Modeling Unexplained Noise in GEO600 Data.

Kelsey Dewey

Department of Physics, University of Florida

P.O. Box 118440, Gainesville, FL 32611-8440

Dr. Ik Siong Heng

Institute for Gravitational Research, University of Glasgow

Glasgow G12 8QQ, United Kingdom

Dr. Stefan Hild

Institute for Gravitational Research, University of Glasgow

Glasgow G12 8QQ, United Kingdom

Dr. Matthew Pitkin

Institute for Gravitational Research, University of Glasgow

Glasgow G12 8QQ, United Kingdom

Abstract

It has recently come to light that the current noise model in use at GEO600 may be lacking a necessary holographic noise component, a theory which was proposed by Craig Hogan. Even in the event that the necessity of holographic noise is a certainty, there is still a significant quantity of noise which is unexplained by the current noise model. This report seeks to explore several simulated noise sources which complete this noise model. This objective was attempted by using Markov chain Monte Carlo methods of determining the posterior probabilities of interest in each of the various models studied.

1 Introduction

The gravitational wave detector GEO600 is a Michelson type interferometer located near Hannover, Germany. With an armlength of 600 meters, GEO600 is the smallest of the detectors in the LIGO Scientific Collaboration, and as such it has implemented many advanced means of innovation which allow it to take data of equal value to that of the larger detectors. Among these innovations are folded arms which give an effective armlength of 1200 meters, power recycling, signal recycling, and with upgrades that began in 2010, an output mode cleaner and the use of squeezing to limit shot noise. These advanced techniques allowed GEO600 to partake in five data taking runs with the larger detectors beginning in 2005, and these techniques will also be applied to the larger detectors in upgrades that will result in Advanced LIGO and Advanced Virgo. They will also allow GEO600 to remain online and searching for gravitational waves while the other detectors have gone offline to receive those upgrades, [3].

In recent years Craig Hogan, director of Fermilab and professor at the University of Chicago, has put forth a theory which proposes that our universe could be a hologram, and this holographic nature may be detectable in the GEO600 data. In short, this 'holographic principle' suggest that there is a possibility that world we live in is a projection from the two dimensional surface of the very edge of the universe [4]. Just as blowing up an image will produce a grainy picture, Hogan believes that this projection from two dimensions to three will lead to space-time possessing a grainy quality. Furthermore, Hogan believes that if the universe is a three dimensional projection from a two dimensional surface, the resulting graininess of space-time will lead to observable effects in the data which GEO600 is collecting. He believes that some of the noise which GEO600's researchers currently cannot explain is, in fact, attributable to the pixelated quality of space-time.

This report aims to describe a method which was developed to create a means of determining the parameters of a model which yield the result closest to the data observed by GEO600. This method made use of Bayesian analysis (for which D.S. Silvia's book [1] is a good general source of reference), as

well as a Markov chain Monte Carlo means of creating probability distributions of the parameters which the given model was varied over. This method was used to determine the best parameters of a variety of models whose subsequent fits to the observed data will be examined and compared.

1.1 Noise Projections and Sources
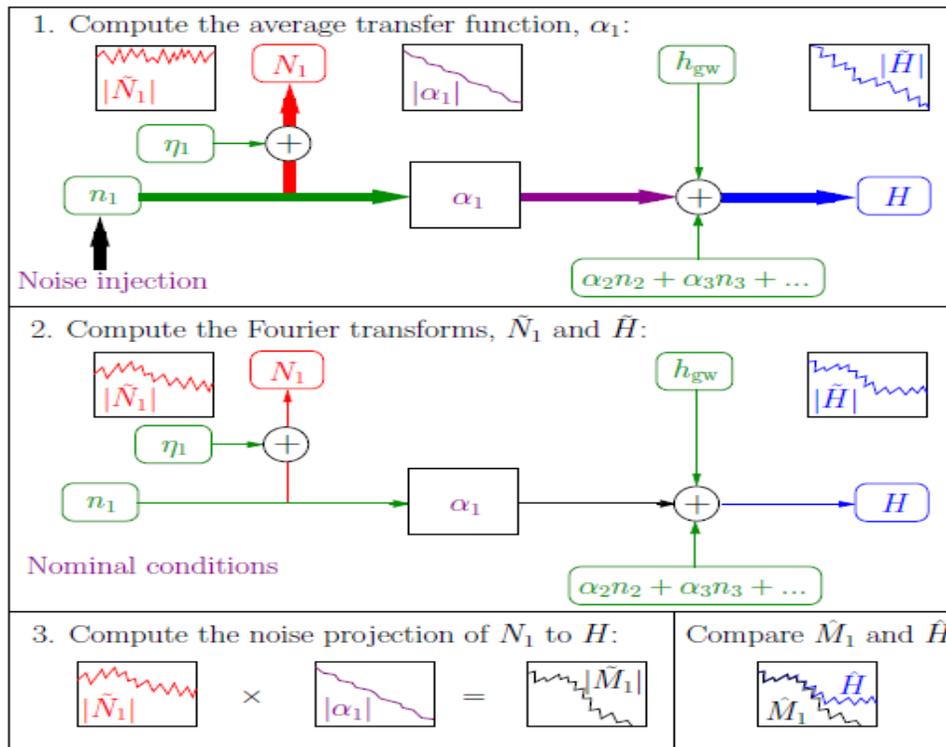
Figure1: This diagram gives an overview of how a noise projection is done. In the bottom right hand corner shows the final comparison between the noise projection, M_1, and the observed strain, H. (Image taken from [2])

　　　　Given that much of what this report is involved with is working with noise projections, a brief summary of how they are calculated will be given, as detailed by Joshua Smith in his 2006 paper, [2]. When a noise source is identified as something which is contributing to the overall strain being detected, a means of identifying how much it is contributing must be identified. Thus, if the overall output channel is measured, its signal may be broken down into gravitational wave signals plus the sum of all mechanical and environmental noise present. Identifying each of those noise signals and determining their effect on the overall output is the target of noise projection. Initially, a transfer function must be calculated, the transfer function is a factor which the noise signal is multiplied by such that the result is compatible with the final output signal, and with this the noise source at hand's contribution to the final output may be determined.

　　　　When calculating a transfer function, there must be injected noise over the relevant frequency ranges, such that the noise source for which the transfer function is being calculated dominates the output signal. This step must be taken to ensure that no transient noise influences the calculation of this transfer function. At this point the discrete Fourier transform is taken of both the noise signal and output signal; this step takes the signals from the time domain into the frequency domain. After these steps have been taken, the noise projection of the noise source onto that of the output signal may be

3

calculated. It is important to note that when these noise projections are added to create the sum of all the noise sources, they are added in quadrature. Thus, the sum of all noise sources whose transfer functions have been calculated and whose effect on the output signal are explained takes the following form (where there are i noise sources, each denoted by a subscripted N):

$$sum\ of\ explained\ noise = \sqrt{(N_1)^2 + (N_2)^2 + \cdots + (N_i)^2}$$

1.2 Holographic Noise

Craig Hogan's holographic noise theory has its origins in black hole thermodynamics. The combined efforts of Stephen Hawking and Jacob Bekenstein led to the determination that the entropy of a black hole is equal to one quarter of the area of its event horizon in Planck units [Phys. Rev. D 78, 087501 (2008)]. This result indicates that the number of microstates of a black hole is proportional to the area of its horizon, not the volume which it contains. Along with others, Craig Hogan has proposed that this may be generalized to apply the entire universe, the overall notion of this 'Holographic principal' being that all the information in the universe is encoded on the two dimensional surface of its boundary, and the world which we observe and live in is a projection from this 2D edge, effectively making our universe a hologram [5].

It has been suggested that the Planck length is the length at which the difference between two locations can no longer be distinguished from one another, that is to say that at this length scale space-time becomes controlled by quantum effects and diverges from the classical notion of continuous space-time. Given that the Planck length is equal to $\sqrt{\dfrac{hG}{2\pi(c^3)}}$, or approximately 1.616×10^{-35}, there is no way to build any apparatus which can make measurements at this scale in order to test this notion. However, if the universe were a hologram, and the two dimensional boundary is the location where the Plank length is the length at which quantum effects on space-time become apparent, then this length, when projected onto the emergent three dimensional volume, may be scaled to a size sufficiently large so as to be detected by the interferometers dedicated to gravitational wave detection. This is the concept which Craig Hogan has proposed [5, 6].

Hogan's theory holds that the quantum indeterminacy of space time will reveal itself in quantum fluctuations of space-time, and that these fluctuations will be large enough to be detected by the gravitational wave detectors. Particularly, Hogan believes that these fluctuations will present themselves in the data collected by GEO600 because it does not have Fabry perot cavities, and is thus designed to measure transverse position changes, which is what Hogan believes the fluctuations will affect. In his most recent paper Hogan has calculated that the 'holographic noise' which these quantum fluctuations lead to is equal to $1.85 \times 10^{-22} / \sqrt{\dfrac{\tau}{sec}}$, [5].

1.3 Bayesian Analysis and Markov Chain Monte Carlo

Bayesian analysis is a means of statistical analysis that is particularly suited for instances, such as those explored in this report, when it is desired to choose a set of parameters such that a reconstructed model best fits the actual observed data. At the most basic level we have the regular product rule found in probability theory, which is:

$$prob(X, Y|I) = prob(X|Y, I) * prob(Y|I)$$

Where X and Y are two propositions, the vertical bar may be read as "given", and I is the relevant background conditions that X and Y depend on. From this well-known equation, Bayes' theorem (immediately below) follows:

$$prob(X|Y, I) = \frac{prob(Y|X, I) * prob(X|I)}{prob(Y|I)}$$

We may see how Bayes' theorem is derived by first multiplying each side of the theorem by the denominator of the right hand side:

$$prob(X|Y, I) * prob(Y|I) = prob(X, Y|I)$$

We then realize that $prob(X, Y|I) = prob(Y, X|I)$, or that the probablilty of X and Y given I, is the same as the probability of Y and X given I. We may now write:

$$prob(Y, X|I) = prob(Y|X, I) * prob(X|I)$$

From here arriving at Bayes theorem is simply a matter of rearranging the above equations. The value of this method can be seen if we presume that X denotes the hypothesis we wish to test and Y denotes our observed data. We may have no way of knowing $prob(X|Y, I)$, however, we may use Bayes' theorem to determine this probability by breaking it down into bits which we may be able to determine. Overall we have that:

$$prob(hypothesis|data, I) \propto prob(data|hypothesis, I) * prob(hypothesis|I)$$

Where the "proportional to" symbol is necessary since the $prob(data|I)$ term has been left out, which is acceptable since this term is simply a normalizing term. When discussing Bayesian analysis the proper names of each of the three above terms are often used, so they are listed here:

$$prob(hypothesis|data, I) = posterior\ probability$$

$$prob(data|hypothesis, I) = likelihood\ function$$

$$prob(hypothesis|I) = prior\ probability$$

The prior probability can be seen as a statement of the knowledge, or lack thereof, of how likely a hypothesis is without analyzing any observed data. In an instance when little to nothing is known about this prior distribution, a flat prior probability, or simply a constant number which spans the range of the likelihood equation may be used. In other instances, when more is known about the prior probability, the prior may be chosen to reflect that knowledge. The likelihood function acts as the means of altering the prior to reflect the experimental values that are observed. Taken together these two return the posterior probability, which is the understanding of the likelihood of the hypothesis given the data that has been collected.

Another application of Bayesian Analysis is in creating an odds ratio. If there are two potential hypotheses, H_0 and H_H, this case the current noise model for GEO600, and one containing a holographic noise component, an odds ratio would be as follows (where {d} is the observed data set):

$$\frac{prob(H_H|\{d\},I)}{prob(H_0|\{d\},I)} = \frac{prob(H_H|I)}{prob(H_0|I)} \times \frac{prob(\{d\}|H_H)}{prob(\{d\}|H_o)}$$

Where the terms again have proper names, from left the right the first term is called the odds ratio, the second is called the prior odds, and the third is the Bayes' Factor. In this case, if the Bayes factor were to be on the order of ten, there would be good support that the model requires a holographic component. This concept is further explored in Graham Woan's 2009 paper [7].

