hz.

Figures 38 and 39 show how the two tests faired on the simulated data sample. Both picked out the 1kHz transient, but only the A-D test uncovered the 4kHz transients, and even then only just.

Generally, it appeared that the A-D test faired better compared to the K-S test, and that it showed better sensitivity. However, both tests performed far worse than any of the parametric tests. This was probably due to the non-parametric distributions only considering the rank order of the data points, rather than their value, so that dramatic outliers were treated the same as marginal outliers. In addition, the empirical distribution function tends to be inaccurate in the tails of the distribution, as this is where the fewest points are; however, this is where the distribution is needed to be most accurate for the detection of transients.

3.6 Post-Processing Results

Appendix C contains both the raw and post-processed results of recursive application of the Likelihood Ratio Interval Analysis for the Normal, Di-Normal, and Gamma distributions. The threshold levels used to generate these results were guessed so as to provide a reasonable quantity of data for post-processing. The values used in practice would have to be estimated by running the algorithms on large quantities of LIGO data.

Tables 1–9 show the raw results before post-processing, while Tables 10–18 list the corresponding post-processed results.

3.7 Trigger Monitor Results

Appendix E lists the output of `BramMon` on a section of LIGO data from science run 2 which include both Samples H1 and H2. The significance listed in these results is the ratio of the significance generated by the appropriate algorithm to the threshold level used. In E.1 and E.2 this threshold level was constant for all frequencies, whereas in E.3 the thresholds generated by `BramThreshold` and shown in Figure 40 were used. Whether using such variable thresholds provides better results is unclear.

A start time relative to the epoch start time is shown for each candidate transient found, together with a GPS time. This relative start time would be approximately in the middle of the interval (so at about 5 seconds for E.1 and E.3, and about 30 seconds for E.2) if it were the cause of the trigger.

4 Conclusion

Of the methods considered, the Likelihood Ratio Interval Analysis based on Normal, Di-Normal, and Gamma distributions performed the best, based both on the ability to detect transients over background noise, and on the speed of calculation. Of these, the Di-Normal distribution provided the best fit to the data samples considered, the Gamma distribution gave the lowest level of background noise, and the Normal distribution gave the most pronounced detection of an anomaly obvious to the eye. These three methods were used to detect a candidate transient in a section of Hanford 4km science run 2 data that was almost certainly the cause of a large `glitch` trigger. The candidate transient was best classified assuming a Normal distribution, yielding a GPS start time of 729337275.15, a duration of 220ms, an average excess power of 3.305×10^{-3} Units^2, and a frequency range of 176Hz–400Hz. A second sample of Hanford 4km science run 2 data with a similarly large `glitch` trigger was also analyzed using these methods, but all failed to detect the cause of the trigger.

The Weibull distribution was demonstrated to give a better fit than both the Normal or Gamma distributions to the time series of power for each frequency, but was too numerically expensive to perform a Likelihood Ratio Interval Analysis with. An approximation to overcome this was attempted by estimating maximum likelihood estimators from moments of the data, but this yielded a worse fit than the Gamma distribution.

A Box-Cox transform was shown to successfully fit all frequencies to a Normal distribution, but at the cost of dramatically reducing the significance of outliers. The extend of this reduction meant that better results were obtained without the transform.

The Normal, Di-Normal, Bivariate Normal, Gamma, K-S, and A-D analyses were implemented in the trigger monitor `BramMon`, with the ability to use per-frequency thresholds. A threshold generation program called `BramThreshold` was written to provide a means of estimating thresholds from LIGO data.

5 Possibilities for Further Investigation

This report is the product of a SURF placement, the duration of which is 10 weeks. If time had permitted, some of the following would have been investigated:

- **High-pass Filtering and Strain Calibration**: This report used the raw data provided by the LIGO observatories. While strain calibration should not have any noticeable effect on the analysis involved here due to the comparably short analysis timescales, it would give more meaning to the power shown in the results. High-pass filtering would be useful in removing some of the bleeding in low frequencies of interest from the even lower, noisy frequencies present in LIGO data. This should help give better sensitivity of the algorithms at low frequencies.

- **Optimize Parameters**: The parameters used to generate the data here (such as length of data either side of a trigger to analyze, size of DFT blocks in the spectrographic analysis, etc) were chosen after a brief search because they gave good results. A detailed study of the effects of these parameters would be useful to ascertain what their optimal values are.

- **Suitable Threshold Values**: Whether the variable thresholds generated by `BramThreshold` provide better results than using a fixed threshold is unclear from this investigation, and would need further study to ascertain an optimal threshold estimation algorithm.

- **Overcoming Non-Stationarity**: So far, the null and alternative hypotheses have considered one and two sections of data respectively, each of which having distributions invariant with time. A simple extension would be to allow the parameters in both hypotheses to be "slowly varying" across the sections of data considered — this should help reduce or eliminate non-stationarity problems. However, a difficult question is how to characterize and implement "slowly varying".

- **Rice Distribution**: Sylvestre [2] suggested that a Rice distribution provided a good model for power time series of LIGO data, with parameters estimated using moments (as was considered for the Weibull distribution in A.5). It is not possible to implement a Likelihood Ratio Interval Analysis based on this distribution in less than complexity $O(n^2)$, even using the moment approximation, although how fast such a method would be in practice is hard to tell.

- **Outlier Analysis**: This report mainly concentrated on detecting intervals of interest. However, the good fit of the Gamma, Weibull, Box-Cox transformed Normal, and possibly Rice distributions could be used for Outlier Analysis. This would involve fitting one of these distribution to the whole of each frequency's time series, possibly using methods described above, and then searching for outliers on this distribution. Post-processing would need to be extended so that anomalies that are close but do not overlap in time are grouped. Non-stationarity would probably still prove to be a problem, but to less of an extent, and could well manifest as missed transients rather then false signals.

- **Wavelets**: The purpose of the spectrographic transform was to find a representation of the data where possible transients were concentrated in relatively few data segments, but where background noise was spread out. Wavelets could possibly offer an even better transform, since they are localized both in time and frequency. This would lead to better results with fewer false signals, provided distributions could be found which fitted the wavelet components well.

6 Acknowledgments

I would like to thank the following people for their help and support during this SURF project:

- Dr. John Zweizig, my mentor, for accepting me on this SURF placement, and for his time and kind help through out it.
- Dr. Szabolcs Márka, for his comments and suggestions, and for taking over as my mentor while John was away.
- Dr. Julien Sylvestre, for discussing with me his research into spectrographic analysis of LIGO data.

I would also like to thank the SURF team for making this opportunity available to me, and the NSF for funding this project.

A The Mathematics

A.1 The Box-Cox Transform

The Box-Cox [3] transform is a non-linear scaling of positive data to better fit a Normal distribution. The transformation is defined as

$$f_\lambda(x) = \begin{cases} \dfrac{x^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \log(x) & \text{for } \lambda = 0 \end{cases} \tag{1}$$

where λ is a real parameter. It can be seen to be continuous in λ, and smooth and monotonically increasing in x.

If $Y_i = f_\lambda(X_i)$, and the Y_i's were independent and Normally distributed with common mean μ and variance σ^2, then the X_i's would have p.d.f.

$$g(\boldsymbol{x} \,|\, \mu, \sigma^2, \lambda) = \left(\prod x_i\right)^{\lambda-1} \left(2\pi\sigma^2\right)^{-n/2} \exp\left\{-\frac{S_{yy} + n(\bar{y} - \mu)^2}{2\sigma^2}\right\} \tag{2}$$

where n is the number of random variables, $y_i = f_\lambda(x_i)$, and

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad\qquad S_{yy} = \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{3}$$

The maximum likelihood estimator (m.l.e.) for λ is therefore the value of λ for which

$$\ell(\lambda) \equiv (\lambda - 1)\sum_{i=1}^{n} \log(x_i) - \frac{n}{2}\log\left(\frac{S_{yy}}{n}\right) \tag{4}$$

attains its maximum.

When used, the Box-Cox transform was applied to each frequency's power time series separately. For each frequency, the maximum likelihood estimate for λ was calculated by minimizing (4) numerically, to 4 significant figures.

A.2 Likelihood Ratio Interval Analysis

Let $\boldsymbol{x}$ be an n dimensional vector denoting the data, and let x_i denote the i^{th} data point. Write $\boldsymbol{x}(t,d)$ for the vector $(x_t, x_{t+1}, \ldots, x_{t+d-1})$, and write $\boldsymbol{x}^*(t,d)$ for all data points in $\boldsymbol{x}$ not included in $\boldsymbol{x}(t,d)$. This algorithm assumes a probability distribution function (p.d.f.) $f(\boldsymbol{x} \,|\, \boldsymbol{\theta})$, with parameters $\boldsymbol{\theta}$. For example, f may be a Normal distribution p.d.f., and then $\boldsymbol{\theta} = (\mu, \sigma^2)$. The shorthand

$$\hat{f}(\boldsymbol{x}) = \sup_{\boldsymbol{\theta}} f(\boldsymbol{x} \,|\, \boldsymbol{\theta}) \tag{5}$$

is used for clarity.

The null hypothesis (H_0) is that all the data come from one common distribution.

The alternative hypothesis (H_1) is that the data come from two independent distributions, with one distribution for a small interval of the data $\boldsymbol{x}(t,d)$, and the other distribution for the rest.

The log Profile Likelihood Ratio statistic is therefore

$$L(\boldsymbol{x}) = 2\log\left(\frac{\sup\limits_{t,d}\left[\hat{f}\big(\boldsymbol{x}(t,d)\big)\hat{f}\big(\boldsymbol{x}^*(t,d)\big)\right]}{\hat{f}(\boldsymbol{x})}\right) \tag{6}$$

with the values of t and d for which the supremum is attained being the m.l.e.'s for the interval of interest. Under H_0, $L(\boldsymbol{x})$ tends to be small, whereas under H_1, larger values of $L(\boldsymbol{x})$ indicate more evidence in support of an inserted interval of a different distribution.

The above was implemented by searching though values of t and d, with d being restricted by a minimum and maximum interval size. If calculation of $\hat{f}(\boldsymbol{x})$ can be done in constant time at each step (for example, by using the moments of $\boldsymbol{x}(t, d)$, which can be updated from the moments of $\boldsymbol{x}(t, d-1)$ or $\boldsymbol{x}(t-1, d)$ in constant time), then the whole algorithm is of complexity $O(n)$.

Various distributions that were used with this method are discussed below.

A.3 The Normal Distribution for Interval Analysis

This method uses the Likelihood Ratio Interval Analysis described previously, using for f the p.d.f. of Normal, independent, identically distributed random variables. That is to say,

$$f(\boldsymbol{x} \mid \mu, \sigma^2) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left\{-\frac{S_{xx} + n(\bar{x} - \mu)^2}{2\sigma^2}\right\} \tag{7}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} {x_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 \tag{8}$$

It follows that

$$\hat{f}(\boldsymbol{x}) = \left(\frac{2\pi e S_{xx}}{n}\right)^{-n/2} \tag{9}$$

Since this only depends on the sum and the sum of the squares of the data (i.e. the moments), lengthening or moving the interval being analyzed by one point simply requires at most two data points being subtracted from one total and added to another, and hence $\hat{f}$ may be calculated in constant time.

A.4 The Gamma Distribution for Interval Analysis

If the original time series data were distributed Normally, then the coefficients of the DFT would also be distributed Normally. Hence, the power (which is proportional to the sum of the squares of the real and imaginary components of the DFT) would be distributed as a scaled χ^2_2 distribution. The Gamma distribution is a generalization of a scaled χ^2 distribution, and the above would correspond to a shape parameter $\alpha = 1$. This motivated the use of a Gamma distribution.

This method again uses the Likelihood Ratio Interval Analysis above, this time with f being the p.d.f. of independent, identically distributed gamma random variables. This gives

$$f(\boldsymbol{x} \mid \alpha, \beta) = \frac{\beta^{n\alpha}\left(\prod x_i\right)^{\alpha-1} e^{-n\beta\bar{x}}}{\Gamma(\alpha)^n} \tag{10}$$

It follows that

$$\hat{f}(\boldsymbol{x}) = \frac{1}{\Gamma(\hat{\alpha})^n} \exp\left\{n\hat{\alpha}\left(\log\frac{\hat{\alpha}}{\bar{x}} - 1\right) + (\hat{\alpha} - 1)\sum_{i=1}^{n}\log x_i\right\} \tag{11}$$

where $\hat{\alpha}$ is given by

$$\log\hat{\alpha} - \Phi(\hat{\alpha}) = \log\bar{x} - \frac{1}{n}\sum_{i=1}^{n}\log x_i \equiv e^{\ell} \tag{12}$$

and Φ is the digamma function, which is the derivative of the log of the gamma function.

Once again, since the above only depends on the sum and the sum of the logs of the data, $\hat{f}$ may be calculated in constant time. However, computation of $\hat{\alpha}$ is computationally quite expensive. A good approximation to (12) in the range of interest is

$$\hat{\alpha} \approx \exp\left\{a\ell^3 + b\ell^2 + c\ell + d\right\} \tag{13}$$

where

$$a = -0.00560202 \qquad b = 0.0143161 \qquad c = -0.872786 \qquad d = -0.484827 \tag{14}$$

This approximation was used for the results in this report.

A.5 The Weibull Distribution for Interval Analysis

Another distribution similar in shape to the Gamma distribution is the Weibull Distribution. When its shape parameter $\alpha = 1$, the distribution is identical to the Gamma distribution with $\alpha = 1$.

The Weibull Distribution has p.d.f.

$$f(\boldsymbol{x} \mid \alpha, \beta) = \alpha^n \beta^{-n\alpha} \left(\prod x_i\right)^{\alpha-1} \exp\left\{-\sum \left(\frac{x_i}{\beta}\right)^{\alpha}\right\} \tag{15}$$

Which gives

$$\hat{f}(\boldsymbol{x}) = f(\boldsymbol{x} \mid \hat{\alpha}, \hat{\beta}) \tag{16}$$

where

$$\hat{\beta}^{\hat{\alpha}} = \frac{1}{n}\sum_{i=1}^{n} x_i^{\hat{\alpha}} \tag{17}$$

and $\hat{\alpha}$ is a solution to

$$g(\alpha) \equiv \frac{\sum x_i^{\alpha} \log x_i}{\sum x_i^{\alpha}} - \frac{1}{\alpha} - \sum \log x_i = 0 \tag{18}$$

As $\alpha \to 0$, $g(\alpha) \to -\infty$, and as $\alpha \to \infty$, $g(\alpha) \to \max\{\log x_i\} - \frac{1}{n}\sum \log x_i > 0$. The derivative of g is

$$g'(\alpha) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{i} \big(\log(x_i) - \log(x_j)\big)^2 x_i^{\alpha} x_j^{\alpha}}{\left(\sum_{i=1}^{n} x_i^{\alpha}\right)^2} + \frac{1}{\alpha^2} > 0 \tag{19}$$

Hence, $g(\alpha)$ is shown to be strictly increasing, so equation (18) has exactly one solution.

This was implemented by using a numerical search for $\hat{\alpha}$. Unfortunately, not only is this search slow, it also has complexity $O(n)$, giving the overall Likelihood Ratio Interval Search a complexity of $O(n^2)$, which is far slower than the other parametric searches considered. An attempt to evaluate $\hat{f}$ in constant time was made by approximating $\hat{\alpha}$ and $\hat{\beta}$ using moments. The raw moments of a Weibull distribution are

$$\mu_1' = \beta\Gamma\big(1 + \alpha^{-1}\big) \tag{20}$$

$$\mu_2' = \beta^2\Gamma\big(1 + 2\alpha^{-1}\big) \tag{21}$$

which suggested using the approximation

$$\hat{\beta} = \frac{\frac{1}{n}\sum x_i}{\Gamma(1 + \hat{\alpha}^{-1})} \tag{22}$$

where $\hat{\alpha}$ is the solution of

$$h(\alpha) \equiv \log\Gamma\big(1+2\alpha^{-1}\big) - 2\log\Gamma\big(1+\alpha^{-1}\big) = \log\left(\frac{1}{n}\sum {x_i}^2\right) - 2\log\left(\frac{1}{n}\sum x_i\right) \tag{23}$$

A reasonable approximation to $h(\alpha)$ is

$$h(\alpha) \approx \frac{1}{a\alpha^2 + b\alpha} \tag{24}$$

where

$$a = e^{-0.5} \qquad\qquad b = e^{-0.3265} \tag{25}$$

This gives

$$\hat{\alpha} \approx \frac{1}{2a}\left(b^2 + \frac{4a}{\log\left(\frac{1}{n}\sum {x_i}^2\right) - 2\log\left(\frac{1}{n}\sum x_i\right)}\right)^{1/2} - \frac{b}{2a} \tag{26}$$

Since equations (22) and (26) involve only the sum and sum of the squares of the data, they can be used to approximate $\hat{f}$ in constant time.