In order to find certain posterior probability estimation, the likelihood function times the prior probability must be integrated over all the possible data points. This is where Markov chain Monte Carlo, shortened to MCMC henceforth, is often made use of. A MCMC method of creating a posterior probability will use a Markov chain to sample from the established prior probability. An initial point from the prior is selected and its likelihood value is calculated. Using the value of this initial point, a step is taken and a new point is chosen, then the likelihood value at this new point is calculated and is then added to the integral being calculated. Another step is now taken from this new point and continuing in this way a Markov chain is constructed. However, the Markov chain created in this process is constructed such that its stationary distribution is the integrand desired, resulting in the posterior distribution of interest [8]. The exact means of the MCMC used in this report is detailed in the next section.

2 Methods and Results

The first task which had to be completed was smoothing the GEO600 data so that it was no longer unnecessarily spikey. As previously discussed, calibration lines are inserted into the data such that noise projections can be successfully created. These calibration lines may be seen in Figure 2 in the observed strain of GEO600 (yellow curve), and they are particularly prevalent from approximately 600 to 700Hz. These lines would have interfered with calculating a proper fit to the observed strain in the GEO600 data since they varied immensely from the actual data, thus they had to be removed in order to create a reasonable fit to the strain in any of the various models tested.

These lines had to be removed efficiently, and thoroughly, however the smoothing process used had to avoid removing any significant portions of data. To do this, a rolling median of the data was calculated. A rolling median calculates the median of the data over a specified range of frequency bins. For instance, if this range were 11, then the kth point in the rolling median would be the median of the data from values k-5 to k+5. To determine an appropriate range to calculate this median over, several values were tested. It was found that smaller values left the data too spikey, and ultimately a value of 100 frequency bins was used. A median was used so that any significant increases in the value of the data were still reflected after smoothing. After this rolling median was calculated, a threshold at which to discard data points exceeding the threshold had to be determined. Several thresholds were considered, values between 1.05 times the rolling median and 5 times the rolling median were tested as thresholds, and 1.2 times the rolling threshold was chosen in the end, since it left significant changes of the data intact, but removed much of the spikes which would have interfered. This means that the final data which was used in all the calculations had discarded any datum which was larger than 1.2 times the rolling median calculated over the data values at the neighboring 100 frequency bins. The final process in this smoothing was to remove the data points between 607 to 715.5 Hz because it was not sufficiently smoothed by the previous process. The following figure shows the results of this process.
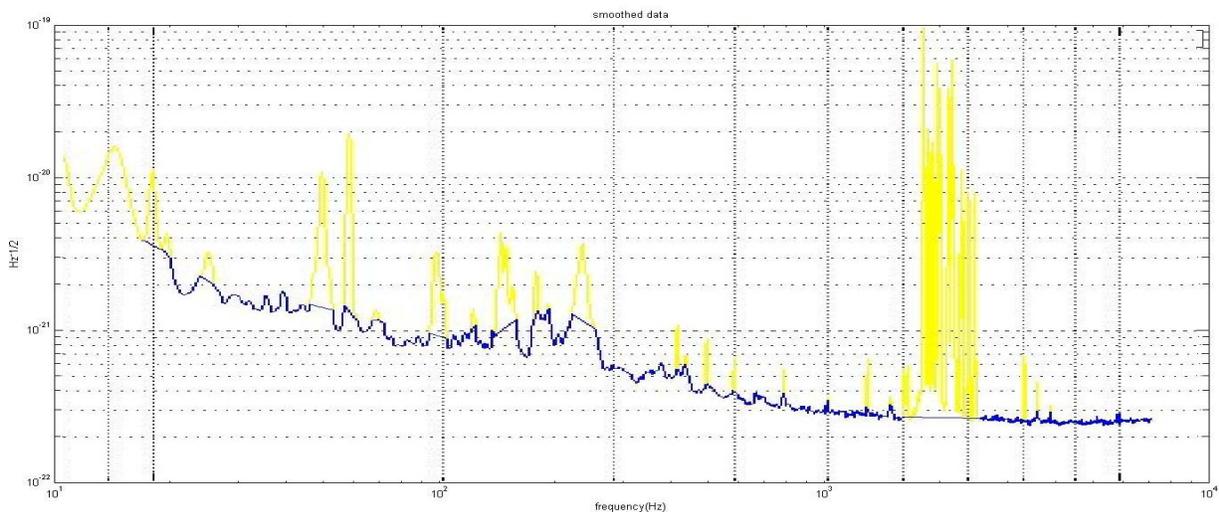

Figure2: This figure shows the results of the smoothing process as detailed above. The yellow curve represents the observed strain observed by GEO600, and the blue curve is the resulting smoothed curve that was henceforth used as the observed strain in GEO600. It is worth noting that no data values were altered by this process, data points were merely discarded in the instance that they were excessively large.

After the data was sufficiently smoothed, the next task to be competed was generating a means of creating a simple initial model to complete the existing GEO600 noise model. The following figure shows the existing noise model, as well as the actual strain which GEO600 has observed.
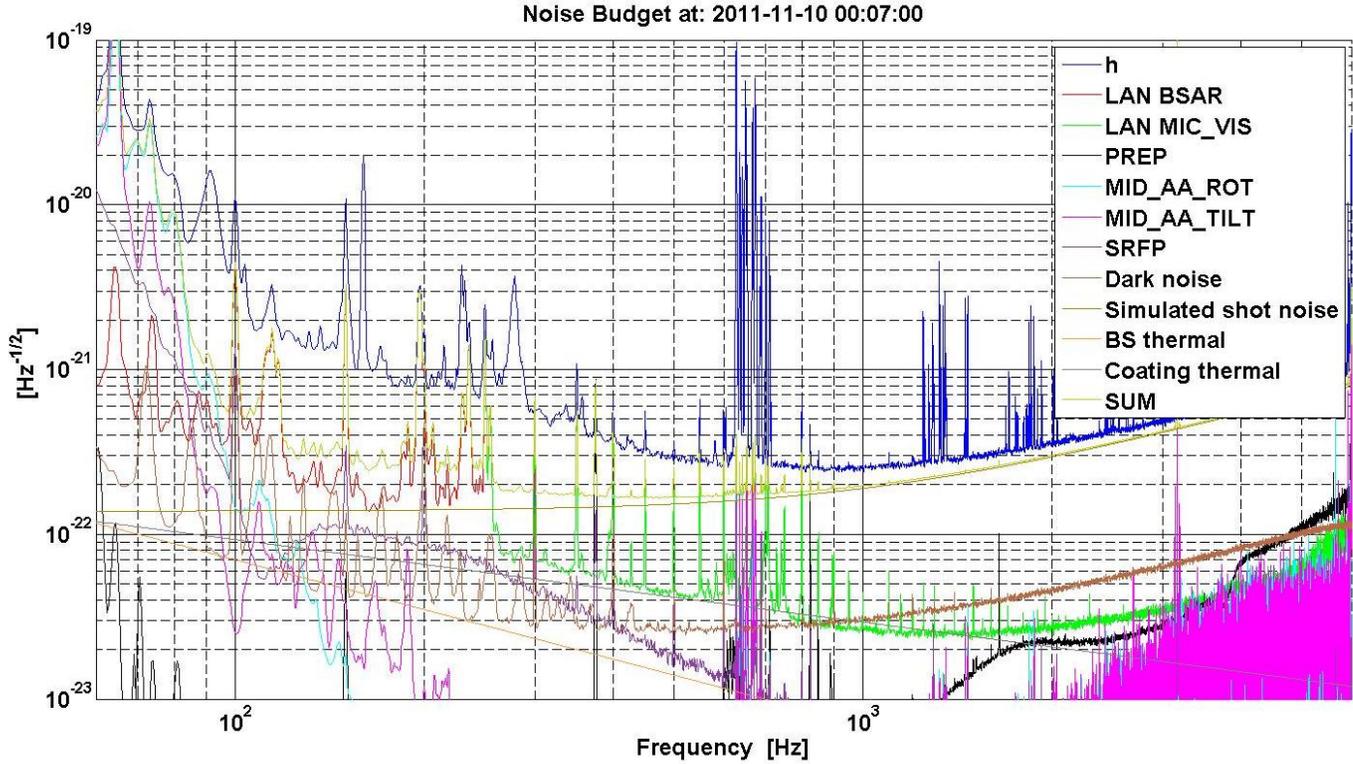
Figure 3: Shows the current GEO600 noise budget. SUM indicates the sum of each of the noise sources, and h indicates the actual strain observed by GEO600. Each of the other curves is an explained noise source which has been modeled by the noise projection techniques described in 1.1.

As shown in the figure, the sum of the explained noise sources does not match the strain observed by GEO600, indicating that there exists substantial unexplained noise. For the purposes of this report, only the frequencies from approximately 100 to 1000 Hz were used in the models. In this frequency range, the missing noise appeared to behave in $\frac{1}{frequency}$ type fashion. Thus, the first model which was chosen to approximate the missing noise was a power law modeling of the unexplained noise. The equation used was:

$$power\ law\ equation = \frac{1 \times 10^y}{frequency^z}$$

Initially this term was added to the SUM term taken from the data presented in Figure 3, and the result was compared to the observed strain, h. To create the sum of the explained noise and the power law equation, each source is added in quadrature, as is done with each explained noise source. Thus this simple model was written as:

$$power\ law\ model = \sqrt{(SUM)^2 + \left(\frac{1 \times 10^y}{frequency^z}\right)^2}$$

8

To compare this model to the observed strain, the natural logarithm of a Gaussian likelihood distribution was used. The logarithm was used due to extremely small nature of the values of the noise model, using the logarithmic function allowed for more manageable calculations, as well as values which our means of computing were better suited to working with. The following equation is the likelihood function which was used, where plm is the above power law model, h is the strain observed by GEO600 smoothed as explained above, and σ^2 is the variance, which was found to be 10^{-22} (this is its value in all further instances of its use).

$$natural \log of\ gaussian\ likelihood = ln\left(\frac{1}{\sqrt{2\ \sigma^2\pi}}e^{\frac{-(h-plm)^2}{2\sigma^2}}\right)$$

This likelihood was calculated at each of the data points of h and SUM between 10^2 and 10^3 Hz. Thus to compute the overall likelihood at any given value of the power law model, the values of this likelihood were summed over all of these data points. The values were summed because this is a logarithmic likelihood, if a logarithmic likelihood were not being used, each of these probabilities would be multiplied since $prob(a, b, …) = prob(a) \times prob(b) \times …$ .Thus $ln(prob(a, b, …)) = ln(prob(a) \times prob(b) \times …) = ln(prob(a)) + ln(prob(b)) + \cdots$.

Initially, a matlab script was created to calculate the value of these likelihoods while varying the parameters of the power law model, y and z. The value of y was varied from -20 to -17, and the value of z was varied from 0.5 to 2.5, and in each range there were 100 values of the parameters selected to calculate the likelihood at. Overall this evaluated the likelihood function at 10,000 values of the parameters. Once these values were calculated, the maximum value of the likelihood was determined, as well as the values of the parameters which yielded this value. The value of the power law model was then calculated for these values of the parameters and this was subsequently compared to the observed strain. The figure below shows the results of this method creating a simple model which completes the noise model of GEO600. The values of the parameters which maximize the likelihood were found to be: y=-18.1010 and z=1.2828, giving a power law equation equal to:

$$\frac{1\times10^{-18.1010}}{frequency^{1.2828}}.$$
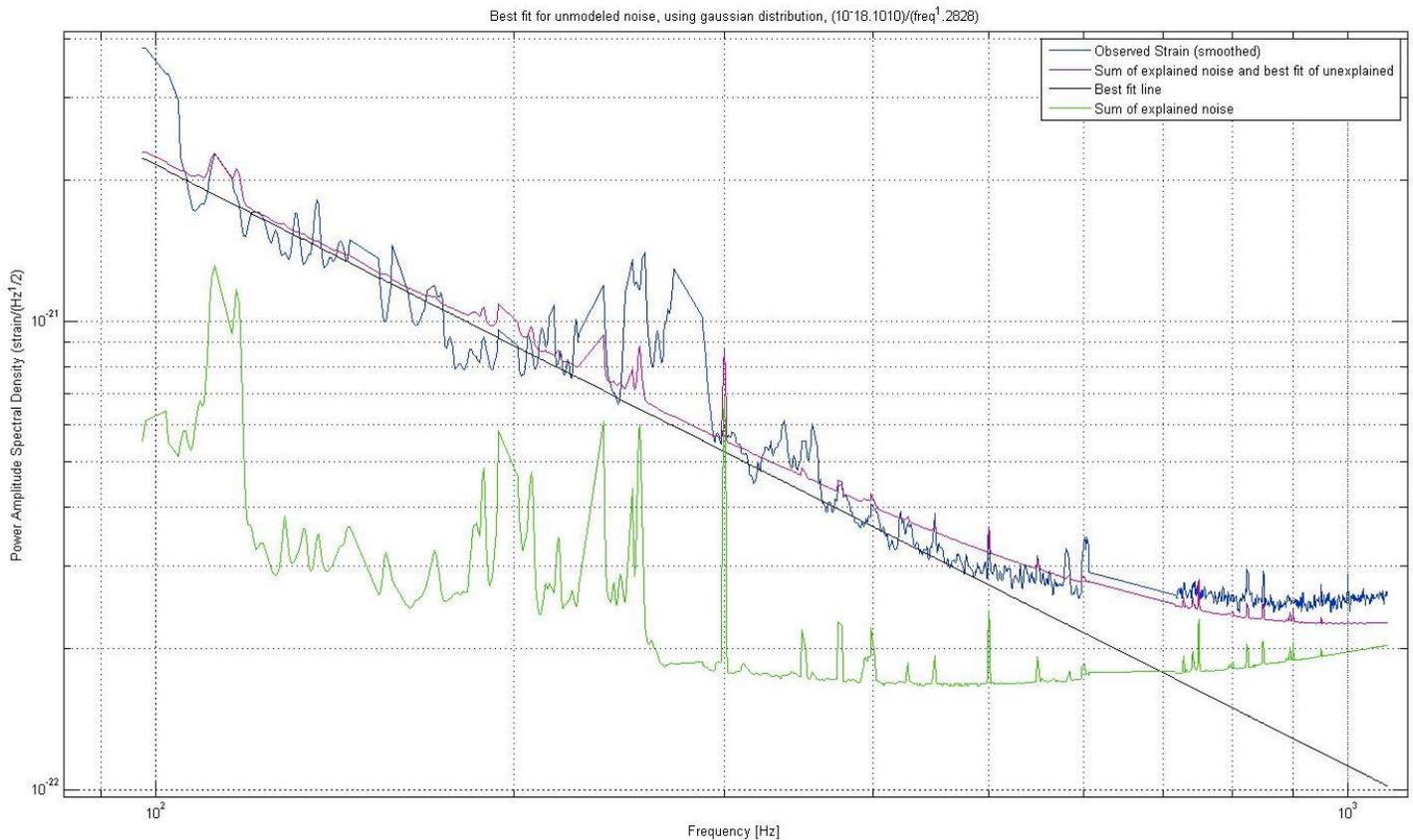
Figure 4: This figure is the result of the initial creation of a simple model to complete the GEO600 noise model. The blue curve indicates the actual strain which GEO600 has observed (after being smoothed), the green curve indicates the sum of the explained noise (SUM), the black line is the value of the power law equation at the parameter values which maximize the likelihood function, y=-18.1010 and z=1.2828, and the red curve is the resultant fit using the power law model.

After this method of choosing parameters to create the best fit was completed, the next step was to create a method which would make use of an MCMC. The first model that the MCMC program created for and used on was the power law model, the same as above, so that a comparison could be made between the two, and the MCMC could be tested for functionality. The first step of the MCMC program was to set up the prior distribution which the program would sample from. First the prior ranges are set, in this case the y values ranged from -18.6 to -18.0, the narrowing of the range allowing for a more targeted exploration of the parameter space. The z parameter ranged from 0.9 to 1.5, again a more narrow range to allow sampling over a more targeted area. The prior distribution was established from these ranges, the range of a parameter being that parameter's maximum value minus its minimum value. In each case, the prior probability was chosen to be $\frac{1}{range\ of\ parameter}$. As discussed in section 1.3, the prior probability is to be multiplied by the likelihood function in order to yield the posterior probability.

The next step in the MCMC was to establish the proposal covariance matrix. In this case it was decided that the parameters had no covariance, and this matrix was simply a diagonal matrix with the

10

initial values that the MCMC would use for sampling on the diagonal. For instance, in this model this matrix was $\begin{bmatrix} initial\ y\ value & 0 \\ 0 & initial\ z\ value \end{bmatrix}$. From this point it is necessary to have a means of sampling from these prior ranges. A common way of choosing a random vector from a multivariate normal distribution is to use the Cholesky decomposition of the covariance matrix. The Cholesky decomposition is a matrix **A** such that **AA**$^{\text{T}}$ is equal to the covariance matrix of the distribution, and this is the method that was used in this case, the Cholesky decomposition of the above proposal decomposition was taken and then used each time a new point was sampled from the prior ranges.

Once the Cholesky decomposition was established the MCMC began. The first step of the MCMC was to evaluate the likelihood function at initial values of the parameters y and z. After this was evaluated, the posterior distribution was calculated by multiplying the value of the likelihood function by the prior probabilities. A logarithmic Gaussian likelihood was again being used, so the posterior probability now took the form:

$$posterior\ probability$$
$$= \ln\left(\frac{1}{\sqrt{2\ \sigma^2 \pi}} e^{\frac{-(h-plm)^2}{2\sigma^2}}\right) + \ln\left(\frac{1}{y\ range}\right) + \ln\left(\frac{1}{z\ range}\right)$$

The natural logarithm of the two prior probabilities was added to natural logarithm of the likelihood function since this is equivalent to multiplying the likelihood function by the prior probabilities and subsequently taking the natural logarithm of the result.

From this point, a new set of values of the two parameters was drawn from the two prior distributions. This was done by creating a random matrix the size of the Cholesky decomposition, with each entry in the matrix greater than zero and less than one. The Cholesky decomposition and this matrix were then multiplied together. The entry of the resulting matrix corresponding to the y variable was then added to the value of y that had just been used to evaluate the likelihood function and the same was done for z. It is in this way that the size of the Cholesky decomposition controlled the size of the steps that the Markov chain took as it explored the parameter spaces. Each new parameter value was chosen simply by adding to the previous parameter value a random multiple of the value of its corresponding row in the Cholesky decomposition. Thus an important part of choosing the entries in the proposal covariance matrix (which is what determines the Cholesky decomposition) was to choose an appropriate scaling factor that allowed sufficiently small steps through the parameter space so that each area was explored thoroughly. These scaling values were chosen through some trial and error and by ensuring that whatever was chosen allowed for a sufficient amount of exploration of the parameter space, while simultaneously not being so large as to never converge to a value that maximizes the posterior probability. For the power law model, 0.005999 was chosen as the scaling value for the y parameter, and .01555 was chosen for the z parameter.

Once a set of new values for the parameters were chosen, it was made sure that these new values did not fall outside the ranges of their parameters, if one was smaller than its parameter's minimum or larger than its maximum then the set was rejected and a new set was chosen. Once an

appropriate set of values was found, the likelihood function was then evaluated using these new parameter values. The posterior probability was once again calculated by adding the natural logarithm of the prior probabilities to the new value of the likelihood function. At this point a ratio was taken between the new posterior probability and the previous one. If this ratio was greater than one, signifying that the new point had a greater probability than the previous one, then this point was accepted and the next set of parameters could be chosen by adding a random multiple of the Cholesky decomposition to this set of parameters. If the ratio was less than one, then the new set of parameters was rejected and a new set was chosen based on the previous set.

This process was repeated 20,000 times. The first 10,000 times was known as the burn in, these iterations were used to explore the parameter space and arrive at the area of the parameter space which maximized the posterior probabilities. The last 10,000 iterations were the ones which were used in the final posterior probabilities. The set of values that each parameter took on over the course of this process is known as the chain of that parameter. Over the course of this process a ratio of the accepted values in the chain versus the total number of values that were tested was kept, both for the burn in and the iterations after the burn in known as the acceptance ratio of the chain. For the power law model the acceptance ratio during the burn in was 0.14, and that of the chain was 0.13. The maximum y value was found by this method to be -18.1415, which was reasonably similar to the previous method's -18.1010. The maximum z value found by this method was 1.2635, also similar to the previous method's 1.2828. The following figure shows probability density functions of the y and z parameters, and the subsequent figure displays the values of the y and z chains, as well as the chain of the posterior probability values, the figure after that shows the result of the model using the parameter values that maximize the posterior probability.
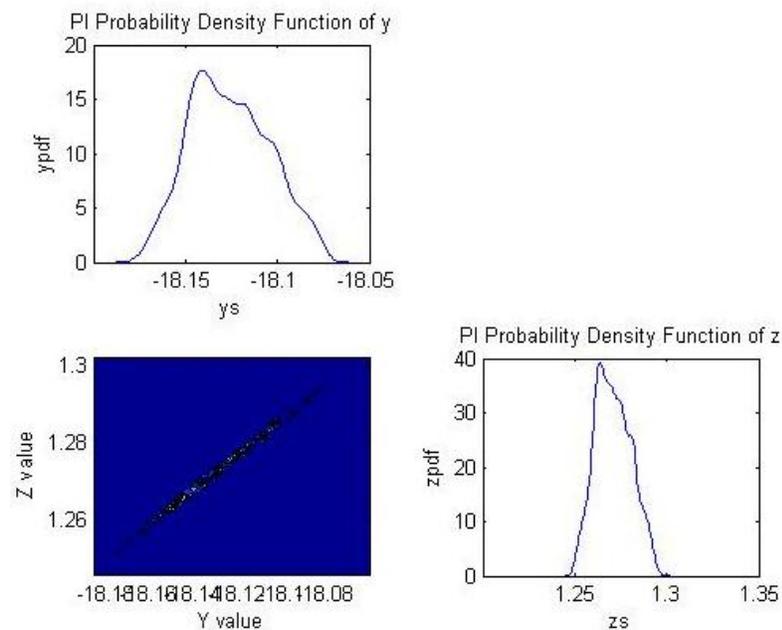