Unfortunately, approximating $\hat{\alpha}$ and $\hat{\beta}$ using moments caused the Weibull distribution to show a much worse fit to the data than the correctly calculated values, to such an extent that the Gamma distribution became a better alternative. Hence, this method of approximating $\hat{\alpha}$ and $\hat{\beta}$ was abandoned; the results shown in this report were generated using the correctly calculated $\hat{f}$.

A.6 The Di-Normal and Bivariate Normal Distributions for Interval Analysis

As noted above for the Gamma distribution, if the time series data were Normally distributed, then the individual real and imaginary components of the DFT would also be Normally distributed. This method of analysis uses the Likelihood Ratio Interval Analysis on the time series generated by the real and imaginary parts of the DFT coefficients of the spectrographic transform; each frequency is considered independently.

Write $\boldsymbol{x}$ for the series of real parts and $\boldsymbol{y}$ for the imaginary parts. In both the Di-Normal and Bivariate Normal distributions, it is assumed that each data point is independent from all other data points, apart from possibly the corresponding real or imaginary part with the same index.

For the Bivariate Normal distribution, the real and imaginary parts are allowed to be correlated. This yields a p.d.f.

$$f(\boldsymbol{x},\boldsymbol{y}\,|\,\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\sigma_{12}) = \Big(4\pi^2\big(\sigma_1^2\sigma_2^2 - \sigma_{12}^2\big)\Big)^{-n/2}\exp\left\{-\frac{Z}{2\big(\sigma_1^2\sigma_2^2 - \sigma_{12}^2\big)}\right\} \tag{27}$$

where

$$Z = \sigma_2^2\big(S_{xx} + n(\bar{x}-\mu_1)^2\big) + \sigma_1^2\big(S_{yy} + n(\bar{y}-\mu_2)^2\big) - 2\sigma_{12}\big(S_{xy} + n(\bar{x}-\mu_1)(\bar{y}-\mu_2)\big) \tag{28}$$

and

$$S_{xy} = \sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) \tag{29}$$

This gives

$$\hat{f}(\boldsymbol{x},\boldsymbol{y}) = \left(\frac{4\pi^2 e^2}{n^2}\left(S_{xx}S_{yy} - S_{xy}^2\right)\right)^{-n/2} \tag{30}$$

As before, this only involves the sum and the sum of the squares of the two data sets, and the sum of the cross terms, and hence $\hat{f}$ may be calculated in constant time.

For the Di-Normal distribution, the real and imaginary parts of each coefficient are assumed to be independent, which forces $\sigma_{12} = 0$ in equations (27) and (28). This produces

$$\hat{f}(\boldsymbol{x},\boldsymbol{y}) = \left(\frac{4\pi^2 e^2}{n^2} S_{xx}S_{yy}\right)^{-n/2} \tag{31}$$

again giving $\hat{f}$ in constant time.

A.7 The Shifted Gamma Distribution for Interval Analysis

If X has a Gamma distribution with parameters α and β, then $Y = d + X$ has a shifted Gamma distribution with parameters α, β, and d. The p.d.f. for a sample of n independent Shifted Gamma random variables is

$$f(\boldsymbol{x} \,|\, \alpha, \beta, d) = \frac{\beta^{n\alpha} \left(\prod (x_i - d) \right)^{\alpha - 1} e^{-n\beta(\bar{x} - d)}}{\Gamma(\alpha)^n} \tag{32}$$

which gives

$$\hat{f}(\boldsymbol{x}) = f\left(\boldsymbol{x} \,|\, \hat{\alpha}, \hat{\beta}, \hat{d}\right) \tag{33}$$

where

$$\hat{\beta} = \frac{\hat{\alpha}}{\bar{x} - \hat{d}} \qquad\qquad \hat{\alpha} = \frac{\sum \dfrac{\bar{x} - \hat{d}}{x_i - \hat{d}}}{\sum \dfrac{\bar{x} - x_i}{x_i - \hat{d}}} \tag{34}$$

and $\hat{d}$ is the solution to the equation

$$\log\left(\frac{\sum \dfrac{1}{x_i - d}}{\sum \dfrac{\bar{x} - x_i}{x_i - d}} \right) + \frac{1}{n} \sum \log(x_i - d) = \Phi\left(\frac{\sum \dfrac{\bar{x} - d}{x_i - d}}{\sum \dfrac{\bar{x} - x_i}{x_i - d}} \right) \tag{35}$$

This gives calculating $\hat{f}$ a complexity of $O(n)$, which in practice takes longer than accurately calculating $\hat{f}$ for the Weibull distribution.

Similarly to the Weibull distribution, one possibility considered to calculate $\hat{f}$ in constant time was to estimate $\hat{\alpha}$, $\hat{\beta}$, and $\hat{d}$ from the moments of the data. The mean (μ), variance (σ^2), and skewness (γ_1) for a Shifted Gamma distribution are

$$\mu = \frac{\alpha}{\beta} + d \qquad\qquad \sigma^2 = \frac{\alpha}{\beta^2} \qquad\qquad \gamma_1 = \frac{2}{\sqrt{\alpha}} \tag{36}$$

which suggested using the approximation

$$\hat{d} = \bar{x} - \frac{2{S_{xx}}^2}{nS_{xxx}} \qquad\qquad \hat{\beta} = 2\frac{S_{xx}}{S_{xxx}} \qquad\qquad \hat{\alpha} = \frac{4{S_{xx}}^3}{n{s_{xxx}}^2} \tag{37}$$

where

$$S_{xxx} = \frac{1}{n} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^3 = \frac{2}{n^3} \left(\sum x_i \right)^3 - \frac{3}{n^2} \left(\sum x_i \right) \left(\sum {x_i}^2 \right) + \frac{1}{n} \sum {x_i}^3 \tag{38}$$

Of course, Gamma random variables are never negative, which adds the constraint $\hat{d} < \min\{x_1, \ldots, x_n\}$. This approximation in practice gave a worse fit to the data than the non-shifted Gamma distribution!

Another approach considered was the crude approximation $\hat{d} = \min\{x_1, \ldots, x_n\}$. This gave identical results to the non-shifted Gamma distribution when applied to LIGO data.

A.8 The Gamma/Normal Distribution for Interval Analysis

A variant on the Gamma and Normal Likelihood Ratio Interval Analyses is the Gamma/Normal Interval Analysis. For this, the null hypothesis (H_0) is that all the data have a Gamma distribution, and the alternative hypothesis (H_1) is that all but an interval have a Gamma distribution, while the data within that interval have a Normal distribution.

This gives a log Profile Likelihood Ratio of

$$L(\boldsymbol{x}) = 2\log\left(\frac{\sup\limits_{t,d}\left[\hat{f}_n\big(\boldsymbol{x}(t,d)\big)\hat{f}_g\big(\boldsymbol{x}^*(t,d)\big)\right]}{\hat{f}_g\big(\boldsymbol{x}\big)}\right) \tag{39}$$

where $\hat{f}_n$ is $\hat{f}$ for a Normal distribution (see A.3), and $\hat{f}_g$ is $\hat{f}$ for a Gamma distribution (see A.4).

In practice, this test gave identical results to the Gamma Likelihood Ratio Interval Analysis when used on the LIGO data samples considered.

A.9 Kolmogorov-Smirnov and Anderson-Darling Interval Analysis

The Kolmogorov-Smirnov [4] test (referred to as the K-S test) is a non-parametric test that uses an empirical distribution function from the data to estimate the underlying distribution. The test used here is the two-tailed, two-sample test, described in *A Knowledgebase for Extragalactic Astronomy and Cosmology* [6].

The empirical distribution function for a sample $\boldsymbol{x}$ of n points is

$$F_{\boldsymbol{x}}(t) = \sum_{i=1}^{n} I[x_i \leq t] \tag{40}$$

where $I[test]$ denotes an indicator function, taking the value 1 if *test* is true, or 0 if *test* is false. The empirical distribution function is therefore the cumulative probability function for drawing a point at random from the sample $\boldsymbol{x}$.

The K-S statistic for testing if two data sets (denoted $\boldsymbol{x}$ and $\boldsymbol{y}$, consisting of n and m points respectively) are drawn from the same distribution is

$$K(\boldsymbol{x},\boldsymbol{y}) = \frac{nm}{\sqrt{n+m}}\max\left\{\max_{i=1,\ldots,n}\left|F_{\boldsymbol{x}}(x_i) - F_{\boldsymbol{y}}(x_i)\right|, \max_{j=1,\ldots,m}\left|F_{\boldsymbol{x}}(y_j) - F_{\boldsymbol{y}}(y_j)\right|\right\} \tag{41}$$

This statistic would be zero for a perfect fit, and tends to be large for a worse fit. The critical values do not depend on the distribution being sampled from, and are approximately 1.4 for a 95% confidence level, or 2.0 for a 99.9% confidence level.

The algorithm implemented scans through the data to find the interval with the largest K-S statistic, in a similar manner to the Likelihood Ratio Interval Analysis. However, calculating the K-S statistic for each interval has complexity $O(n)$, so the total complexity of this test is $O(n^2)$, which is much worse than the parametric tests above. Although noticeably slower, the K-S test still seemed to perform fast enough for 60 seconds of data to be usable.

The Anderson-Darling [5] test (referred to as the A-D test) is similar to the K-S test, but gives more weight to the tails of the distribution. Denoting by $\boldsymbol{x} \cup \boldsymbol{y}$ the sample consisting of all points in $\boldsymbol{x}$ and all points in $\boldsymbol{y}$, and by $x_{(i)}$ the i^{th} smallest element of $\boldsymbol{x}$, the square of the A-D statistic is

$$A(\boldsymbol{x},\boldsymbol{y})^2 = -n - \sum_{i=1}^{n}\frac{2i-1}{d}\Big[\log\big(F_{\boldsymbol{x}\cup\boldsymbol{y}}(x_{(i)})\big) + \log\big(1 - F_{\boldsymbol{x}\cup\boldsymbol{y}}(x_{(n+1-i)})\big)\Big] \tag{42}$$

This also yields a search algorithm of complexity $O(n^2)$, and in practice performed only slightly slower than the K-S Interval Analysis.

A.10 Kolmogorov-Smirnov Goodness of Fit Test

To test how well a distribution with cumulative distribution function $F(x \,|\, \boldsymbol{\theta})$ fitted data $\boldsymbol{x}$ (consisting of n samples), a generalized Kolmogorov-Smirnov [4] goodness of fit test was used. This test is similar to the K-S two sample test discussed above. Given an estimator $\hat{\boldsymbol{\theta}}(\boldsymbol{x})$ for $\boldsymbol{\theta}$, the generalized K-S goodness of fit statistic is

$$K = \sqrt{n} \max_{i=1,\ldots,n} \left\{ \left| \frac{i}{n} - F\big(x_{(i)} \,|\, \hat{\boldsymbol{\theta}}(\boldsymbol{x})\big) \right| \right\} \tag{43}$$

The larger the value of K, the worse the fit, and a value of 0 indicates a perfect fit. The test is "generalized", as the classical K-S test does not allow $F(x \,|\, \boldsymbol{\theta})$ to have parameters best fitted from the data. Allowing this only changes the critical values, which then become dependent on the distribution being tested against. From numerical Monte Carlo simulation, an approximate 99.99% confidence level for all the parametric distributions considered was determined to be 1.4.

The estimator $\hat{\boldsymbol{\theta}}$ used for the results in this report was the same estimator as was used for the Likelihood Ratio Interval Analysis, so that the K-S statistic gave an indication of how well both the distribution and the estimate of $\boldsymbol{\theta}$ performed during Likelihood Ratio Interval Analysis.

B Data Samples

B.1 Sample H1

Sample H1 was taken from the Hanford 4km Interferometer dark port (`H1:LSC-AS_Q`). It consisted of 60 seconds of data, starting from GPS time 729337245, and was chosen because a `glitch` trigger occurred 30.22 seconds into the sample with a strength of 7.94. This sample formed part of LIGO science run 2.

B.2 Sample H2

Sample H2 was similarly taken from the Hanford 4km Interferometer dark port (`H1:LSC-AS_Q`). It consisted of 60 seconds of data, starting from GPS time 729337710, and was chosen because a `glitch` trigger occurred 29.43 seconds into the sample with a strength of 6.84. This sample formed part of LIGO science run 2.

B.3 Sample S1

Sample S1 was a 60 second sample of simulated data, with the following components:

- Normally distributed noise with variance 1.0 through out the sample.
- A 30Hz sine wave with amplitude 100 through out the sample.
- A 1kHz sine Gaussian centered at 31 seconds, with s.d. width 50ms and maximum amplitude 0.7.
- A 4kHz sine Gaussian centered at 25 seconds, with s.d. width 10ms and maximum amplitude 1.0.
- A 4kHz sine Gaussian centered at 26 seconds, with s.d. width 10ms and maximum amplitude 0.8.

C Tables

List of Tables

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
64	33.87	34.12	640.6
64	38.05	38.21	514.9
192	30.21	30.30	475.1
224	30.20	30.35	2934.
256	30.20	30.37	3899.
256	33.43	33.55	495.4
288	30.20	30.35	4024.
288	33.43	33.55	1638.
320	30.21	30.35	2536.
320	33.54	33.70	667.6
352	30.21	30.35	1482.
352	33.66	33.70	338.8
384	30.20	30.35	1461.
384	33.59	33.70	1256.
576	33.59	33.70	389.6
608	33.54	33.70	915.6
640	33.49	33.70	316.7
672	30.20	30.29	502.6
704	30.15	30.29	479.3
768	33.68	33.71	533.2
800	33.54	33.70	429.7
800	5.461	5.617	360.8
928	5.742	5.867	365.1
960	5.727	5.977	573.7
1024	5.414	5.664	438.1
1056	5.555	5.773	711.2
1088	5.352	5.602	535.0
1088	5.633	5.883	385.1
1120	1.430	1.617	310.3
1216	1.398	1.648	322.2
1248	1.367	1.602	422.4
1280	5.555	5.805	382.6
1312	5.461	5.695	366.4
1376	5.430	5.680	573.9
1408	5.414	5.648	813.9
1440	5.414	5.648	502.7
1472	1.352	1.586	497.4
1504	1.336	1.586	434.6
1536	1.336	1.586	422.1
1536	5.336	5.586	337.8
1568	1.445	1.695	585.2
1632	5.336	5.570	422.8
1664	5.320	5.539	686.1
1664	5.664	5.914	362.2
1696	5.367	5.617	393.3
1696	5.773	5.930	619.3
1728	1.398	1.555	488.6
1728	5.773	5.977	324.7

Table 1: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Normal distribution, on Sample H1. A significance threshold of 300 was used. Continues on next page.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
1760	5.211	5.461	334.9
1824	5.430	5.664	623.2
1824	34.96	35.21	355.5
1856	5.523	5.773	691.2
1856	5.805	5.930	372.4
1856	34.96	35.21	399.1
1888	5.523	5.773	573.2
1952	5.523	5.773	336.6
1984	1.336	1.586	403.2
2016	1.383	1.617	452.5
2016	1.742	1.898	445.6
2048	1.367	1.617	343.6
2112	5.398	5.648	300.3
2144	1.352	1.555	576.7
2144	41.40	41.59	323.5
2240	5.430	5.680	356.3

Table 1 Continued.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
0	42.68	42.93	328.2
32	42.68	42.93	428.0
64	42.66	42.91	480.9
64	39.73	39.98	420.0
64	43.74	43.98	334.3
64	33.05	33.21	366.9
96	42.66	42.91	365.7
992	57.52	57.77	377.6
992	35.16	35.41	439.0
1024	35.20	35.40	587.4
1152	35.21	35.46	334.2
1312	57.48	57.73	355.4
1408	35.24	35.49	342.5
1664	59.10	59.34	344.2
1664	57.38	57.63	343.2
1696	59.16	59.32	327.8
1792	59.18	59.30	302.0
1824	57.45	57.70	345.0
1856	57.55	57.80	404.7
2048	57.51	57.76	355.0
2496	28.48	28.68	344.1
2528	28.52	28.62	379.3
7136	23.37	23.62	320.7
7168	23.37	23.62	458.0

Table 2: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Normal distribution, on Sample H2. A significance threshold of 300 was used.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
992	30.91	31.09	4114.
1024	30.91	31.09	1671.
3968	24.99	25.02	426.2
4000	24.98	25.02	2262.
4000	25.98	26.02	1217.
4032	24.98	25.02	400.9

Table 3: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Normal distribution, on Sample S1. A significance threshold of 300 was used.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
64	33.88	34.12	105.0
224	30.20	30.35	181.4
256	30.20	30.37	232.7
288	30.20	30.35	331.3
288	33.41	33.55	112.8
320	30.20	30.35	161.0
384	30.18	30.35	138.1
1056	5.555	5.773	130.6
1088	5.555	5.773	121.3
1376	5.414	5.648	110.4
1408	5.414	5.648	137.0
1440	5.414	5.648	122.5
1568	1.336	1.570	126.2
1664	5.320	5.539	103.9
1824	5.430	5.664	112.8
1856	5.461	5.695	106.6
1888	5.383	5.617	122.2

Table 4: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Gamma distribution, on Sample H1. A significance threshold of 100 was used.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
64	43.74	43.98	118.7
64	42.66	42.84	136.0
7136	23.37	23.60	108.0
7168	23.37	23.60	144.7

Table 5: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Gamma distribution, on Sample H2. A significance threshold of 100 was used.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
992	30.91	31.09	361.1
1024	30.91	31.09	180.8
4000	24.98	25.02	144.4

Table 6: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Gamma distribution, on Sample S1. A significance threshold of 100 was used.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
64	33.87	34.12	133.9
224	30.20	30.35	235.4
256	30.20	30.37	349.7
288	30.20	30.35	385.6
288	33.43	33.55	130.3
320	30.20	30.35	147.5
384	30.18	30.35	154.1
960	5.727	5.977	136.5
1056	5.555	5.773	150.7
1088	5.555	5.773	139.3
1216	1.398	1.648	106.3
1280	5.555	5.805	102.1
1376	5.430	5.680	148.6
1408	5.414	5.648	175.9
1440	5.414	5.648	130.4
1472	1.352	1.602	106.3
1504	1.336	1.586	110.2
1536	1.336	1.586	133.3
1536	5.336	5.586	120.6
1568	1.336	1.586	161.1
1664	5.289	5.539	133.8
1664	5.664	5.914	102.2
1696	1.367	1.617	118.7
1696	5.367	5.617	125.5
1696	5.711	5.930	101.1
1728	1.352	1.555	108.9
1824	5.414	5.664	140.6
1856	5.523	5.773	130.5
1888	5.383	5.617	139.4
2016	1.383	1.617	104.3
2144	1.352	1.602	110.1

Table 7: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Di-Normal distribution, on Sample H1. A significance threshold of 100 was used.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
64	43.73	43.98	142.2
64	42.59	42.84	156.3
64	39.73	39.98	103.8
7168	23.37	23.62	116.3

Table 8: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Di-Normal distribution, on Sample H2. A significance threshold of 100 was used.

Frequency (in Hz)	Start Time (in s)	End Time (in s)	Significance
992	30.91	31.09	378.6
1024	30.91	31.09	185.0
4000	24.98	25.02	155.7

Table 9: Results of recursively applying a Likelihood Ratio Interval Analysis, assuming a Di-Normal distribution, on Sample S1. A significance threshold of 100 was used.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
1.336	1.695	585.3	$9.48\ \times 10^{-5}$	1104 - 1136: 0.05494
.	.	.	.	1200 - 1264: 0.1297
.	.	.	.	1456 - 1584: 0.3697
.	.	.	.	1712 - 1744: 0.0839
.	.	.	.	1968 - 2064: 0.2672
.	.	.	.	2128 - 2160: 0.09461
1.742	1.898	437.3	9.808×10^{-6}	2000 - 2032:
5.211	5.977	813.4	1.375×10^{-4}	784 - 816: 0.02538
.	.	.	.	912 - 976: 0.05586
.	.	.	.	1008 - 1104: 0.1034
.	.	.	.	1264 - 1328: 0.07431
.	.	.	.	1360 - 1456: 0.1980
.	.	.	.	1520 - 1552: 0.04523
.	.	.	.	1616 - 1776: 0.2461
.	.	.	.	1808 - 1904: 0.1662
.	.	.	.	1936 - 1968: 0.03992
.	.	.	.	2224 - 2256: 0.04570
30.15	30.37	4025.	3.305×10^{-3}	176 - 400: 0.9951
.	.	.	.	656 - 720: 0.004916
33.43	33.71	1639.	5.073×10^{-4}	240 - 400: 0.9387
.	.	.	.	560 - 656: 0.03929
.	.	.	.	752 - 816: 0.02201
33.87	34.12	640.7	4.156	48 - 80:
34.96	35.21	399.1	1.503×10^{-5}	1808 - 1872:
38.05	38.21	515.	4.586	48 - 80:
41.4	41.59	323.1	1.043×10^{-5}	2128 - 2160:

Table 10: Results of post-processing Table 1.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
23.37	23.62	458.1	5.170×10^{-4}	7120 - 7184:
28.48	28.68	379.3	2.343×10^{-5}	2480 - 2544:
33.05	33.21	366.9	4.812	48 - 80:
35.16	35.49	587.5	2.909×10^{-5}	976 - 1040: 0.4060
.	.	.	.	1136 - 1168: 0.2358
.	.	.	.	1392 - 1424: 0.3582
39.73	39.98	420.	5.055	48 - 80:
42.66	42.93	481.	204.7	0 - 112:
43.74	43.98	334.7	7.721	48 - 80:
57.38	57.8	404.8	4.018×10^{-5}	976 - 1008: 0.1222
.	.	.	.	1296 - 1328: 0.1502
.	.	.	.	1648 - 1680: 0.1751
.	.	.	.	1808 - 1872: 0.3496
.	.	.	.	2032 - 2064: 0.2029
59.1	59.34	344.2	2.781×10^{-5}	1648 - 1712: 0.6894
.	.	.	.	1776 - 1808: 0.3106

Table 11: Results of post-processing Table 2.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
24.98	25.02	2263.	0.1992	3952 - 4048:
25.98	26.02	1217.	0.07300	3984 - 4016:
30.91	31.09	4115.	0.1317	976 - 1040:

Table 12: Results of post-processing Table 3.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
1.336	1.586	126.2	2.173×10^{-5}	1520 - 1584:
5.32	5.773	136.9	8.117×10^{-5}	1040 - 1104: 0.1595
.	.	.	.	1360 - 1456: 0.3990
.	.	.	.	1648 - 1680: 0.1015
.	.	.	.	1808 - 1904: 0.3400
30.18	30.37	331.4	2.459×10^{-3}	208 - 336: 0.9645
.	.	.	.	368 - 400: 0.03552
33.41	33.55	112.9	2.049×10^{-4}	272 - 304:
33.87	34.12	111.3	4.156	48 - 80:

Table 13: Results of post-processing Table 4.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
23.37	23.62	146.9	5.170×10^{-4}	7120 - 7184:
42.66	42.84	136.1	9.160	48 - 80:
43.73	43.98	120.	7.432	48 - 80:

Table 14: Results of post-processing Table 5.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
24.98	25.02	144.4	0.1191	3984 - 4016:
30.91	31.09	361.2	0.1317	976 - 1040:

Table 15: Results of post-processing Table 6.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
1.336	1.648	161.1	7.974×10^{-5}	1200 - 1232: 0.07553
.	.	.	.	1456 - 1584: 0.4729
.	.	.	.	1680 - 1744: 0.2173
.	.	.	.	2000 - 2032: 0.1111
.	.	.	.	2128 - 2160: 0.1232
5.289	5.977	175.8	9.352×10^{-5}	944 - 976: 0.04992
.	.	.	.	1040 - 1104: 0.1076
.	.	.	.	1264 - 1296: 0.06501
.	.	.	.	1360 - 1456: 0.2821
.	.	.	.	1520 - 1552: 0.07084
.	.	.	.	1648 - 1712: 0.1650
.	.	.	.	1808 - 1904: 0.2595
30.18	30.37	385.6	2.459×10^{-3}	208 - 336: 0.9645
.	.	.	.	368 - 400: 0.03552
33.43	33.55	130.3	2.219×10^{-4}	272 - 304:
33.87	34.12	133.9	4.156	48 - 80:

Table 16: Results of post-processing Table 7.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
23.37	23.62	116.3	3.467×10^{-4}	7152 - 7184:
39.73	39.98	103.8	5.055	48 - 80:
42.59	42.84	156.3	7.01	48 - 80:
43.73	43.98	142.2	7.432	48 - 80:

Table 17: Results of post-processing Table 8.

Start Time (in s)	End Time (in s)	Significance	Excess Power	Frequency composition
24.98	25.02	155.6	0.1191	3984 - 4016:
30.91	31.09	378.7	0.1317	976 - 1040:

Table 18: Results of post-processing Table 9.

D Figures

List of Figures

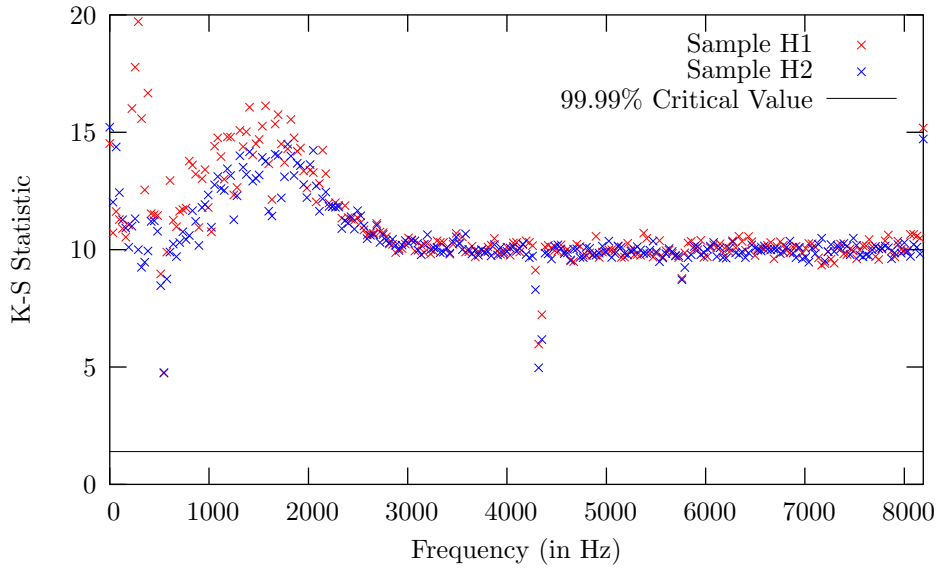

Figure 1: The goodness of fit of a Normal distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

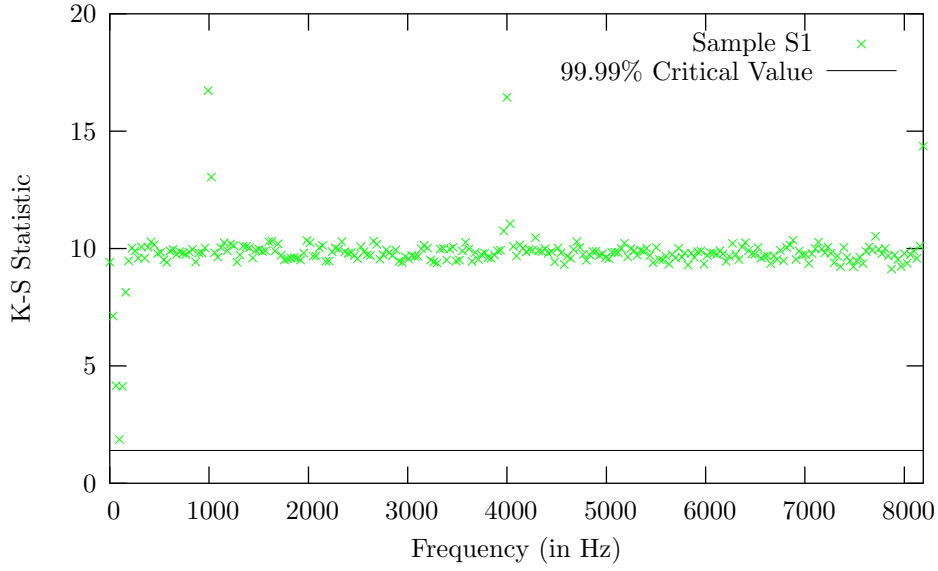

Figure 2: The goodness of fit of a Normal distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

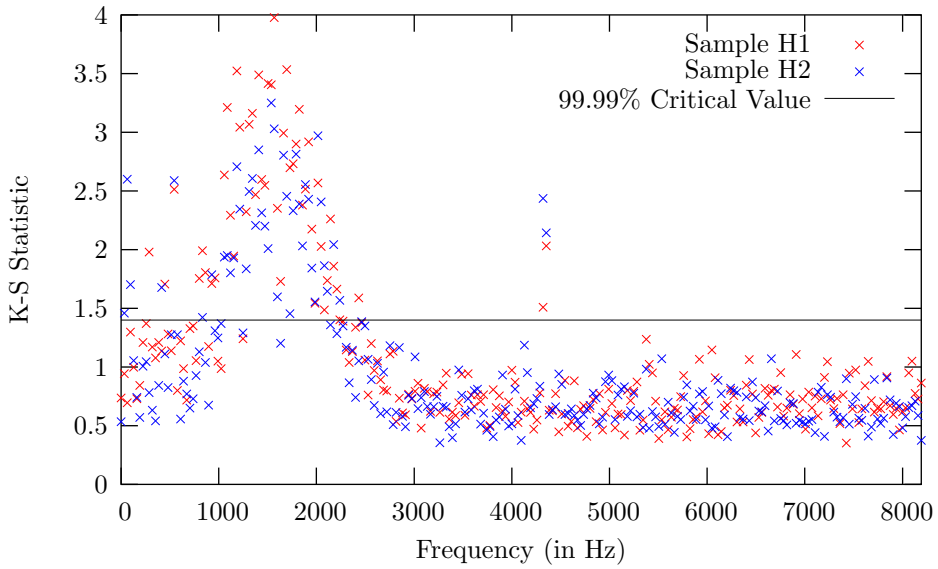

Figure 3: The goodness of fit of a Gamma distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

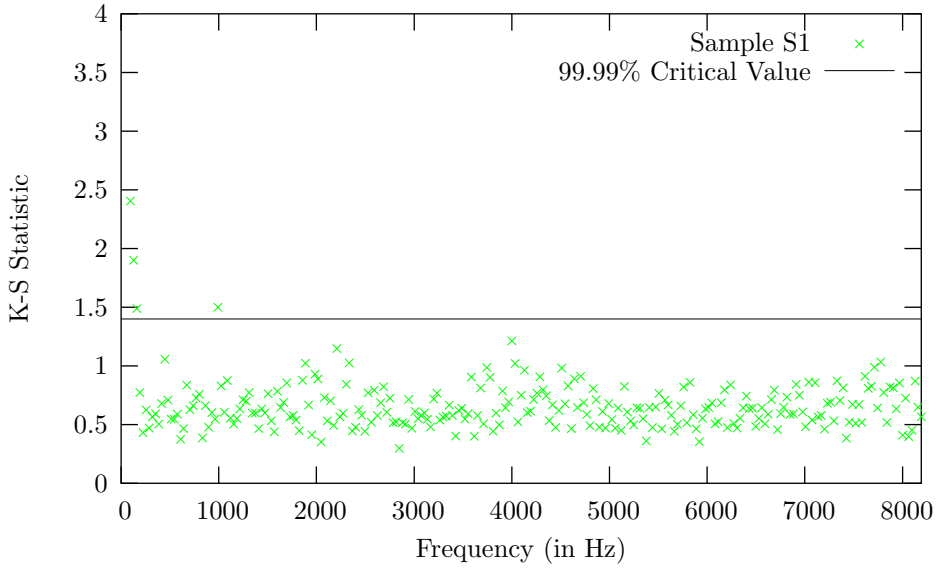

Figure 4: The goodness of fit of a Gamma distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

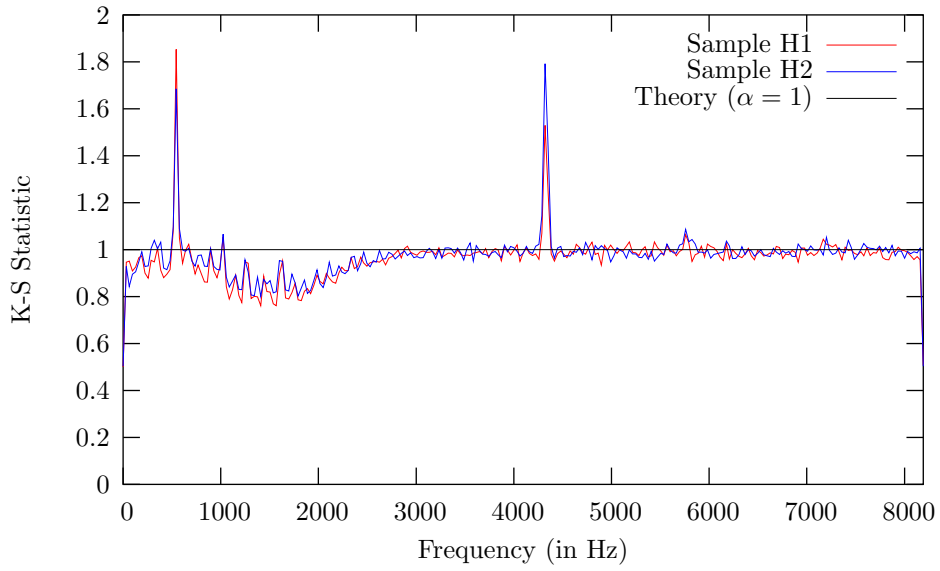

Figure 5: The α values used to fit a Gamma distribution to the power time series of various frequencies. These values were used to construct Figure 3.