Figure 5: The figure to the left shows the probability density functions of the y and z parameters in the power law model. The top left figure is the y pdf. The bottom left figure is a contour plot of the overall posterior probability. The bottom right figure is the z pdf.
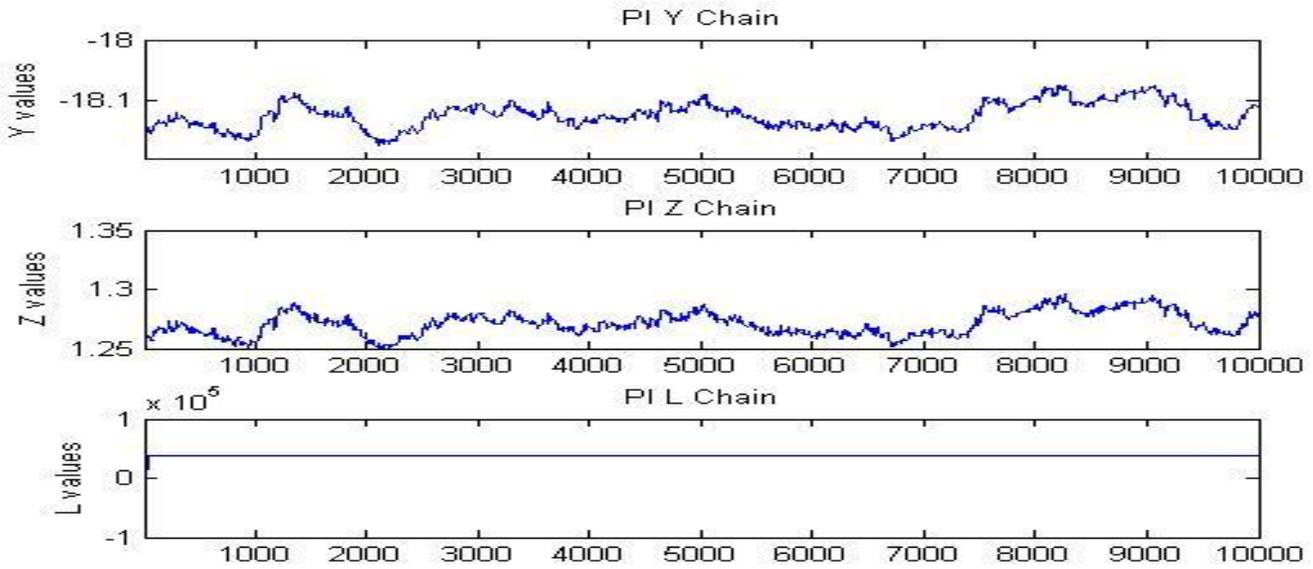
Figure 6: The figure above shows the values which y, z and the posterior probability (shortened to L for likelihood), took on over the course of the last 10,000 iterations of the MCMC. The top plot is the y chain, the middle is the z chain, and the bottom is the posterior probability chain. As can be seen in the posterior probability chain, by the time the burn in portion of the MCMC had been completed, the values of y and z which maximized the posterior had already been established, and the posterior probability was nearly stabilized.
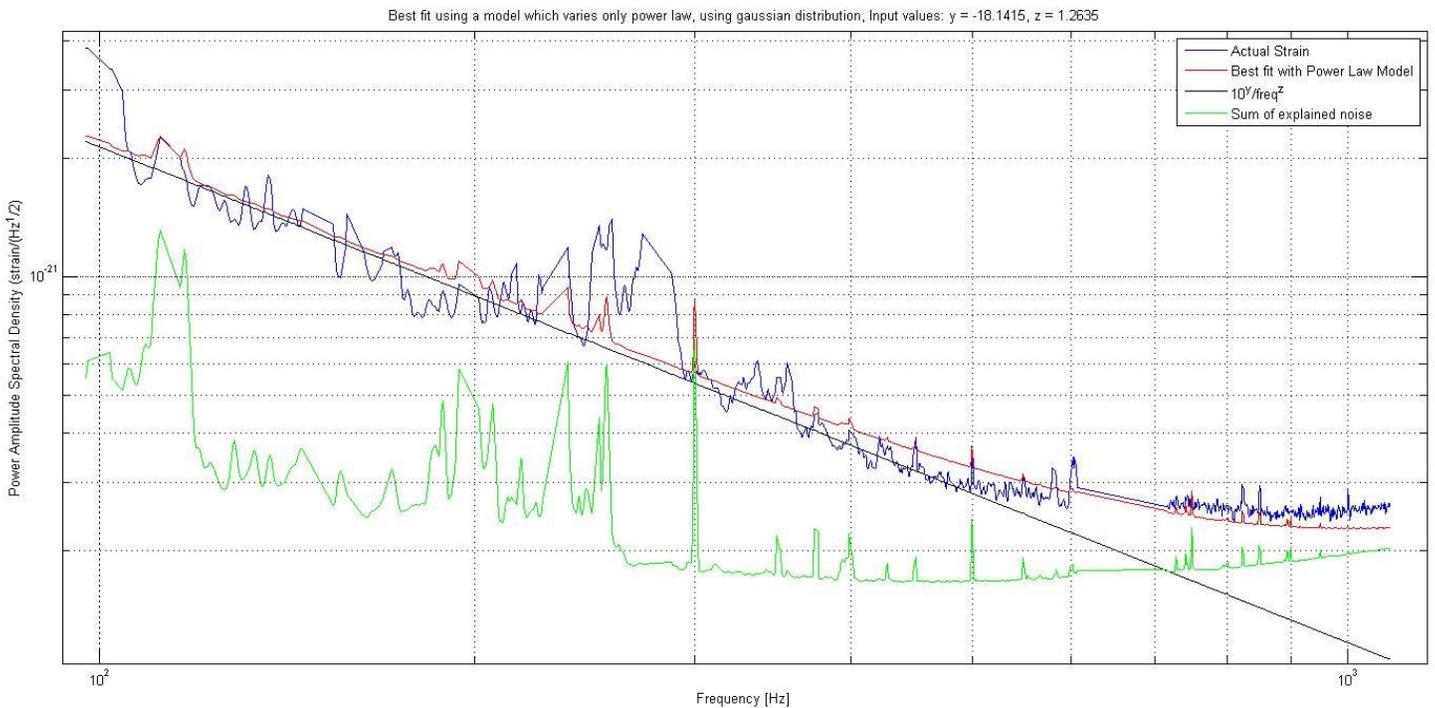
Figure 7: The figure above is the result of the MCMC method of completing the GEO600 noise model. The blue curve indicates the actual strain which GEO600 has observed (after being smoothed), the green curve indicates the

sum of the explained noise (SUM), the black line is the value of the power law equation at the parameter values which maximize the likelihood function, y=-18.1415 and z=1.2635, and the red curve is the resultant fit using the power law model. This figure should be compared to Figure 4, which was the result of the initial method of completing the GEO600 noise budget using the same model, the power law model.

The next model tested was one that simply broke SUM down into five of its largest components and made no attempt to complete the noise budget (as a shorthand name this model was called brksumnoplm, for broken apart SUM with no power law model). This was done so as to have a point of comparison for later models which would have the broken downsum as well as several other components attempting to complete the noise budget. The first step was to determine which of the noise projections made the largest contributions to the values of SUM. These were found to be LAN BSAR and LAN MIC_VIS, both of which are types of laser amplitude noise, shot noise, dark noise, and SFRP. The result of this breakdown is shown below where the red curve indicates the reconstructed sum and the blue curve indicates its actual value.
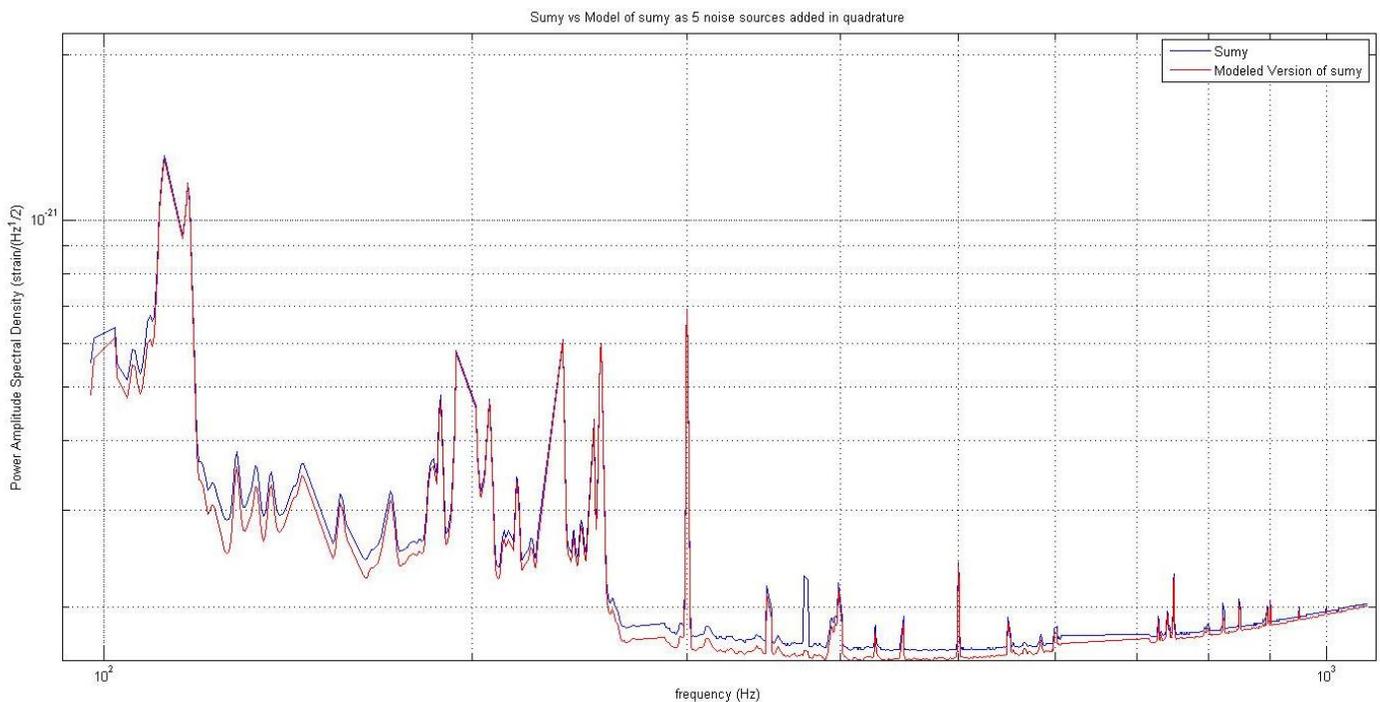
Figure 8.

This model then took the form:

$$brksumnoplm = \sqrt{(a * lanb)^2 + (b * lanm)^2 + (c * shot)^2 + (d * dark)^2 + (e * srfp)^2}$$

Where lanb is the values of LAN BSAR and lanm is the values of LAN MIC_VIS. The coefficient in front of each noise projection is a measure of the error in each of those projections, and these are the values which will be determined by the MCMC. Each of these coefficients a, b, c, d, and e scale these projections until they give the best fit to the data. The method for this MCMC was precisely the same as

14

that of the power law model, merely the power law model was interchanged for the brksumnoplm model and the number of parameters was expanded. A logarithmic Gaussian distribution was again used as the likelihood function and the prior probabilities were once again set as $\frac{1}{parameter\ range}$. The scaling factors chosen to control the size of the Cholesky decomposition were 0.0009 for each error parameter: a, b, c, d, and e. Initially the parameters were permitted to range from zero to one hundred; however a more reasonable range would be within twenty percent above or below their current value, thus a second MCMC was run with a range between 0.8 and 1.2. The differences which altering the ranges caused can be seen in the figures below.

Model using a sum value that is broken down into dark,shot,lan bsar,srfp, and lan mic-vis noise (reffered to as brksumnopl model). Best fit, using gaussian distribution Values of parameters: a = 1.8345, b = 0.8536, c = 1.1978, d = 4.1987, e = 10.6432. Noise curves plotted with their respective coefficients
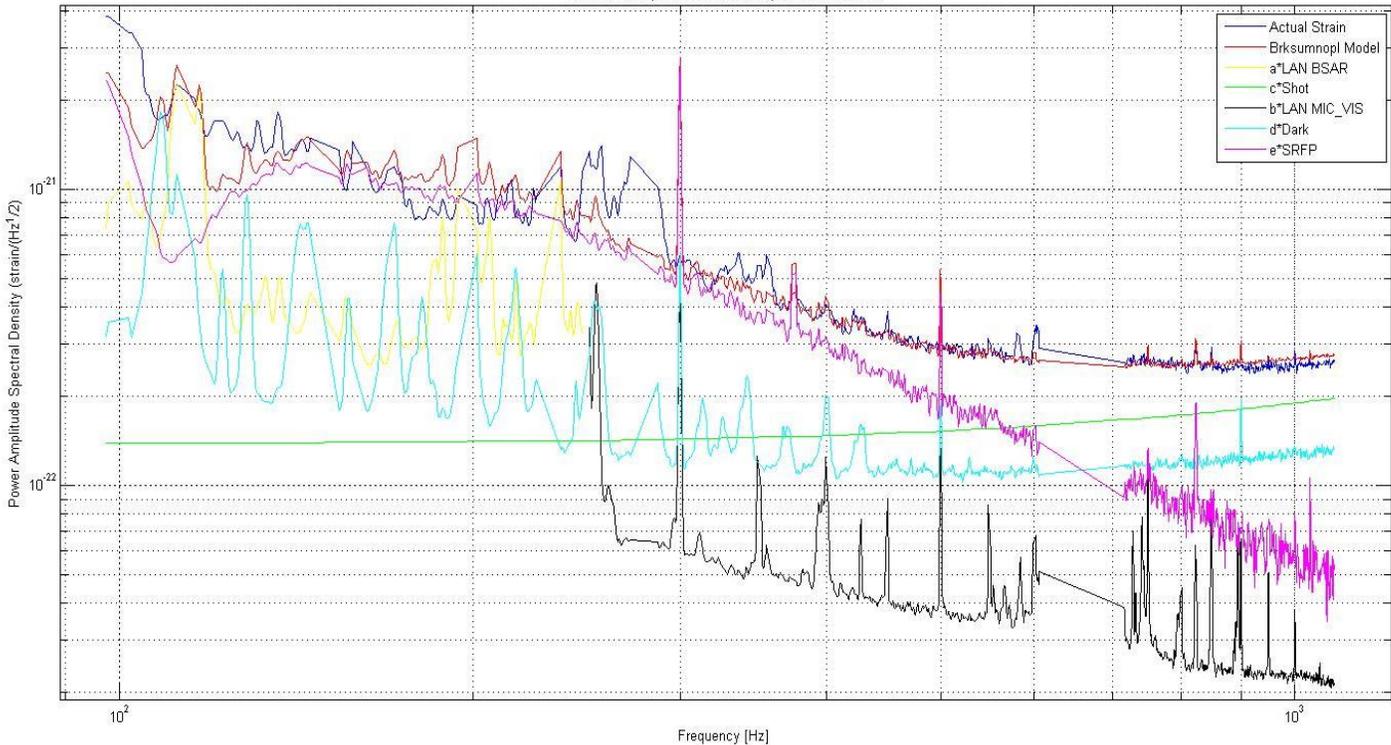
Figures 9 &10: These figures show the results of the MCMC for the model known as brksumnoplm. In each graph the noise projections have been multiplied by their corresponding coefficients determined by the MCMC to maximize the posterior probability. Also in each figure the observed strain is the darker blue curve and the best fit which the model yielded is the red curve. The above graph has permitted the error parameters: a, b, c, d, and e to range from 0 to 100, with the maximizing parameters taking the following values: a=1.8345, b=0.8563, c=1.1978, d=4.1967, and e=10.6432. In this instance the acceptance ratio during the burn in was 0.10 and that of the chain was 0.09. The figure below has limited the range of the error parameters: a, b, c, d, and e to 0.8 to 1.2, with the maximizing parameters taking the values: a=1.1999, b=1.1066, c=1.1999, d=1.1937, and e=1.994. In this case the acceptance ratio during the burn in was 0.36 and that of the chain was 0.24.