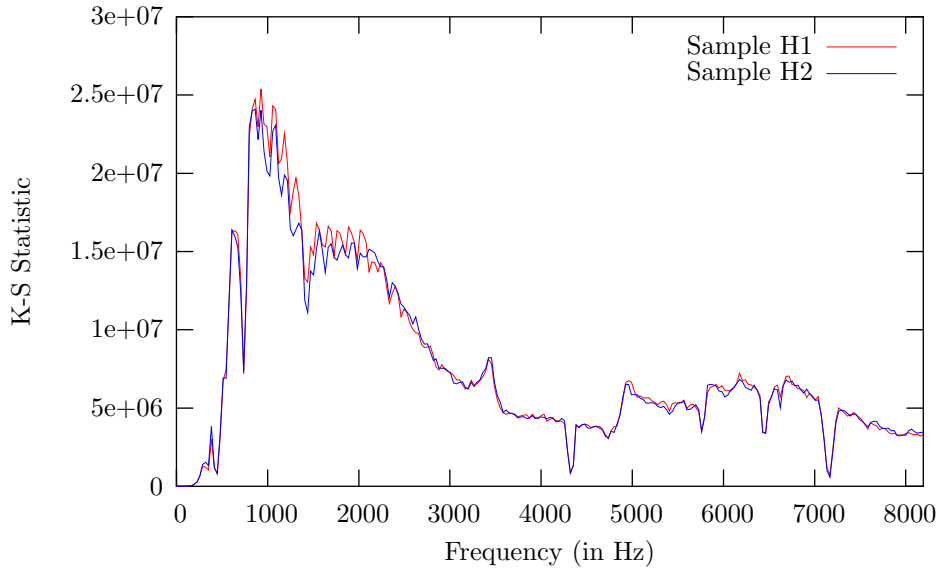

Figure 6: The β values used to fit a Gamma distribution to the power time series of various frequencies. These values were used to construct Figure 3.

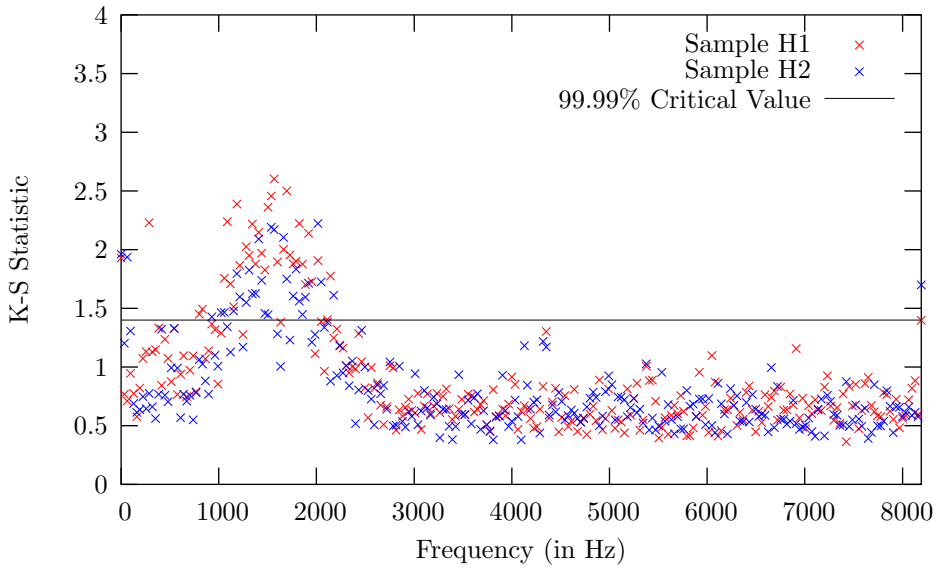

Figure 7: The goodness of fit of a Weibull distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

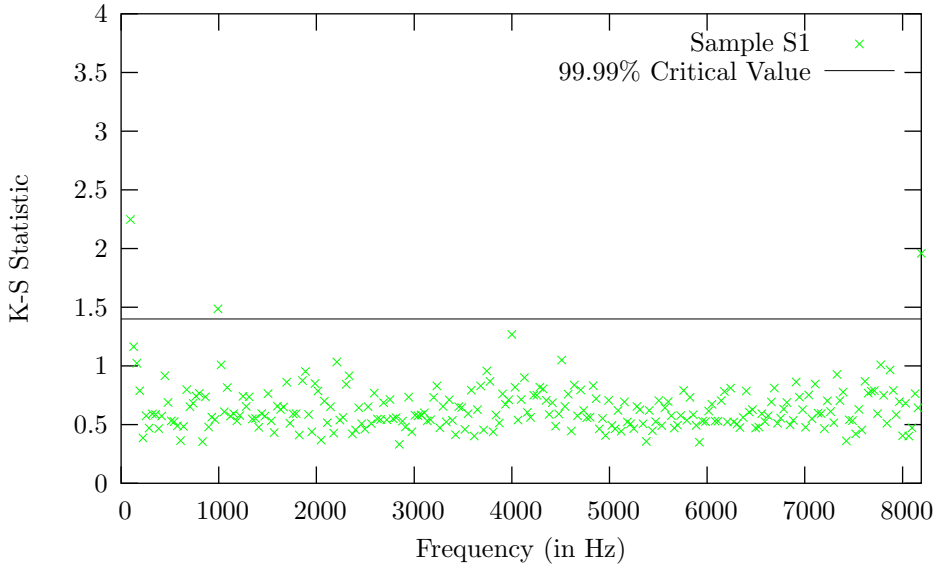

Figure 8: The goodness of fit of a Weibull distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

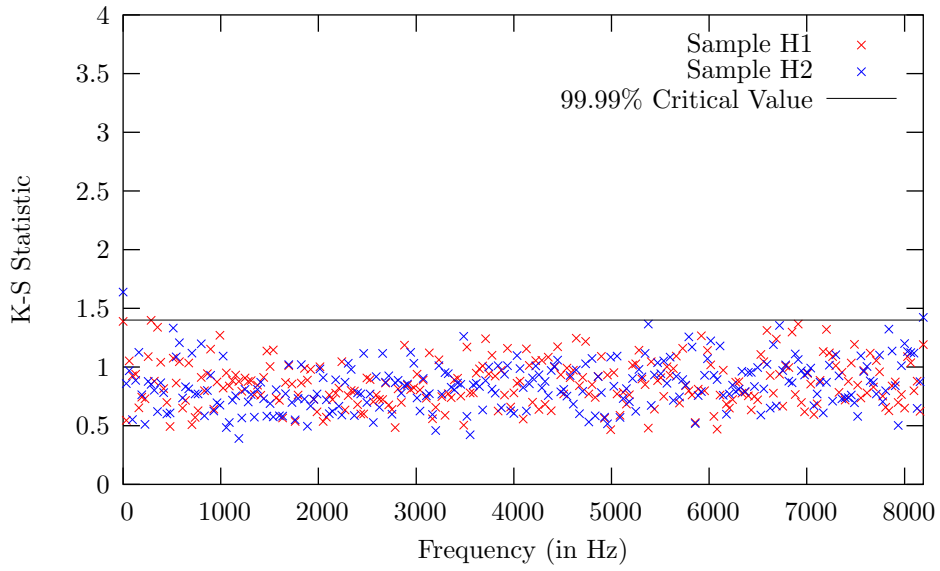

Figure 9: The goodness of fit of a Normal distribution to the Box-Cox transformed power time series of various frequencies, from a spectrographic transform of width 512.

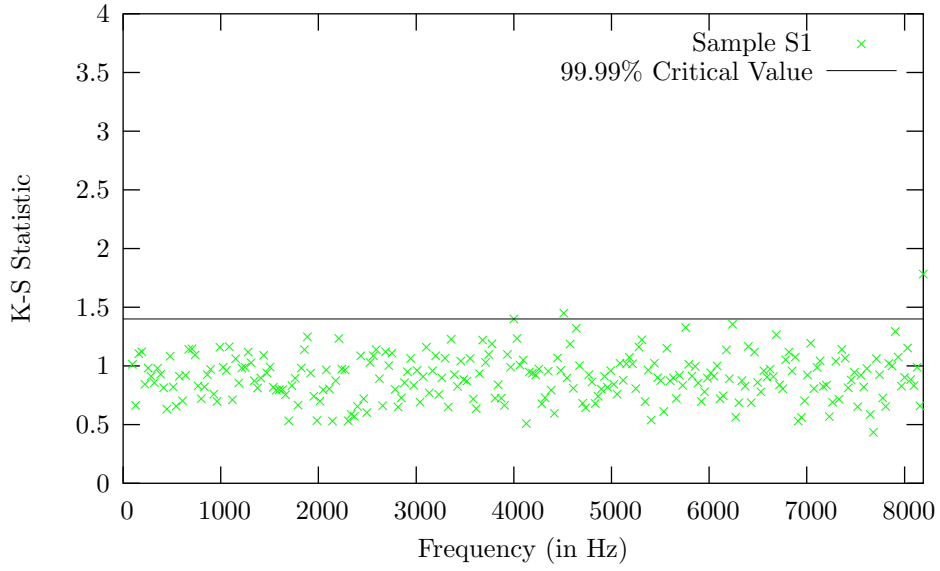

Figure 10: The goodness of fit of a Normal distribution to the Box-Cox transformed power time series of various frequencies, from a spectrographic transform of width 512.

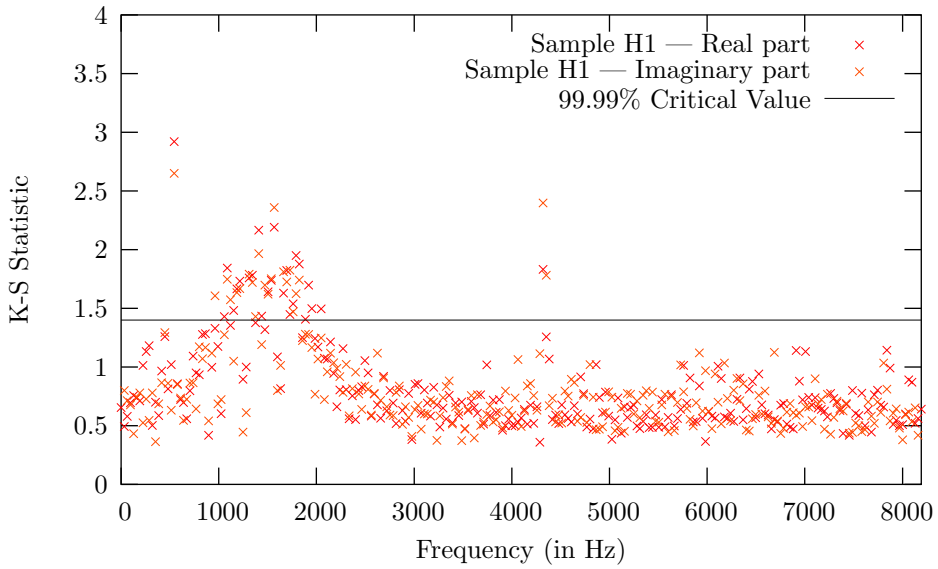

Figure 11: The goodness of fit of a Normal distribution to the time series of real and imaginary parts of the DFT coefficients from a spectrographic transform of Sample H1, using a width of 512, at various frequencies.

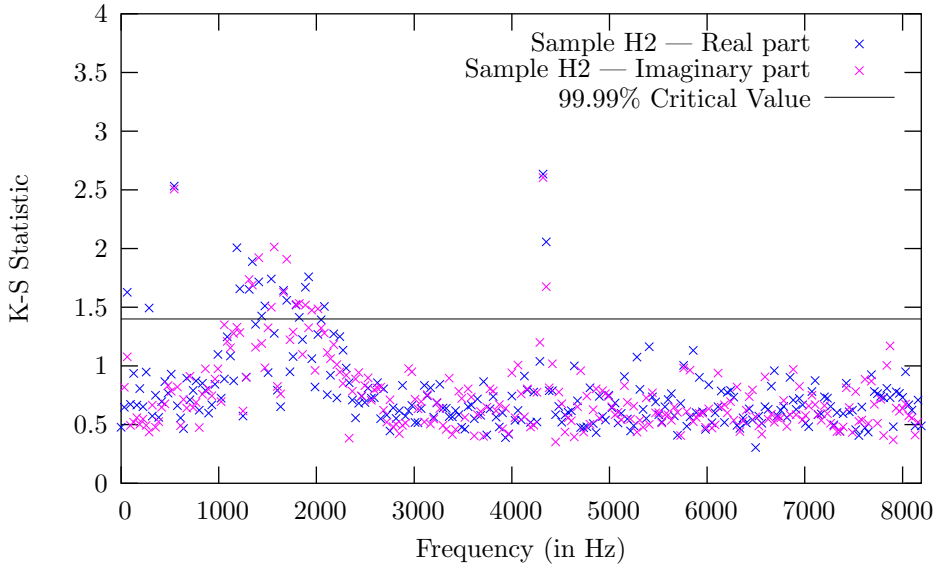

Figure 12: The goodness of fit of a Normal distribution to the time series of real and imaginary parts of the DFT coefficients from a spectrographic transform of Sample H2, using a width of 512, at various frequencies.

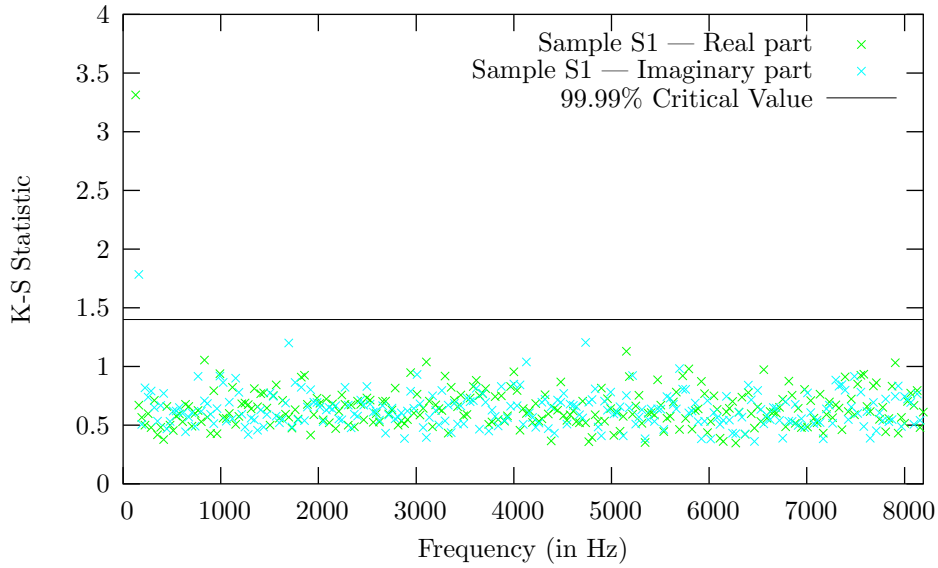

Figure 13: The goodness of fit of a Normal distribution to the time series of real and imaginary parts of the DFT coefficients from a spectrographic transform of Sample S1, using a width of 512, at various frequencies.

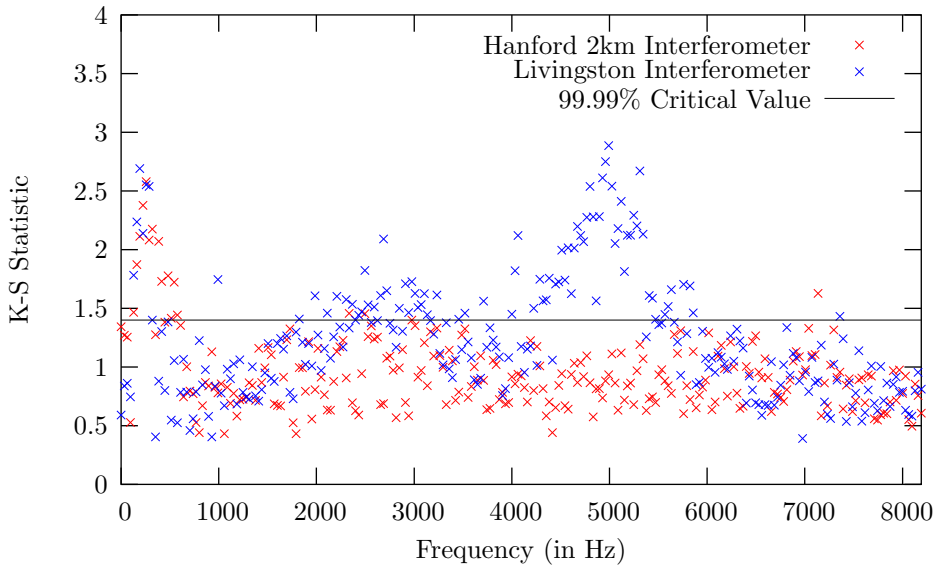

Figure 14: The goodness of fit of a Gamma distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

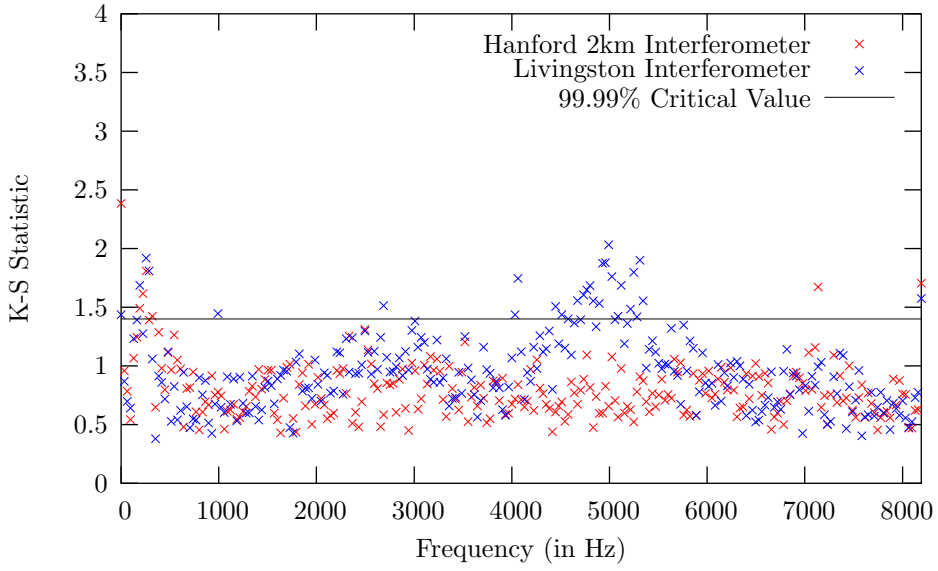

Figure 15: The goodness of fit of a Weibull distribution to the power time series of various frequencies, from a spectrographic transform of width 512.

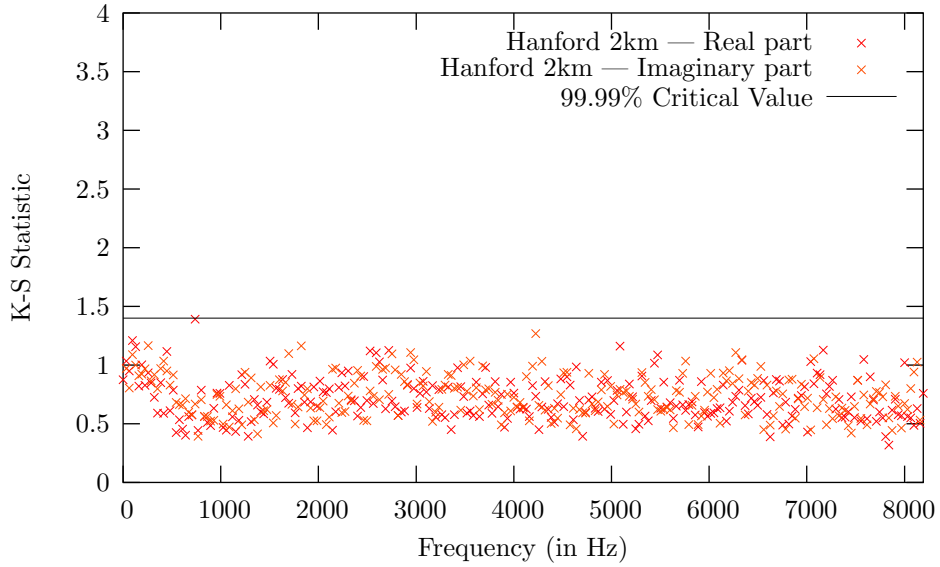

Figure 16: The goodness of fit of a Normal distribution to the time series of real and imaginary parts of the DFT coefficients from a spectrographic transform of a sample of Hanford 2km dark port data, using a width of 512, at various frequencies.

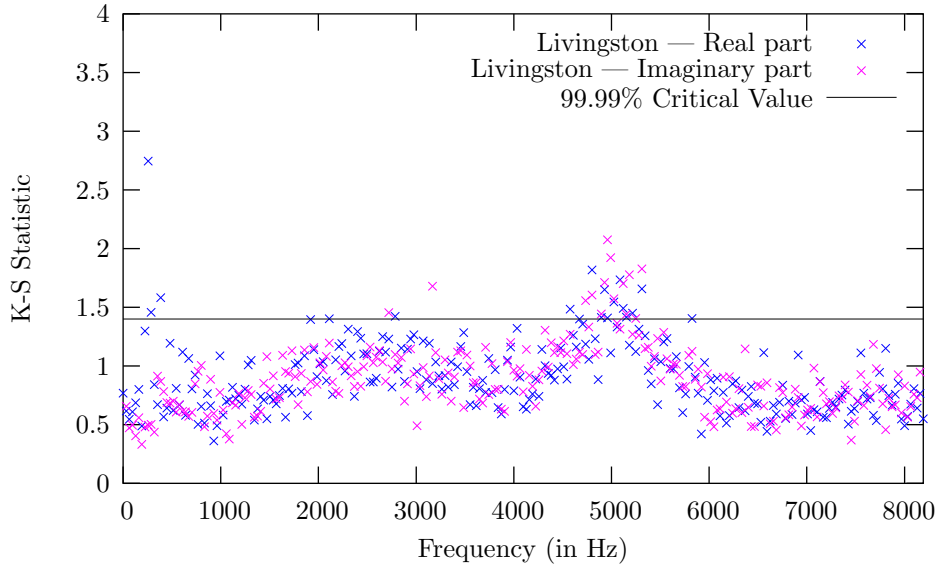

Figure 17: The goodness of fit of a Normal distribution to the time series of real and imaginary parts of the DFT coefficients from a spectrographic transform of a sample of Livingston dark port data, using a width of 512, at various frequencies.

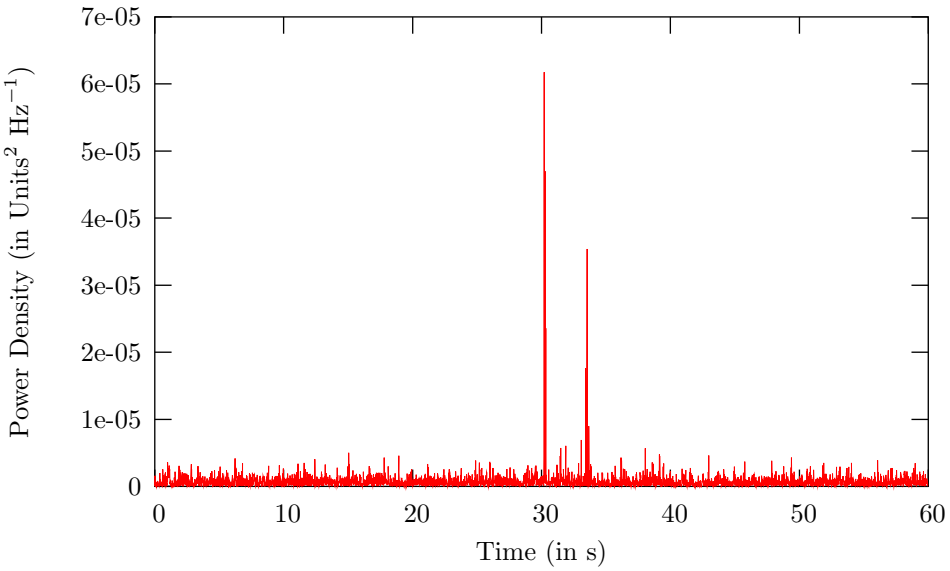

Figure 18: The time series of power in the 288Hz frequency band, for Sample H1. The unit used is uncalibrated strain, as saved in '.gwf' files.

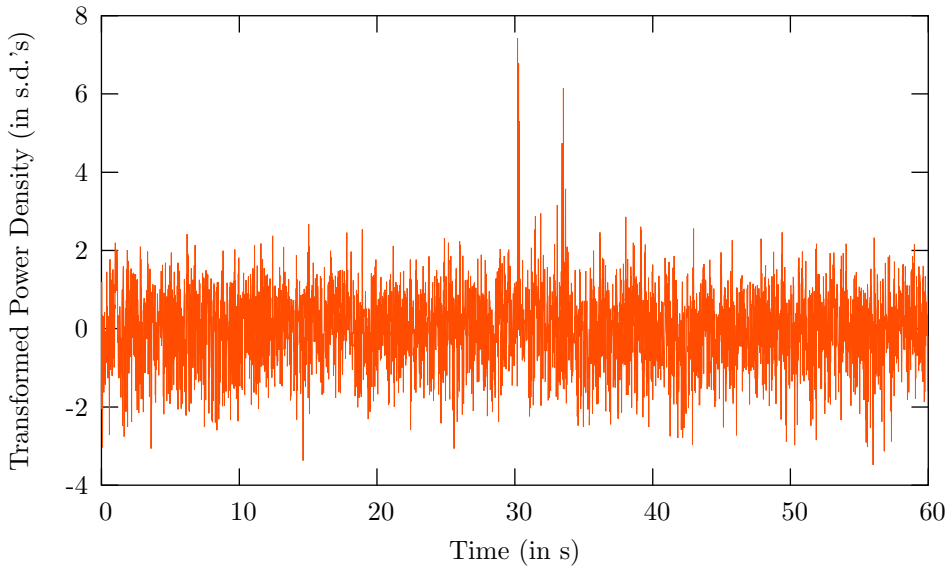

Figure 19: The optimal Box-Cox transform of Figure 18.

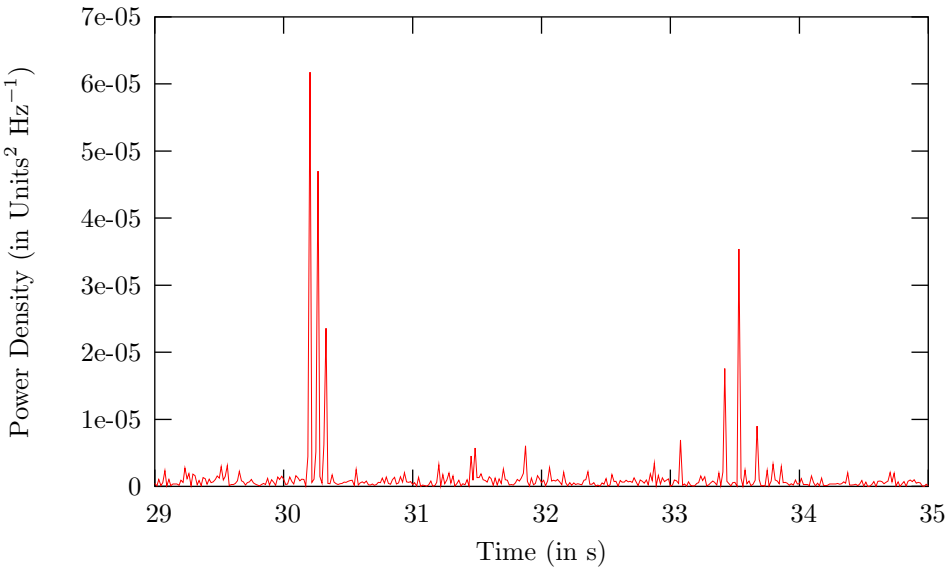

Figure 20: A closeup of the candidate transient shown in Figure 18.

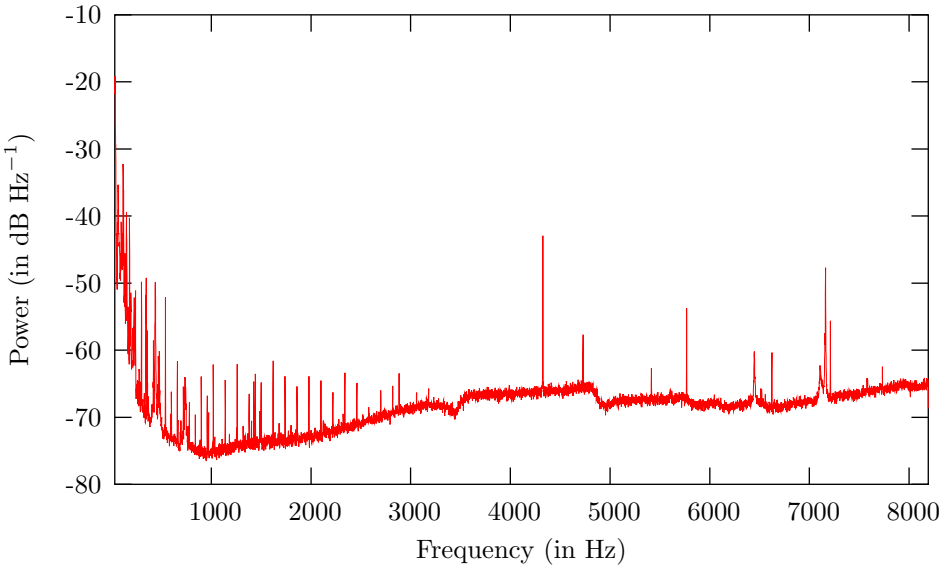

Figure 21: The power spectrum of Sample H1. This was created by Fourier transforming 60 non-overlapping one second intervals, and averaging the power present in each frequency over all intervals.