15

Model using a sum value that is broken down into dark,shot,lan bsar,srfp, and lan mic-vis noise (reffered to as brksum model)
.Best fit, using gaussian distribution Values of parameters: a = 1.1999, b = 1.1066, c = 1.999, d = 1.1937, e = 1.1994
Noise curves plotted with their respective coefficients
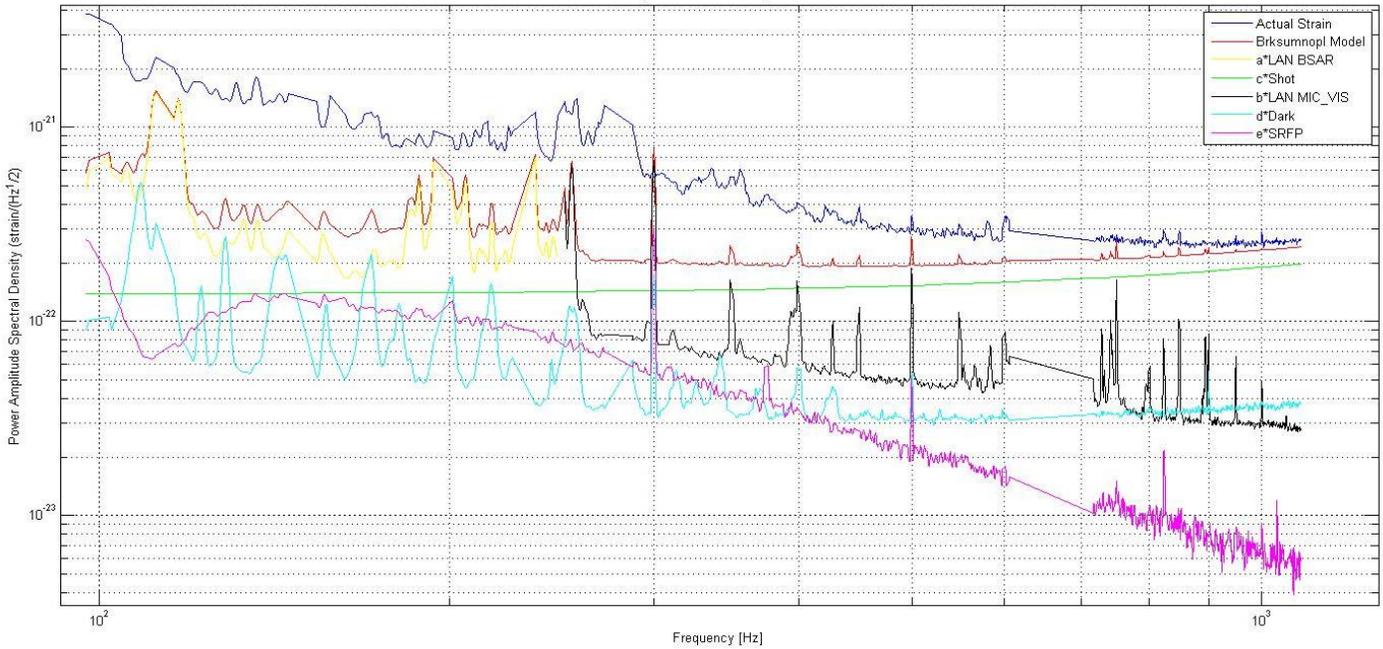
Figure 11: The figure below shows the results of brksumnoplm model with the parameters ranging from 0.8 to 1.2, however unlike the above graphs, the noise projections in this graph have not been multiplied by the coefficients determined by the MCMC to maximize the posterior probability. This graph should be compared to the above figure, and the changes in the noise curves should be noted.

Model using a sum value that is broken down into dark,shot,lan bsar,srfp, and lan mic-vis noise (reffered to as brksum model).
Best fit, using gaussian distribution Values of parameters: a = 1.1999, b = 1.1066, c = 1.999, d = 1.1937, e = 1.1994
Noise curves plotted without their respective coefficients
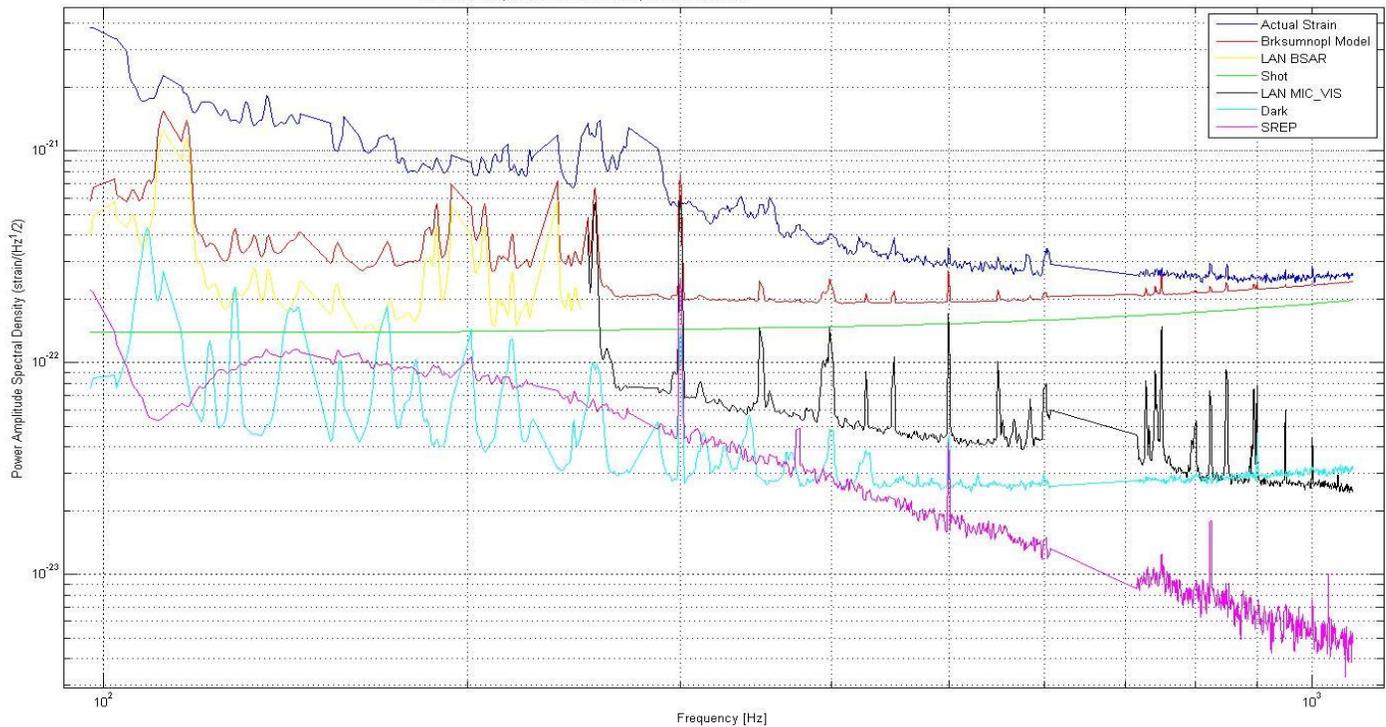
It should be noted that when the parameters a, b, c, d, and e are limited to a reasonable range, the resulting fit to the observed strain is quite poor, indicating that the noise model is indeed incomplete and other forms of noise projections are necessary. If a prior probability that reflected more knowledge about the model were developed, rather than using the flat prior of $\frac{1}{parameter\ range}$, then the ranges of these error parameters could be determined with more accuracy. However, given that is not the case, for the other models making use of a broken down representation of SUM the 0.8 to 1.2 range was used since it is a practical range for the error of these values. Below are the a, b, c, d, and e pdfs and chains for the brksumnoplm model, with the error parameters ranging from 0.8 to 1.2.
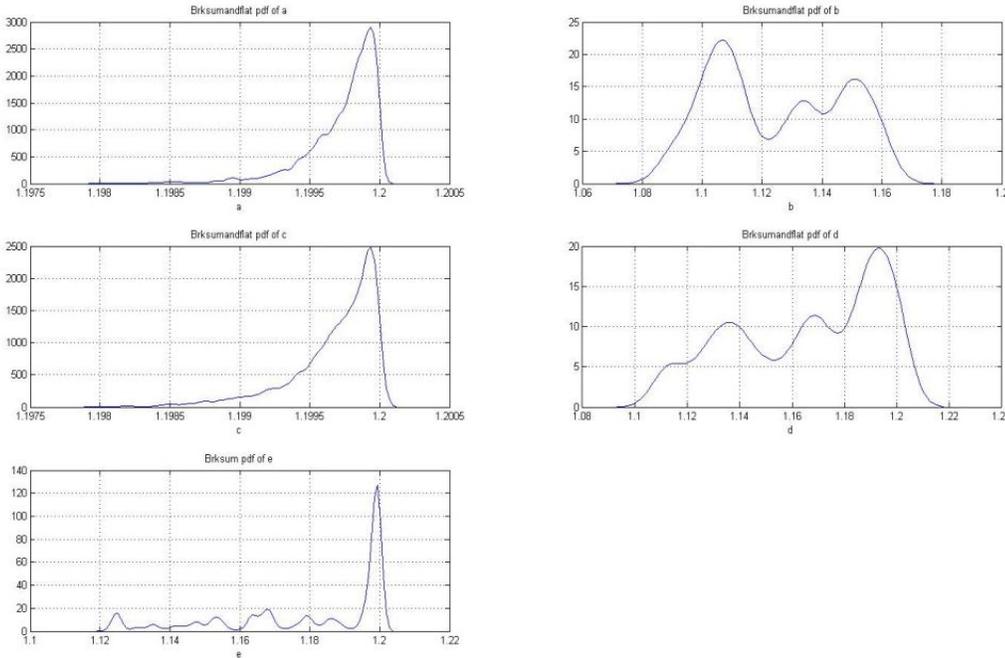
Figures 12 & 13: The figure to the left shows the pdfs of the error parameters a, b, c, d, and e, and the below figure shows their chains, as well as the posterior probability chain. As can be seen the graph, the posterior probability reaches a reasonably stable level by approximately the 3000th iteration.