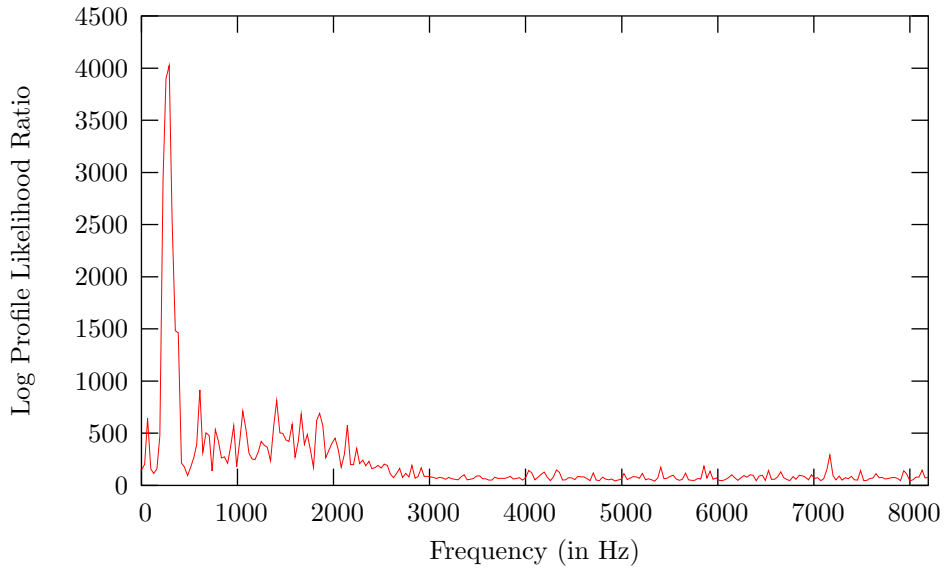

Figure 22: The likelihood of a transient being present in Sample H1 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Normal distribution.

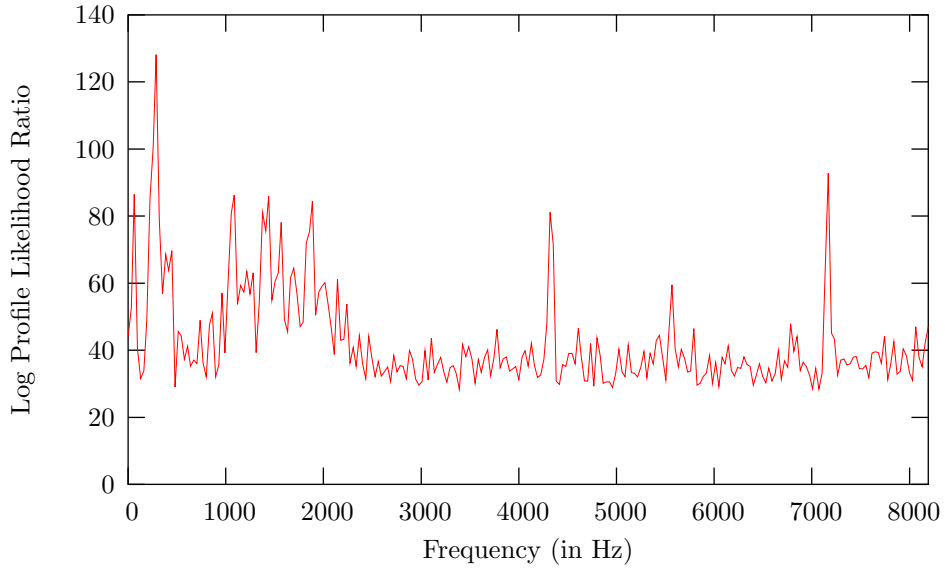

Figure 23: The likelihood of a transient being present in Sample H1 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Normal distribution. Each frequency band was Box-Cox transformed independently prior to analysis.

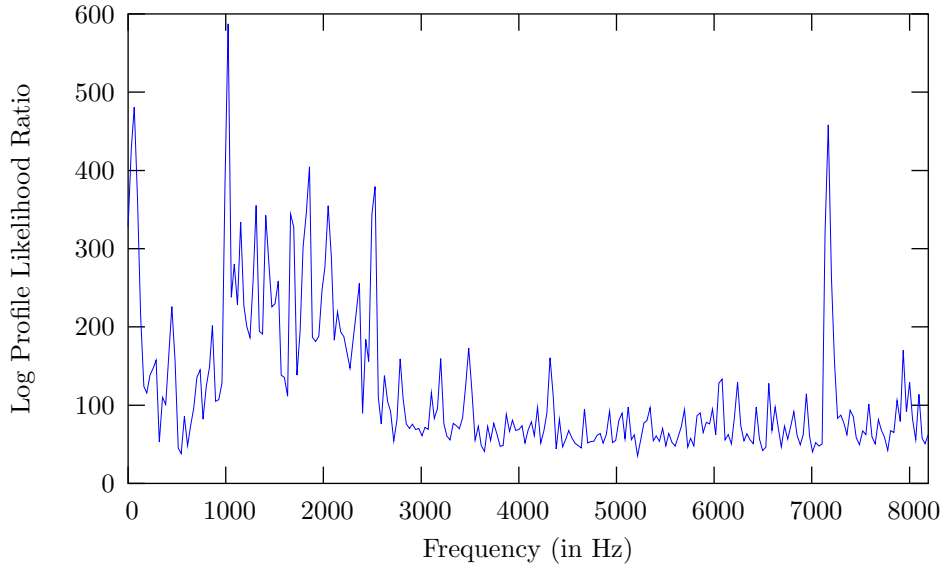

Figure 24: The likelihood of a transient being present in Sample H2 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Normal distribution.

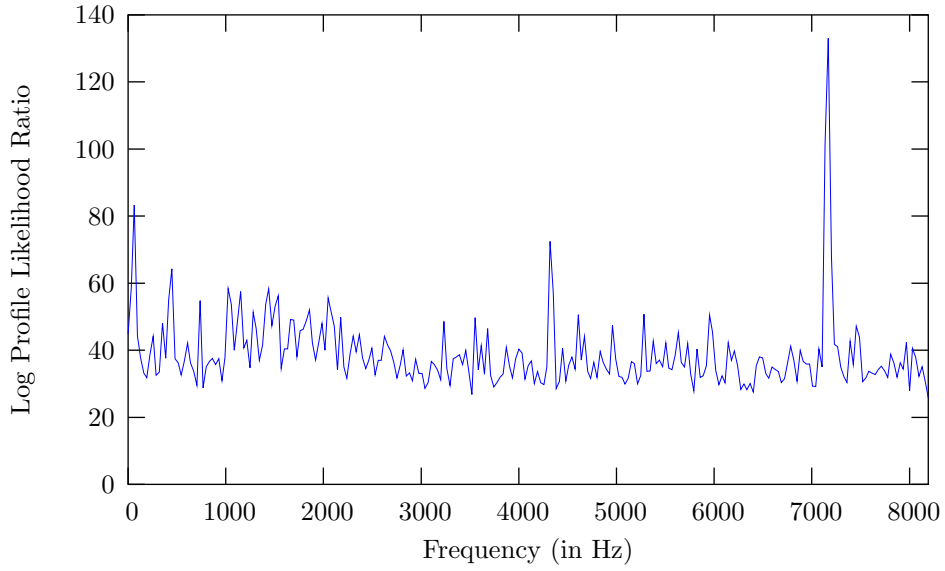

Figure 25: The likelihood of a transient being present in Sample H2 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Normal distribution. Each frequency band was Box-Cox transformed independently prior to analysis.

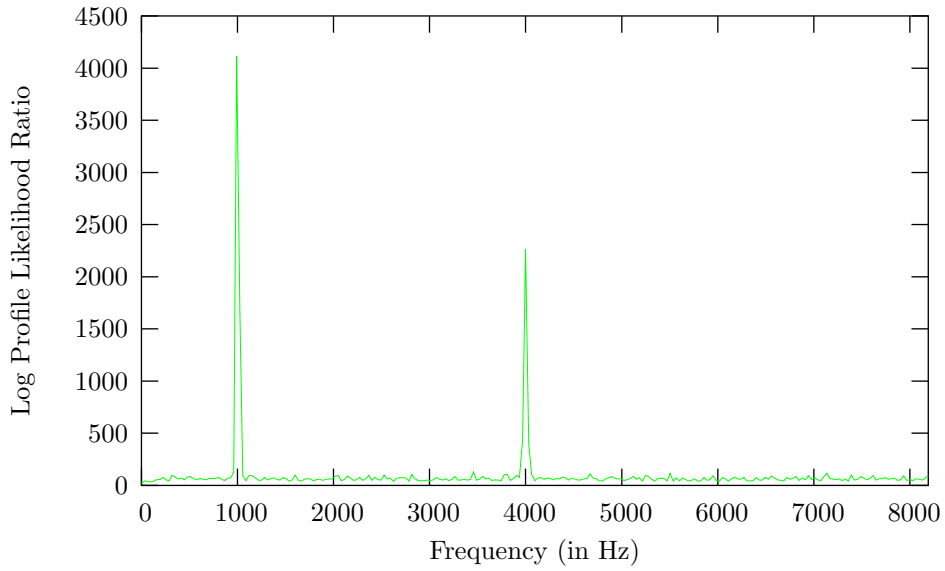

Figure 26: The likelihood of a transient being present in Sample S1 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Normal distribution.

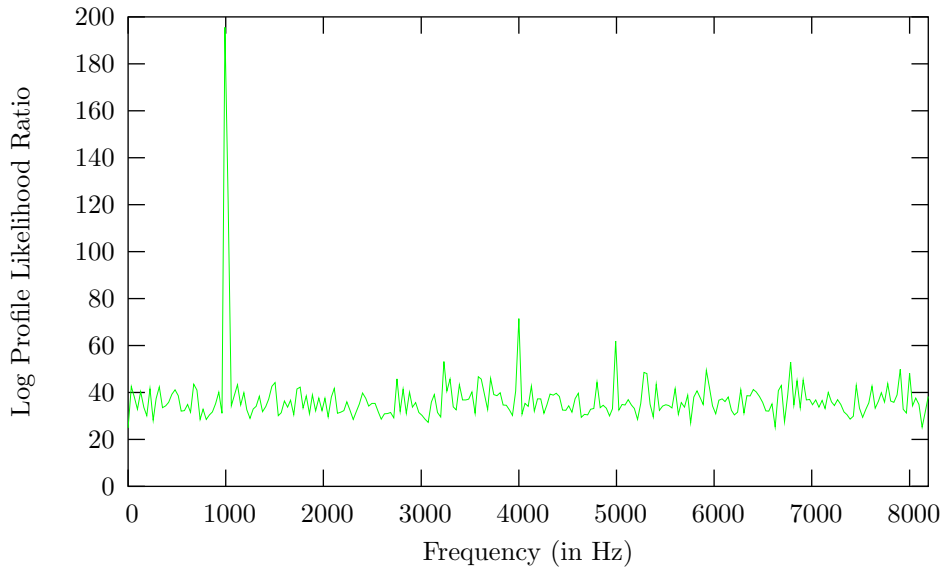

Figure 27: The likelihood of a transient being present in Sample S1 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Normal distribution. Each frequency band was Box-Cox transformed independently prior to analysis.

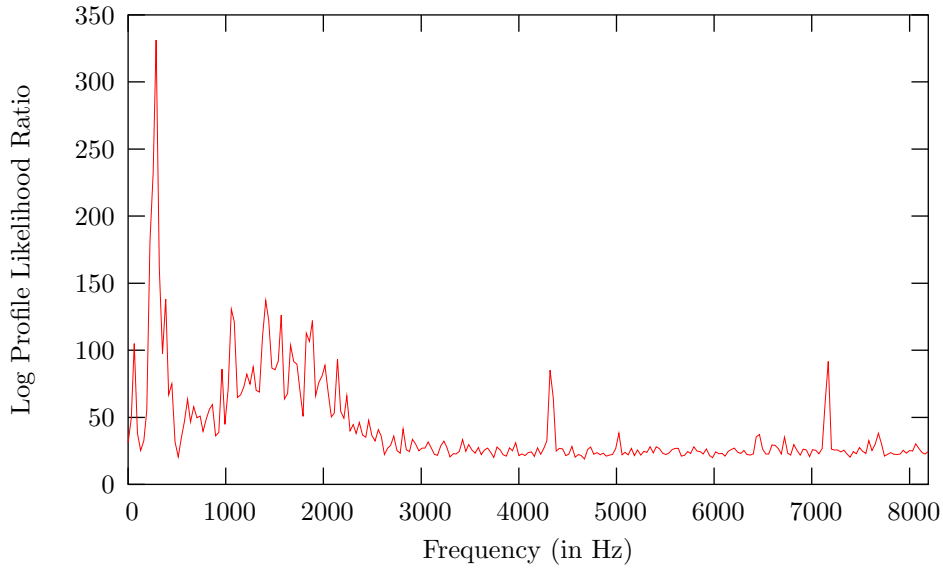

Figure 28: The likelihood of a transient being present in Sample H1 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Gamma distribution.

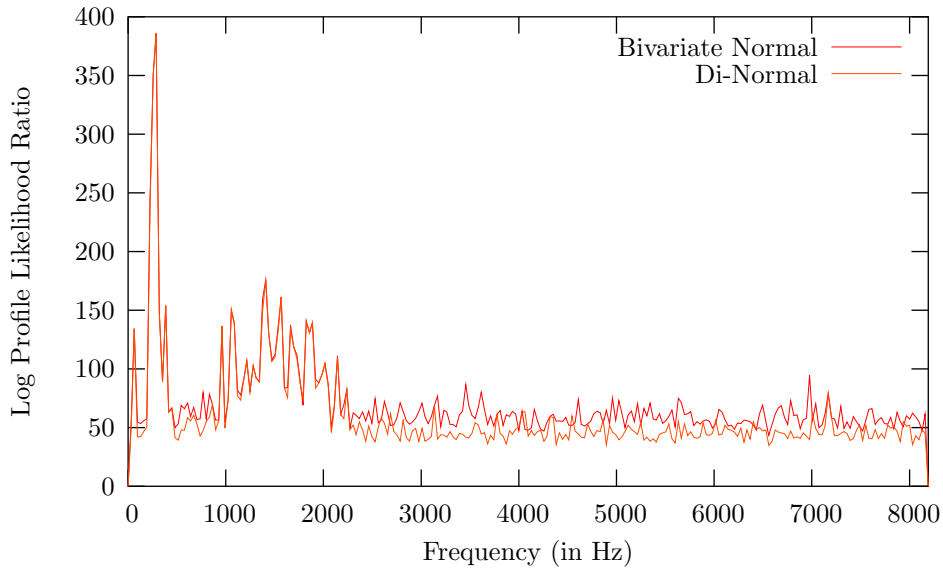

Figure 29: The likelihood of a transient being present in Sample H1 for various frequency bands, using the Likelihood Ratio Interval Analysis on the DFT coefficients.

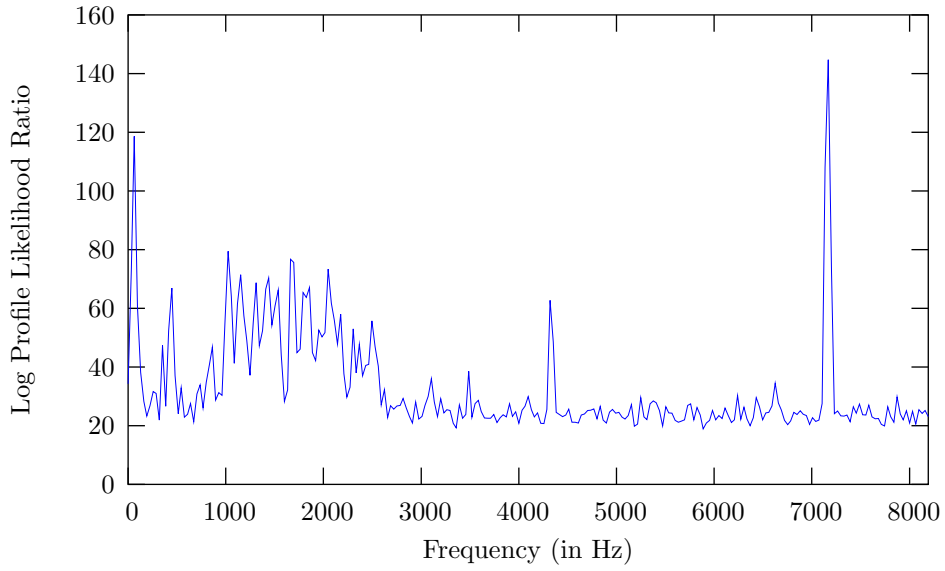

Figure 30: The likelihood of a transient being present in Sample H2 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Gamma distribution.

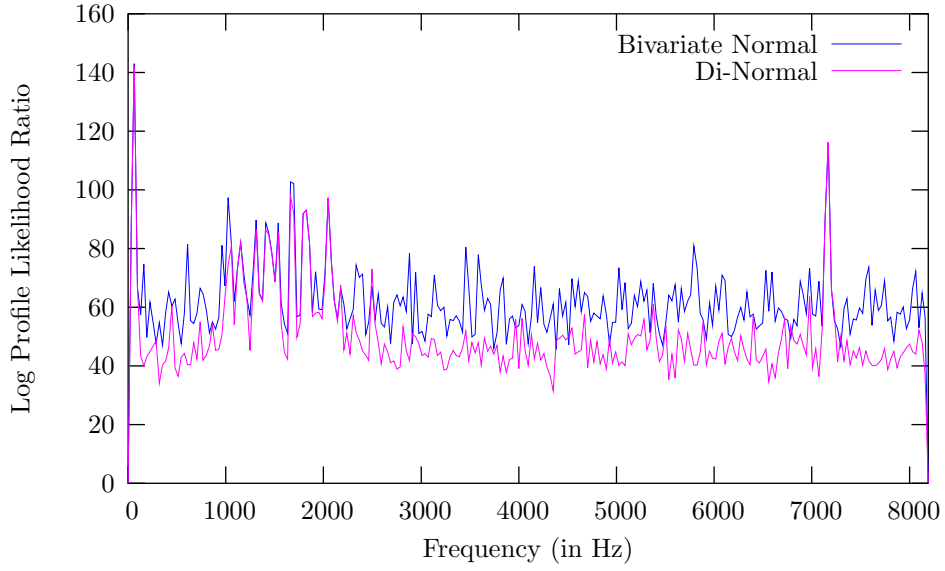

Figure 31: The likelihood of a transient being present in Sample H2 for various frequency bands, using the Likelihood Ratio Interval Analysis on the DFT coefficients.

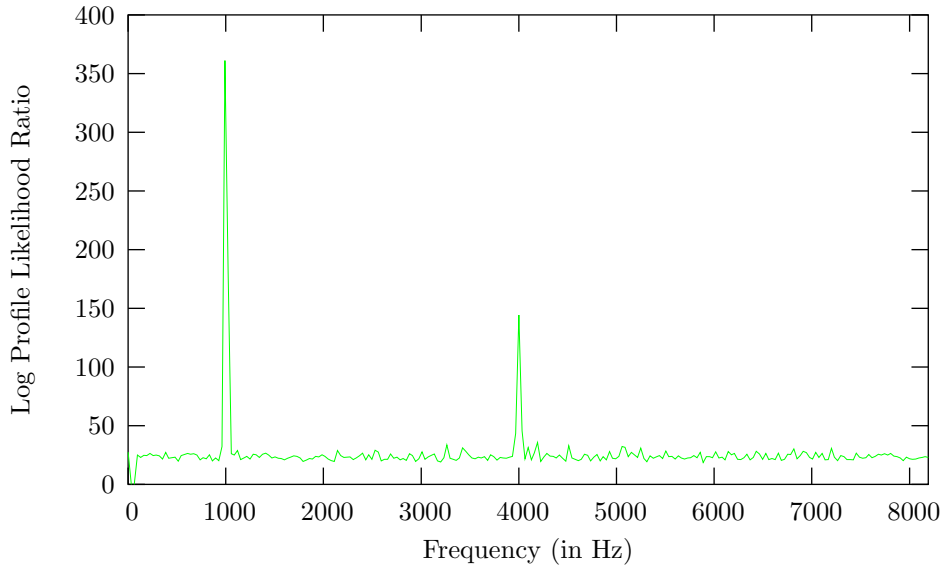

Figure 32: The likelihood of a transient being present in Sample S1 for various frequency bands, using the Likelihood Ratio Interval Analysis, and assuming a Gamma distribution.

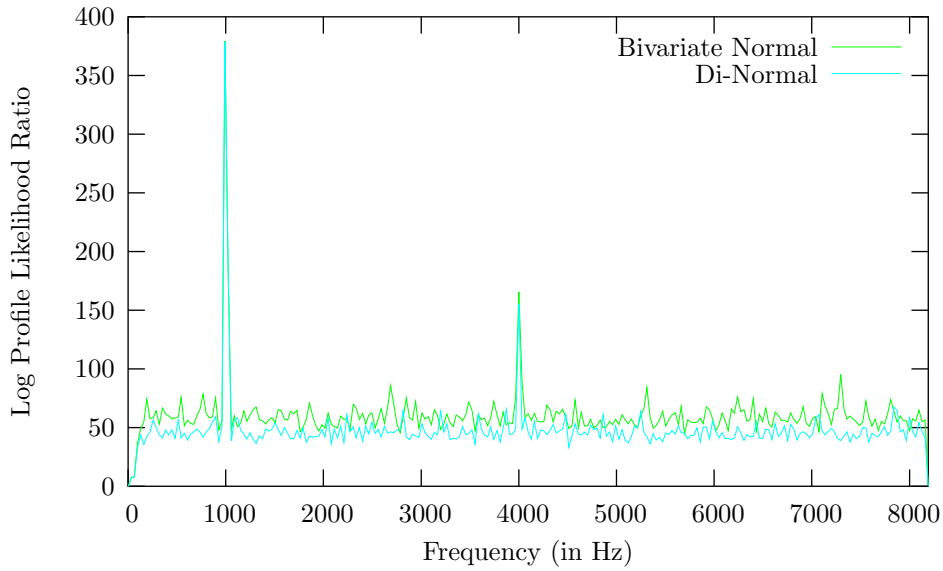

Figure 33: The likelihood of a transient being present in Sample S1 for various frequency bands, using the Likelihood Ratio Interval Analysis on the DFT coefficients.

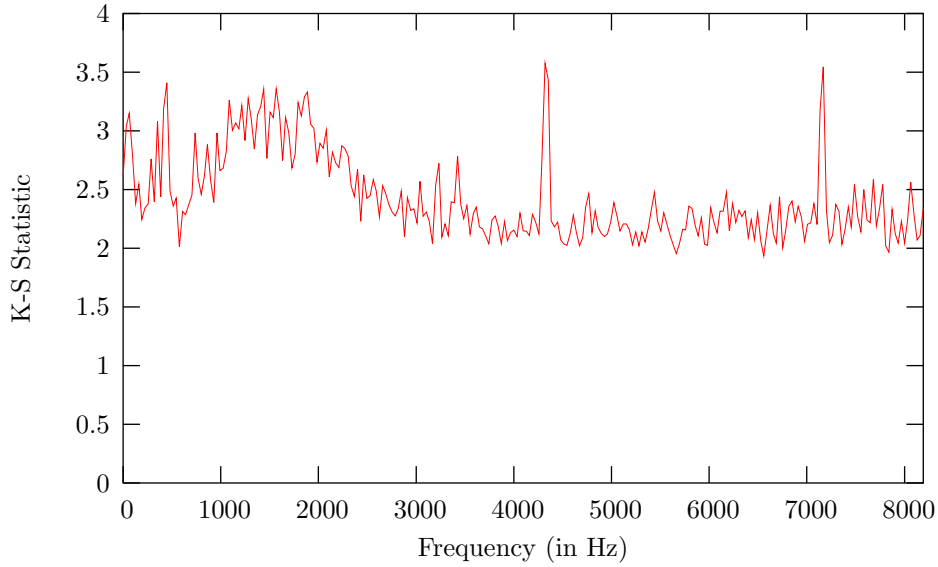

Figure 34: The likelihood of a transient being present in Sample H1 for various frequency bands, using the K-S Interval test.

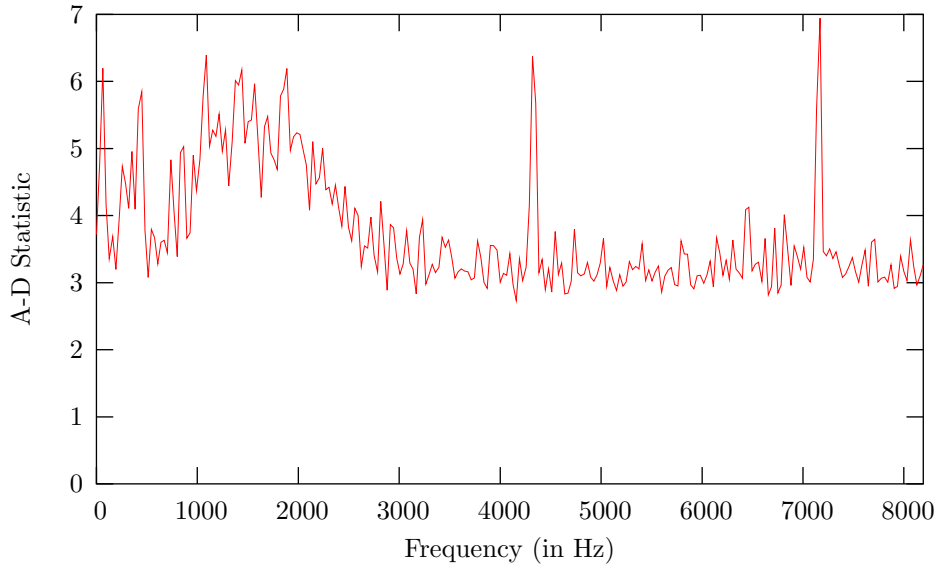

Figure 35: The likelihood of a transient being present in Sample H1 for various frequency bands, using the A-D Interval test.

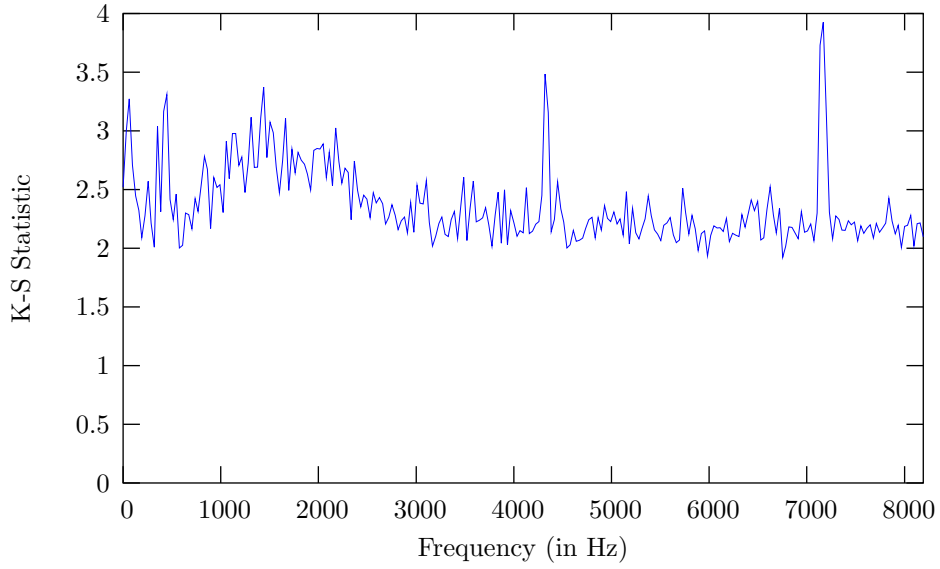

Figure 36: The likelihood of a transient being present in Sample H2 for various frequency bands, using the K-S Interval test.

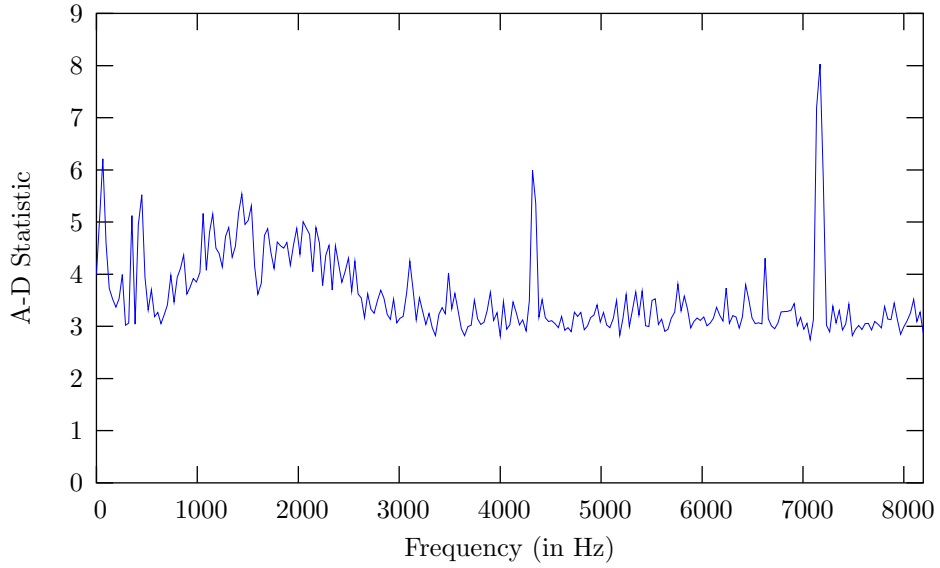

Figure 37: The likelihood of a transient being present in Sample H2 for various frequency bands, using the A-D Interval test.

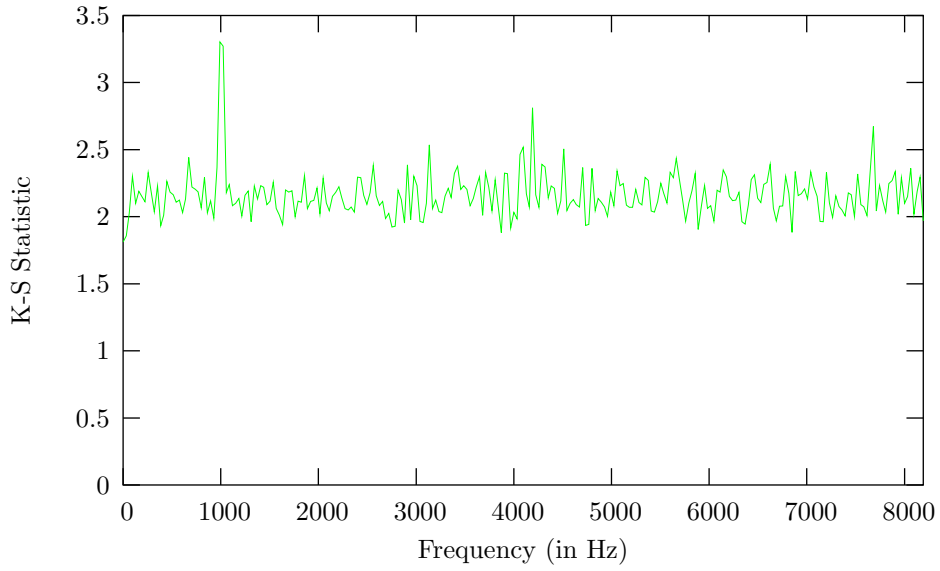

Figure 38: The likelihood of a transient being present in Sample S1 for various frequency bands, using the K-S Interval test.