17
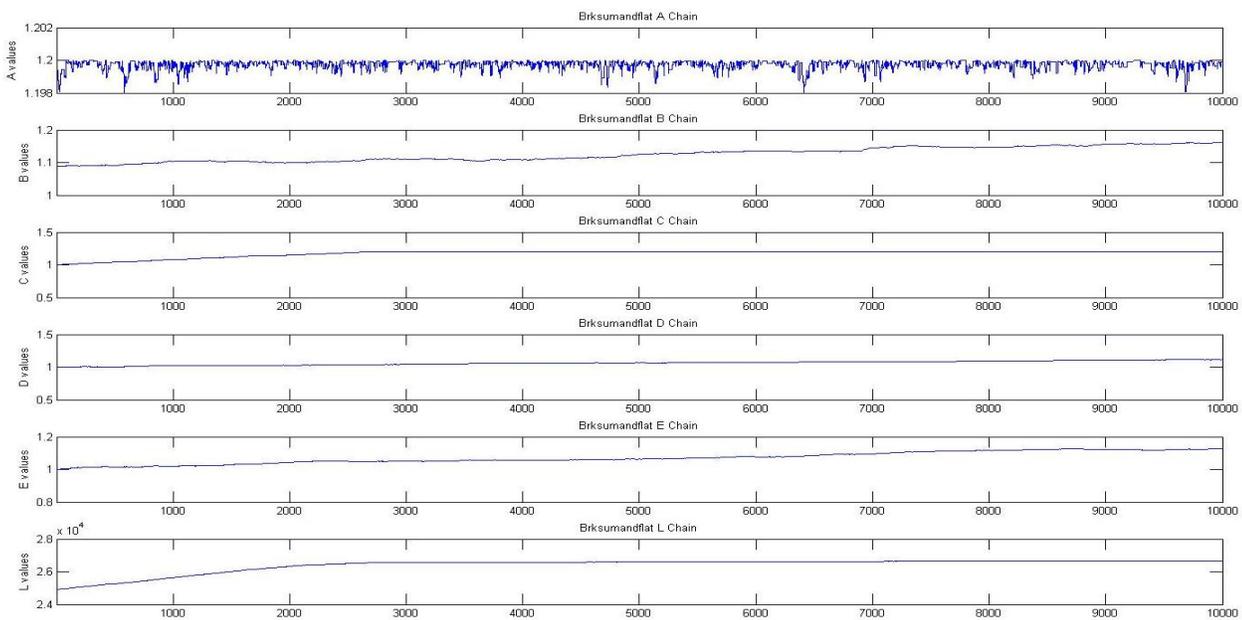
The next model which was tested using the MCMC method was a model which combined the power law model with the broken down version of SUM. Again, the procedure of the MCMC was the same as the simple power law model case, a logarithmic likelihood was again used, and the number of parameters was expanded. In this case the model was referred to as brksum, and it took the form:

$brksum$

$$= \sqrt{(a*lanb)^2 + (b*lanm)^2 + (c*shot)^2 + (d*dark)^2 + (e*srfp)^2 + \left(\frac{1 \times 10^Y}{frequency^z}\right)^2}$$

The scaling factors chosen to control the size of the Cholesky decomposition were chosen to be the same as they had been in both of the previous models; y:0.0005999, z:0.001555, and a, b, c, d, and e: 0.0009. The y parameter was permitted to vary between -18.6 and -18, z from 0.9 to 1.5, and the error parameters from 0.8 to 1.2. The acceptance ratio during the burn in was 0.7 and that of the chain was 0.72, thus more than half of the values sampled were kept. The following figures show the result of this method.
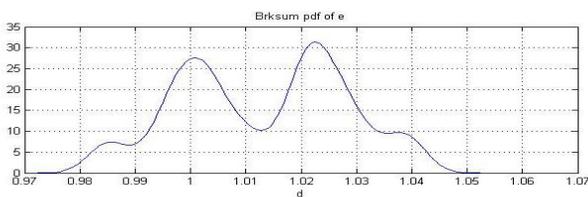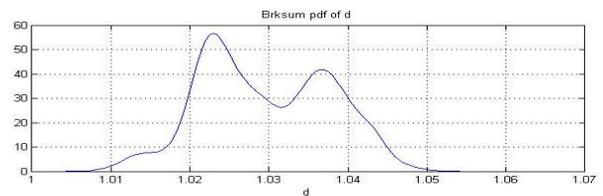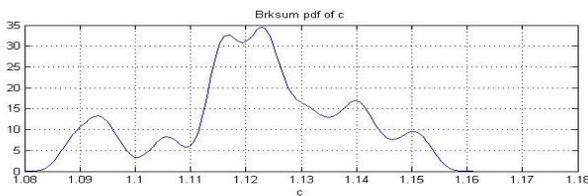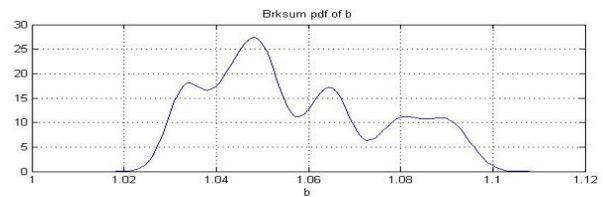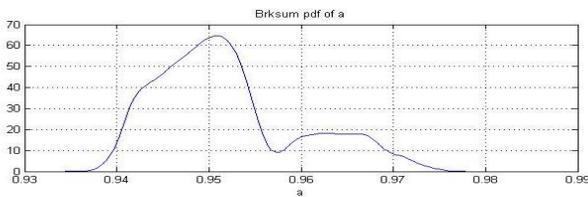
Figures 14 & 15: The figure above is the a, b, c, d, and e pdfs for the brksum model, and the figure below are their chains. The values of the error parameters which were found to maximize the posterior probability were as follows: a=.9511, b=1.0481, c=1.1227, d=1.0229, and e=1.0225.
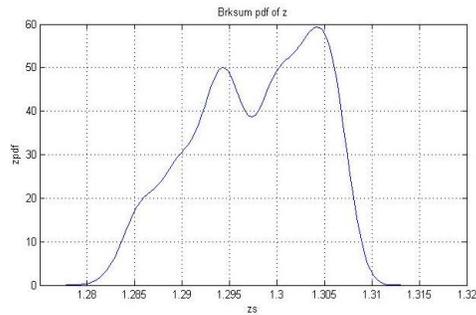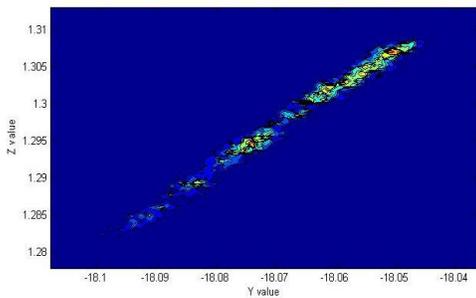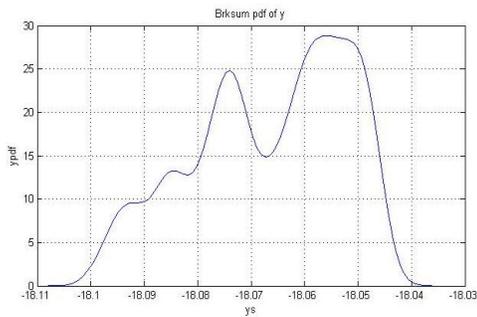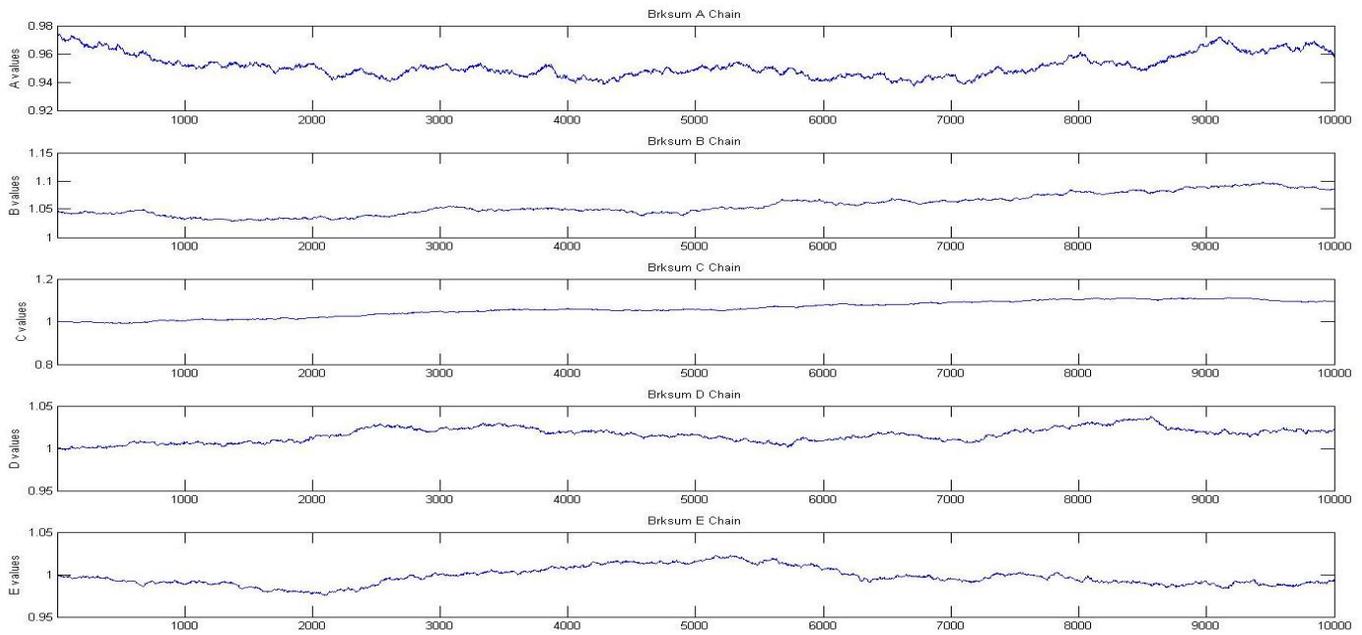
Figures 16 & 17: The figure to the left is the y and z pdfs for the brksum model. In the top left of this figure is the y pdf, the bottom right is the z pdf, and the bottom left is a contour plot of the posterior probability over the y and z paramerers. The figure below displays the y,z and L (posterior probability) chains. The values of y and z which maximize the posterior probabiliy are as follows: y=-18.0558 and z=1.3041.
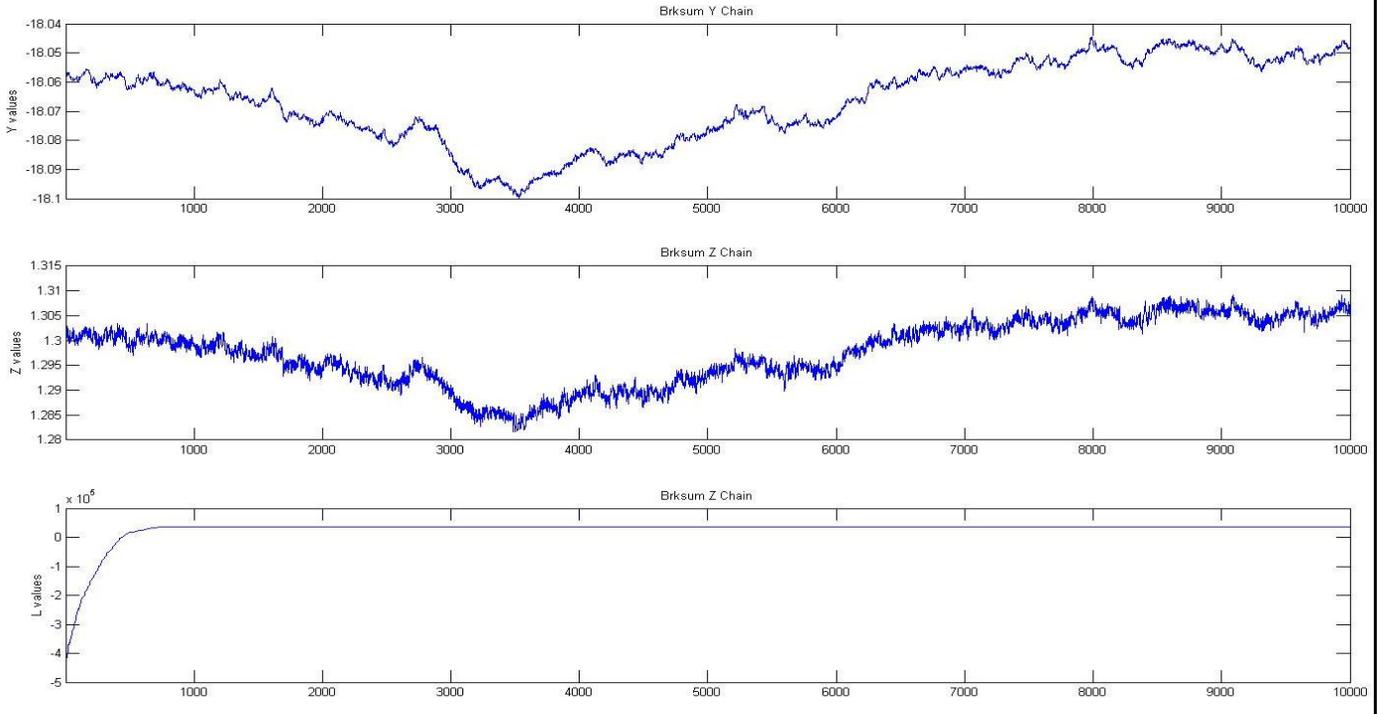
Figures 18 & 19: These figures show the results of the MCMC for the model known as brksum. In the above graph the noise projections have been multiplied by their corresponding coefficients determined by the MCMC to maximize the posterior probability. In the figure below the noise curves have not been multiplied by their coefficients. Also in each figure the observed strain is the darker blue curve and the best fit which the model yielded is the red curve.

Model using a sum value that is broken down into dark,shot,lan bsar,srfp, and lan mic-vis noise, as well as a power law variable (reffered to as brksum model). Best fit, using gaussian distribution Values of parameters: y = -18.0558, z = 1.3041, a = 0.9511, b = 1.0481, c = 1.1227, d = 1.0229, e = 1.0225
Noise curves plotted without their respective coefficients
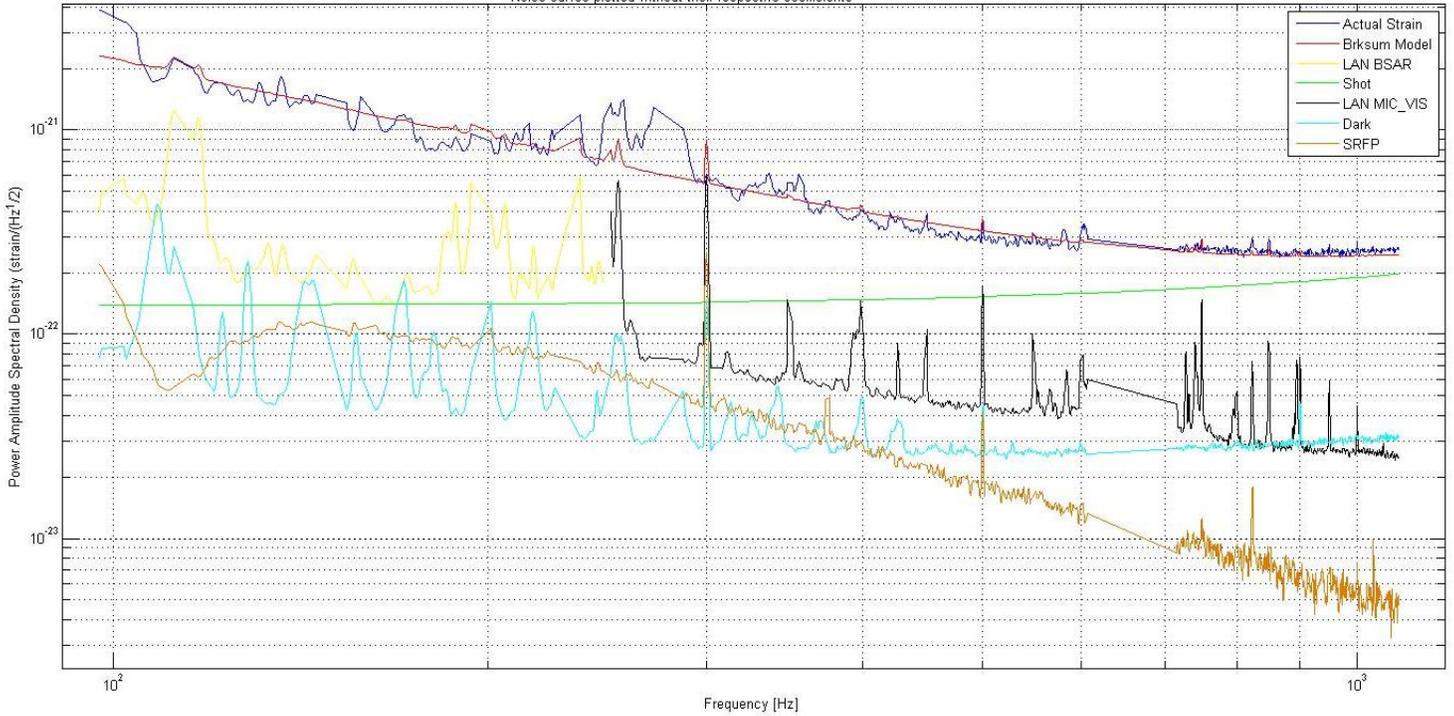
The MCMC method was then used to test a model which combined the broken down version of SUM and a noise source which was completely flat, just a constant number. The idea with this model was to see if the flat noise which maximized the posterior probability would match or be near to Craig Hogan's prediction for holographic noise. The procedure for the MCMC was the same as the previous models, a logarithmic likelihood was used once more, and the ranges of the parameters were expanded. This model was called brksumandflatnoplm, and it took the form:

$$brksumflatnopl$$
$$= \sqrt{(a * lanb)^2 + (b * lanm)^2 + (c * shot)^2 + (d * dark)^2 + (e * srfp)^2 + (flat)^2}$$

The scaling factors chosen to control the size of the Cholesky decomposition were chosen to be the same as they had been in the brksumnoplm model; a, b, c, d, and e: 0.0009, and the scaling factor for the new parameter flat was chosen as 0.009. The flat parameter was permitted to vary between 10^{-23} and 10^{-21}, and again the error parameters ranged from 0.8 to 1.2. The acceptance ratio during the burn in was 0.28 and that of the chain was 0.22. The following figures show the result of this method.
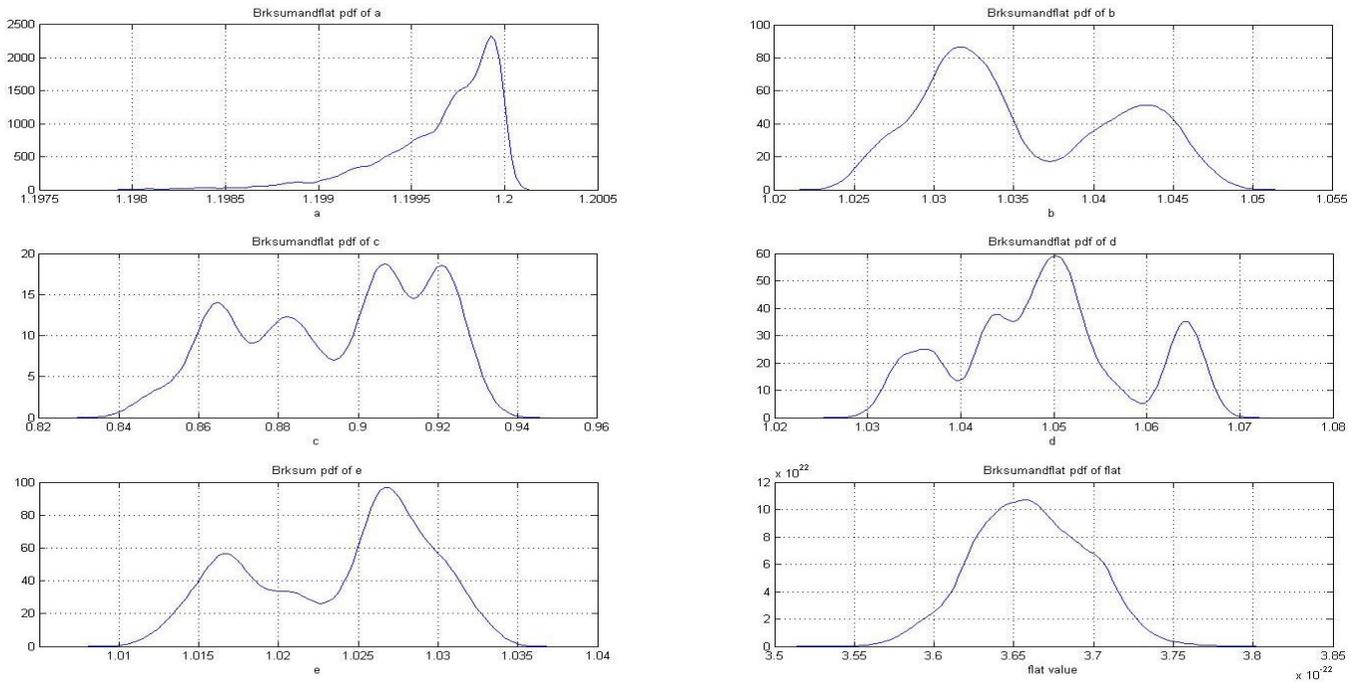
Figures 20 & 21: The above figure shows the pdfs of a, b, c, d, e, and flat, and the figure below shows each of their chains, as well as the chain of the posterior probability. The values of each parameter that maximized the posterior probability were as follows: a=1.1999, b=1.0315, c=0.9067, d=1.0503, e=1.0267, and flat=3.6564×10⁻²². Comparing this value of the flat parameter to Hogan's predicted value of 1.85×10⁻²², we see that flat is only 2.35% larger than Hogan's proposed value.
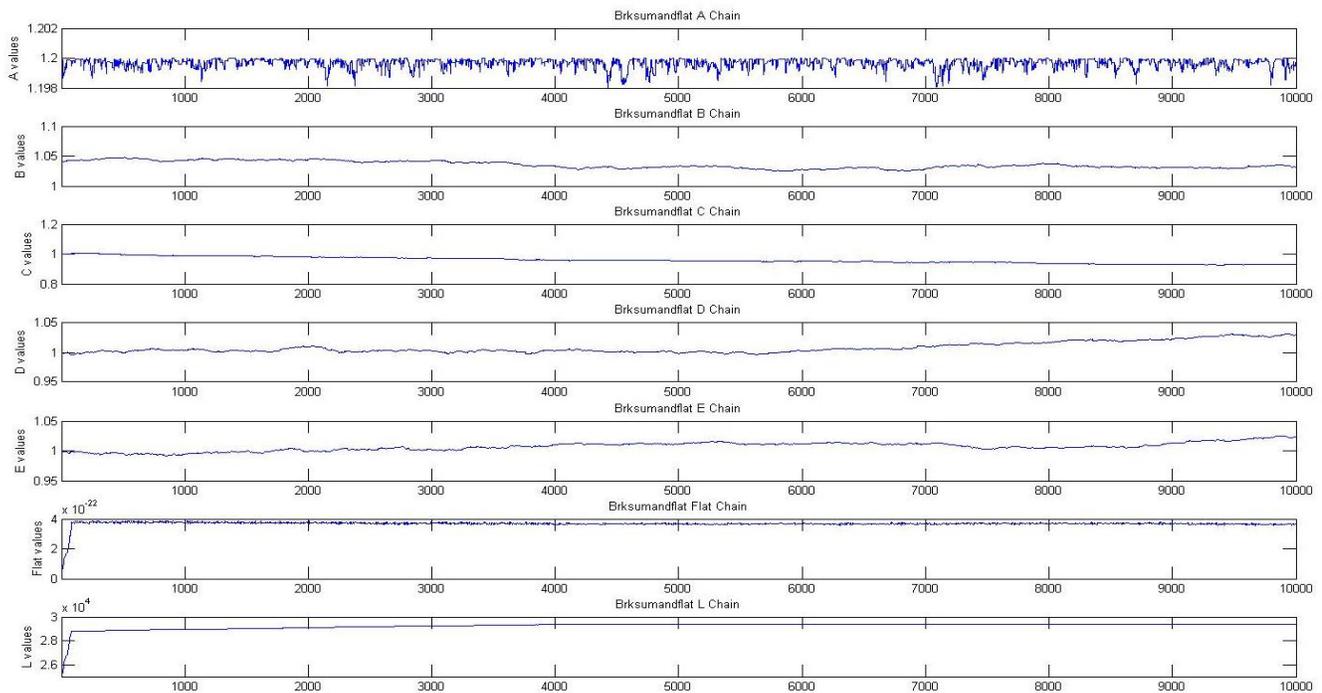
Model using a sum value that is broken down into dark,shot,lan bsar, sfrp, and lan mic-vis noise, as well as varying flat noise (reffered to as brksumflatnopl model).
Best fit, using gaussian distribution Values of parameters: a = 1.1999 b = 1.0315, c = 0.9067, d = 1.0503, e = 1.0267, e = 3.6564e-22
Noise curves plotted with their respective coefficients
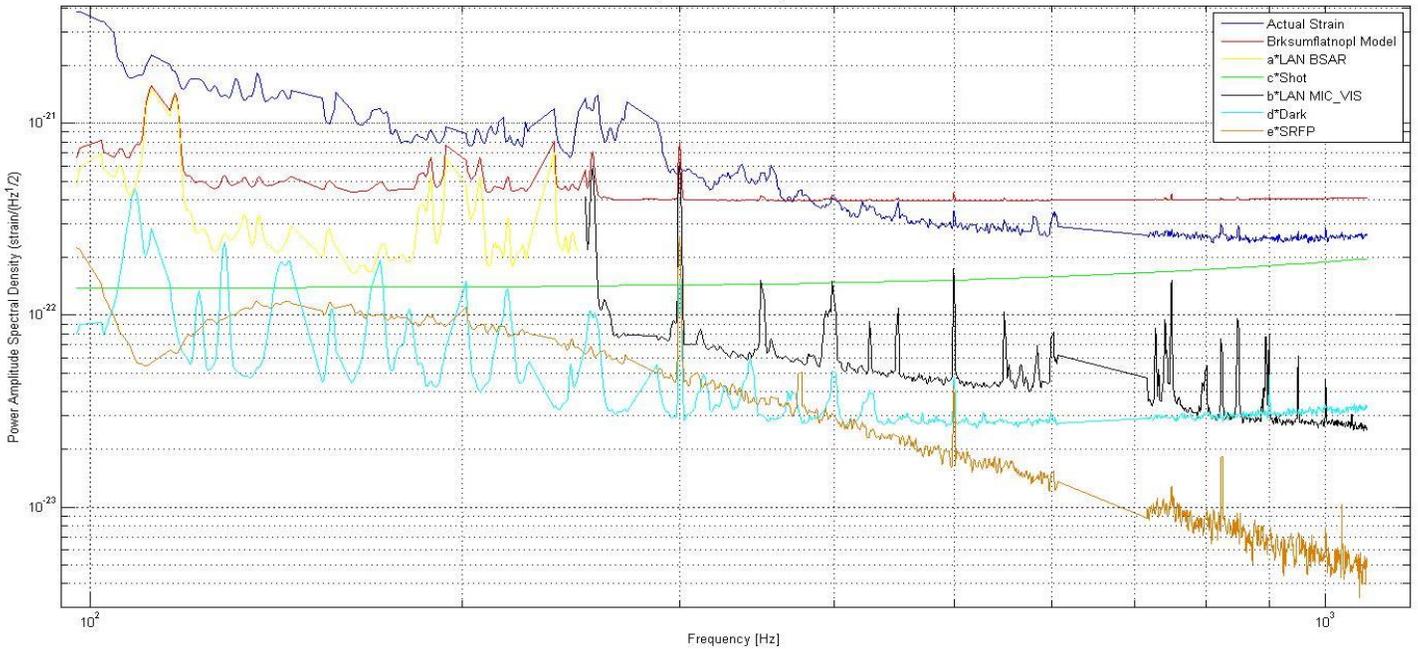
Figures 22 & 23 : These figures show the results of the MCMC for the model known as brksumflatnoplm. In the above graph the noise projections have been multiplied by their corresponding coefficients determined by the MCMC to maximize the posterior probability. In the figure below the noise curves have not been multiplied by their coefficients. Also in each figure the observed strain is the darker blue curve and the best fit which the model yielded is the red curve. It is worth noting that, despite the close match between the flat parameter and Hogan's prediction, the elements of this model alone do not create a good fit to the observed data and there clearly remains noise which this model cannot account for.