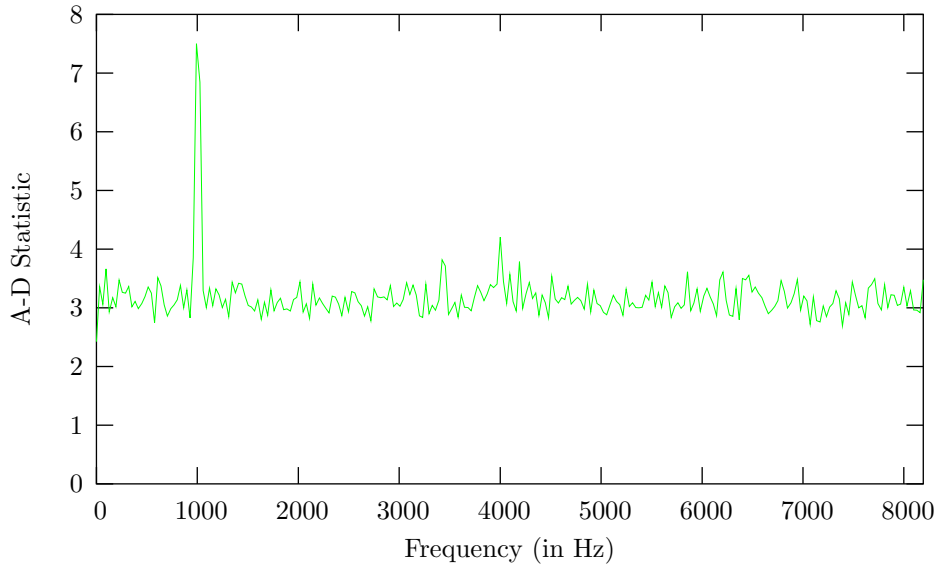

Figure 39: The likelihood of a transient being present in Sample S1 for various frequency bands, using the A-D Interval test.

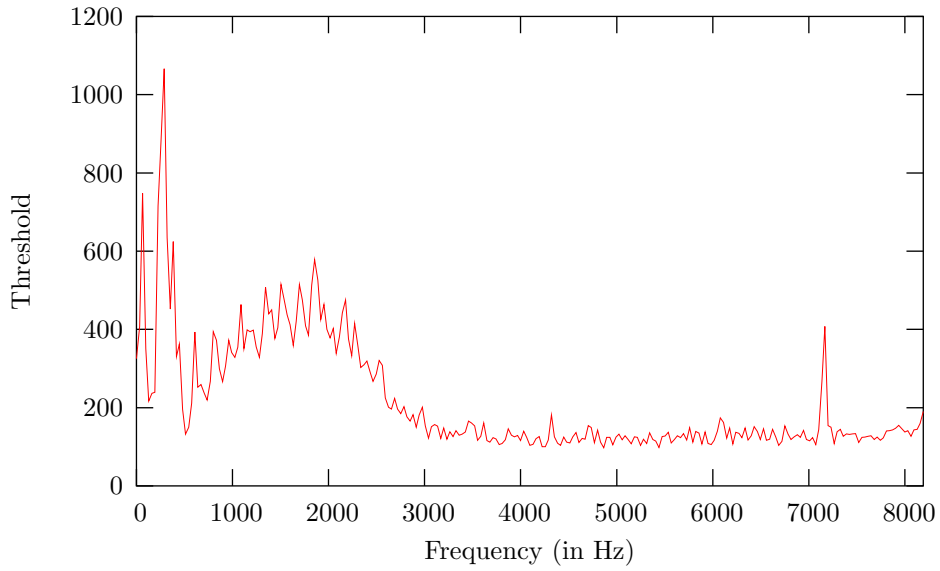

Figure 40: Thresholds generated by `BramThreshold` using Normal analysis, with s.d. of 5. An interval length of 10 seconds, minimum size of 3 samples, maximum size of 0.25 seconds, and spectrogram width of 512 samples were used. The data processed was all Hanford 4km `LSC-AS_Q` data from 729337008 to 729338000.

E BramMon Results

The results listed here are from analyzing `glitch` triggers generated for the Hanford 4km interferometer's dark port (`H1:LSC-AS_Q`) between GPS times 729337008 and 729338000 using `BramMon`. A spectrogram width of 512 samples, minimum interval length of 3 samples, and maximum length of 0.25 seconds were used.

E.1 Normal Analysis, 10s Samples, and Constant Threshold 500

```
Opened trigger table file: Glitch_H1.xml
Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337221 s + 961914063 ns GPS time
    dT: 0.2 s
    Significance: 5.026

1 Candidate Transients Found:
    Start Time: 4.953 s into sample (729337221 s + 915039063 ns GPS time)
    Duration: 0.2188 s
    Significance: 1.164
    Excess Power: 0.1795
    Frequency Decomposition:
        48 - 80: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337275 s + 222290039 ns GPS time
    dT: 0.2293 s
    Significance: 7.942

1 Candidate Transients Found:
    Start Time: 4.969 s into sample (729337275 s + 191040039 ns GPS time)
    Duration: 0.1562 s
    Significance: 2.936
    Excess Power: 0.0008964
    Frequency Decomposition:
        176 - 400: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337278 s + 689270020 ns GPS time
    dT: 0.2021 s
    Significance: 5.955
```

```
2 Candidate Transients Found:
    Start Time: 1.5 s into sample (729337275 s + 189270020 ns GPS time)
    Duration: 0.1562 s
    Significance: 2.537
    Excess Power: 0.0009453
    Frequency Decomposition:
        208 - 400: 1

    Start Time: 4.734 s into sample (729337278 s + 423645020 ns GPS time)
    Duration: 0.125 s
    Significance: 1.233
    Excess Power: 0.0001831
    Frequency Decomposition:
        272 - 304: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337467 s + 671936035 ns GPS time
    dT: 0.2 s
    Significance: 5.103

0 Candidate Transients Found:


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337513 s + 776306152 ns GPS time
    dT: 0.2 s
    Significance: 5.092

0 Candidate Transients Found:


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337739 s + 428771973 ns GPS time
    dT: 0.2075 s
    Significance: 6.836

0 Candidate Transients Found:


------------------------------------------------------------
```

```
Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337969 s + 475280762 ns GPS time
    dT: 0.2 s
    Significance: 5.424

0 Candidate Transients Found:

------------------------------------------------------------
```

E.2 Normal Analysis, 60s Samples, and Constant Threshold 500

```
Opened trigger table file: Glitch_H1.xml
Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337221 s + 961914063 ns GPS time
    dT: 0.2 s
    Significance: 5.026

3 Candidate Transients Found:
    Start Time: 29.95 s into sample (729337221 s + 915039063 ns GPS time)
    Duration: 0.2188 s
    Significance: 1.748
    Excess Power: 1.485
    Frequency Decomposition:
        48 - 80: 1

    Start Time: 54.36 s into sample (729337246 s + 321289063 ns GPS time)
    Duration: 0.25 s
    Significance: 1.278
    Excess Power: 1.159e-05
    Frequency Decomposition:
        1520 - 1552: 1

    Start Time: 58.58 s into sample (729337250 s + 540039063 ns GPS time)
    Duration: 0.2031 s
    Significance: 1.015
    Excess Power: 1.065e-05
    Frequency Decomposition:
        1040 - 1072: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337275 s + 222290039 ns GPS time
    dT: 0.2293 s
    Significance: 7.942
```

```
5 Candidate Transients Found:
    Start Time: 1.109 s into sample (729337246 s + 331665039 ns GPS time)
    Duration: 0.2656 s
    Significance: 1.499
    Excess Power: 1.849e-05
    Frequency Decomposition:
        1456 - 1488: 0.2746
        1520 - 1584: 0.458
        2128 - 2160: 0.2675

    Start Time: 5.188 s into sample (729337250 s + 409790039 ns GPS time)
    Duration: 0.5625 s
    Significance: 1.666
    Excess Power: 7.165e-05
    Frequency Decomposition:
        944 - 976: 0.07195
        1008 - 1104: 0.2082
        1360 - 1424: 0.2393
        1680 - 1712: 0.1067
        1808 - 1904: 0.3738

    Start Time: 29.97 s into sample (729337275 s + 191040039 ns GPS time)
    Duration: 0.1562 s
    Significance: 9.669
    Excess Power: 0.004621
    Frequency Decomposition:
        176 - 400: 0.9976
        688 - 720: 0.002441

    Start Time: 33.19 s into sample (729337278 s + 409790039 ns GPS time)
    Duration: 0.1406 s
    Significance: 2.69
    Excess Power: 0.0001793
    Frequency Decomposition:
        272 - 304: 1

    Start Time: 33.64 s into sample (729337278 s + 862915039 ns GPS time)
    Duration: 0.25 s
    Significance: 1.366
    Excess Power: 4.177
    Frequency Decomposition:
        48 - 80: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337278 s + 689270020 ns GPS time
    dT: 0.2021 s
    Significance: 5.955
```

```
5 Candidate Transients Found:
    Start Time: 1.578 s into sample (729337250 s + 267395020 ns GPS time)
    Duration: 0.7031 s
    Significance: 2.12
    Excess Power: 1.822e-06
    Frequency Decomposition:
        944 - 976: 0.0547
        1008 - 1104: 0.1721
        1360 - 1456: 0.3042
        1648 - 1712: 0.1822
        1808 - 1904: 0.2869

    Start Time: 26.5 s into sample (729337275 s + 189270020 ns GPS time)
    Duration: 0.1562 s
    Significance: 9.534
    Excess Power: 0.004458
    Frequency Decomposition:
        176 - 400: 0.9952
        656 - 720: 0.004757

    Start Time: 29.73 s into sample (729337278 s + 423645020 ns GPS time)
    Duration: 0.2656 s
    Significance: 2.664
    Excess Power: 0.0002915
    Frequency Decomposition:
        240 - 304: 0.8108
        368 - 400: 0.1681
        592 - 624: 0.0211

    Start Time: 30.17 s into sample (729337278 s + 861145020 ns GPS time)
    Duration: 0.25 s
    Significance: 1.337
    Excess Power: 4.167
    Frequency Decomposition:
        48 - 80: 1

    Start Time: 34.36 s into sample (729337283 s + 48645020 ns GPS time)
    Duration: 0.1562 s
    Significance: 1.002
    Excess Power: 4.59
    Frequency Decomposition:
        48 - 80: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337467 s + 671936035 ns GPS time
    dT: 0.2 s
    Significance: 5.103
```

```
1 Candidate Transients Found:
    Start Time: 35.23 s into sample (729337472 s + 906311035 ns GPS time)
    Duration: 0.25 s
    Significance: 1.527
    Excess Power: 1.131e-05
    Frequency Decomposition:
        1968 - 2000: 0.555
        2160 - 2192: 0.445


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337513 s + 776306152 ns GPS time
    dT: 0.2 s
    Significance: 5.092

0 Candidate Transients Found:

------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337739 s + 428771973 ns GPS time
    dT: 0.2075 s
    Significance: 6.836

3 Candidate Transients Found:
    Start Time: 37.72 s into sample (729337747 s + 147521973 ns GPS time)
    Duration: 0.2344 s
    Significance: 1.095
    Excess Power: 3.266e-06
    Frequency Decomposition:
        1008 - 1040: 1

    Start Time: 45.23 s into sample (729337754 s + 663146973 ns GPS time)
    Duration: 0.25 s
    Significance: 1.379
    Excess Power: 7.099
    Frequency Decomposition:
        48 - 80: 1

    Start Time: 46.3 s into sample (729337755 s + 725646973 ns GPS time)
    Duration: 0.2344 s
    Significance: 1.11
    Excess Power: 7.776
    Frequency Decomposition:
        48 - 80: 1


------------------------------------------------------------
```

```
Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337969 s + 475280762 ns GPS time
    dT: 0.2 s
    Significance: 5.424

2 Candidate Transients Found:
    Start Time: 6.438 s into sample (729337945 s + 912780762 ns GPS time)
    Duration: 0.25 s
    Significance: 1.277
    Excess Power: 1.389e-06
    Frequency Decomposition:
        1680 - 1712: 1

    Start Time: 57.31 s into sample (729337996 s + 787780762 ns GPS time)
    Duration: 0.3281 s
    Significance: 1.934
    Excess Power: 9.861e-05
    Frequency Decomposition:
        1328 - 1360: 0.1086
        1552 - 1584: 0.11
        1712 - 1904: 0.6785
        1968 - 2000: 0.1029


------------------------------------------------------------
```

E.3 Normal Analysis, 10s Samples, and `BramThreshold` Threshold from Figure 40

```
Opened trigger table file: Glitch_H1.xml
Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337221 s + 961914063 ns GPS time
    dT: 0.2 s
    Significance: 5.026

1 Candidate Transients Found:
    Start Time: 7.25 s into sample (729337224 s + 211914063 ns GPS time)
    Duration: 0.09375 s
    Significance: 1.391
    Excess Power: 6.124e-07
    Frequency Decomposition:
        5968 - 6000: 1


------------------------------------------------------------
```

```
Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337275 s + 222290039 ns GPS time
    dT: 0.2293 s
    Significance: 7.942

3 Candidate Transients Found:
    Start Time: 4.969 s into sample (729337275 s + 191040039 ns GPS time)
    Duration: 0.1562 s
    Significance: 2.293
    Excess Power: 0.0009003
    Frequency Decomposition:
        176 - 400: 0.9956
        656 - 720: 0.004389

    Start Time: 5.453 s into sample (729337275 s + 675415039 ns GPS time)
    Duration: 0.04688 s
    Significance: 1.081
    Excess Power: 3.672e-05
    Frequency Decomposition:
        4176 - 4208: 1

    Start Time: 8.438 s into sample (729337278 s + 659790039 ns GPS time)
    Duration: 0.04688 s
    Significance: 1.136
    Excess Power: 2.452e-05
    Frequency Decomposition:
        560 - 592: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337278 s + 689270020 ns GPS time
    dT: 0.2021 s
    Significance: 5.955

3 Candidate Transients Found:
    Start Time: 1.5 s into sample (729337275 s + 189270020 ns GPS time)
    Duration: 0.1562 s
    Significance: 1.809
    Excess Power: 0.001181
    Frequency Decomposition:
        176 - 272: 0.8514
        304 - 368: 0.1432
        656 - 720: 0.005438
```

```
    Start Time: 1.984 s into sample (729337275 s + 673645020 ns GPS time)
    Duration: 0.04688 s
    Significance: 1.277
    Excess Power: 3.703e-05
    Frequency Decomposition:
        4176 - 4208: 1

    Start Time: 4.953 s into sample (729337278 s + 642395020 ns GPS time)
    Duration: 0.04688 s
    Significance: 1.024
    Excess Power: 2.086e-05
    Frequency Decomposition:
        752 - 784: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337467 s + 671936035 ns GPS time
    dT: 0.2 s
    Significance: 5.103

1 Candidate Transients Found:
    Start Time: 3.844 s into sample (729337466 s + 515686035 ns GPS time)
    Duration: 0.04688 s
    Significance: 1.009
    Excess Power: 1.236e-05
    Frequency Decomposition:
        7696 - 7728: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337513 s + 776306152 ns GPS time
    dT: 0.2 s
    Significance: 5.092

0 Candidate Transients Found:

------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337739 s + 428771973 ns GPS time
    dT: 0.2075 s
    Significance: 6.836
```

```
1 Candidate Transients Found:
    Start Time: 6.125 s into sample (729337740 s + 553771973 ns GPS time)
    Duration: 0.0625 s
    Significance: 1.326
    Excess Power: 7.46e-06
    Frequency Decomposition:
        3472 - 3504: 1


------------------------------------------------------------

Trigger Info:
    ID: Glitch
    subID: H1:LSC-AS_Q
    Time: 729337969 s + 475280762 ns GPS time
    dT: 0.2 s
    Significance: 5.424

0 Candidate Transients Found:

------------------------------------------------------------
```

F References

References

[1] A. L. Stuver, *A New LIGO Data Analysis Method...*
http://gravity.psu.edu/~stuver/research.htm

[2] J. Sylvestre, *Searching for Bursts: Algorithms, Vetoes, Upper Limits*
http://www.ligo.caltech.edu/docs/G/G020028-00.pdf

[3] *What do we do when data are "Non-normal"?*
NIST/SEMATECH e-Handbook of Statistical Methods
http://www.itl.nist.gov/div898/handbook/pmc/section5/pmc52.htm

[4] *Kolmogorov-Smirnov Goodness-of-Fit Test*
NIST/SEMATECH e-Handbook of Statistical Methods
http://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm

[5] *Anderson-Darling Test*
NIST/SEMATECH e-Handbook of Statistical Methods
http://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm

[6] *Tests for Comparison of Two Independent Samples*
A Knowledgebase for Extragalactic Astronomy and Cosmology
http://nedwww.ipac.caltech.edu/level5/Wall2/Wal4_3.html