Model using a sum value that is broken down into dark,shot,lan bsar,srfp, and lan mic-vis noise, as well as varying flat noise (reffered to as brksumflatnopl model).
Best fit, using gaussian distribution Values of parameters: a = 1.1999 b = 1.0315, c = 0.9067, d = 1.0503, e = 1.0267, e = 3.6564e-22
Noise curves plotted without their respective coefficients
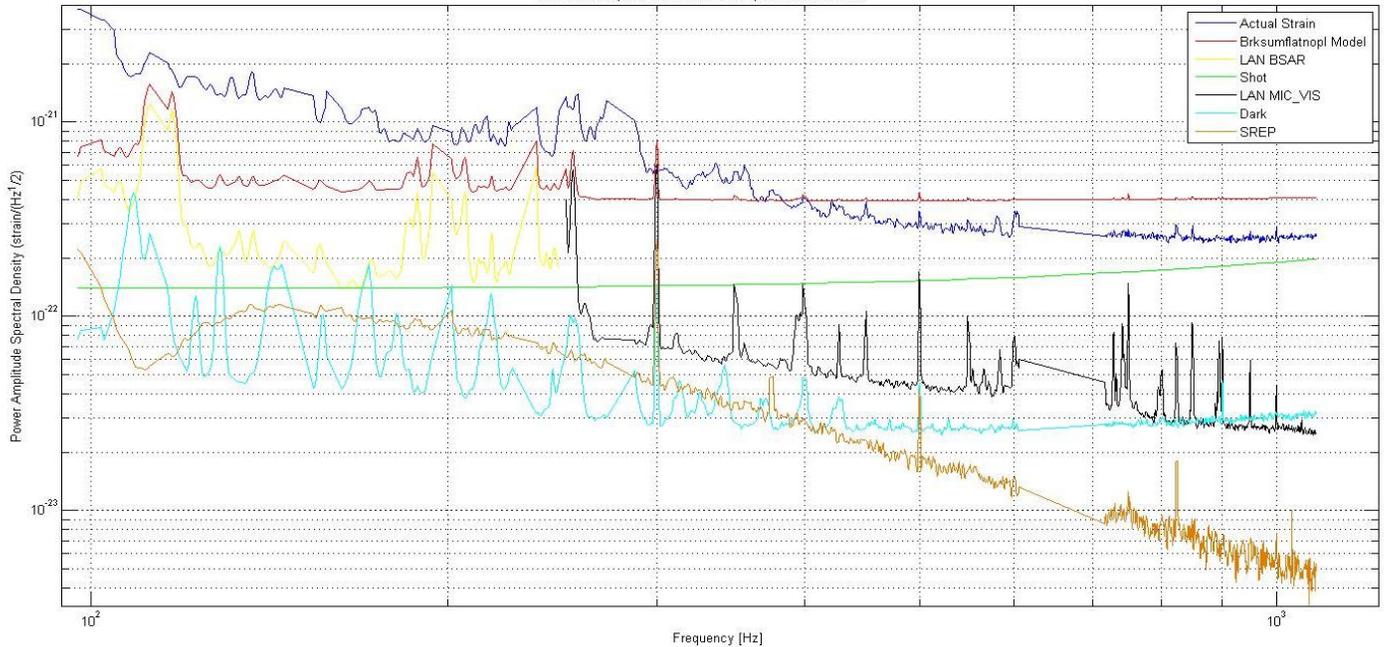
The next model the MCMC method was used to test was called brksumandholo, it combined the broken down SUM and power law models, while also including Hogan's proposed holographic noise component. Brksumandholo had the form:

$$brksumandholo$$

$$= \sqrt{\begin{array}{c}(a * lanb)^2 + (b * lanm)^2 + (c * shot)^2 + (d * dark)^2 + (e * srfp)^2 + \\ \left(\dfrac{1 \times 10^Y}{frequency^Z}\right)^2 + (holo)^2\end{array}}$$

The procedure for the MCMC was the same as the previous models, a logarithmic likelihood was used once more, and the term holo was simply equal to Hogan's prediction of 1.85×10^{-22}. The scaling factors chosen to control the size of the Cholesky decomposition were chosen to be the same as they had been in the brksum model; y:0.0005999, z:0.001555, and a, b, c, d, and e: 0.0009. The y parameter was again permitted to vary between -18.6 and -18, z from 0.9 to 1.5, and the error parameters from 0.8 to 1.2. The acceptance ratio during the burn in was 0.68 and that of the chain was 0.70, thus more than half of the values sampled were kept. The following figures show the result of this method.
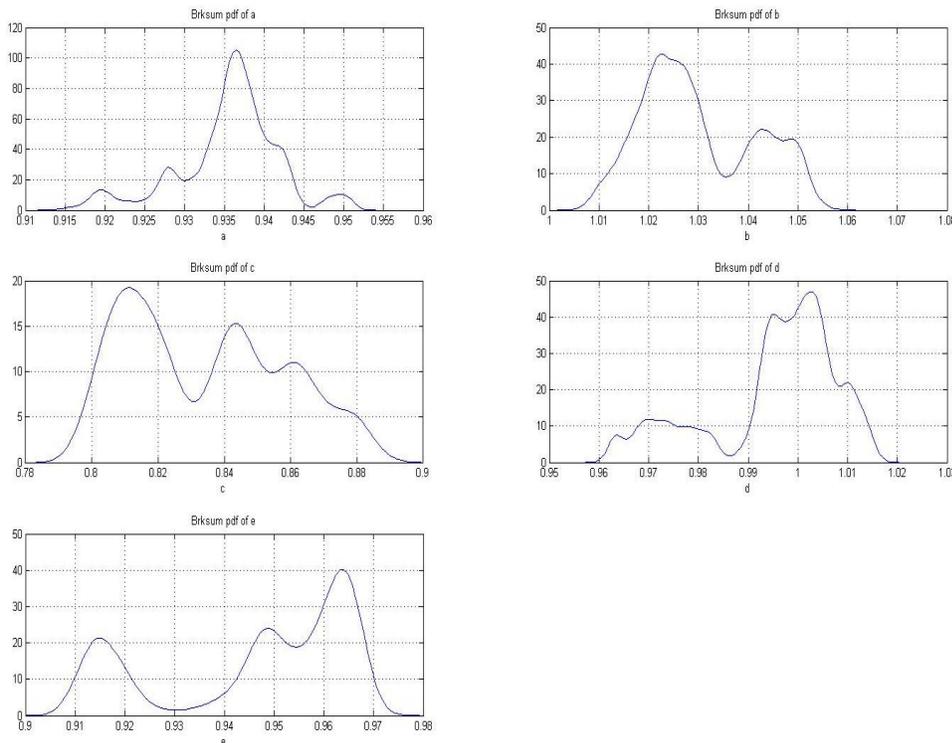

Figures 24 & 25: The figure to the left shows the pdfs of the error parameters a, b, c, d, and e, and the figure below shows their chains. The values of a, b, c, d, and e which maximized the posterior probability are as follows: a=0.9364, b=1.0228, c=0.8114, d=1.0024, and e=0.9640.
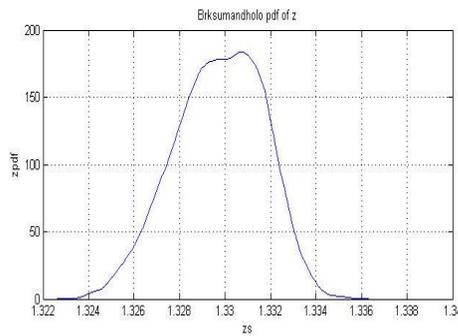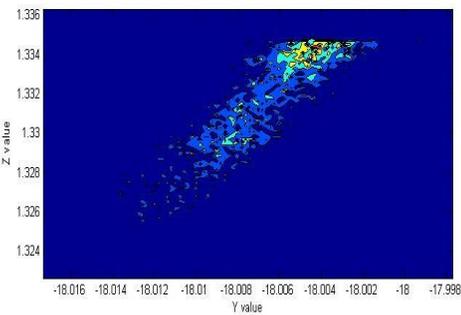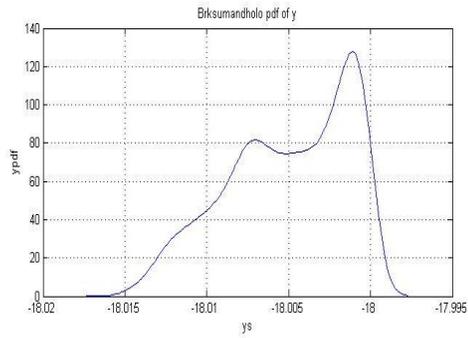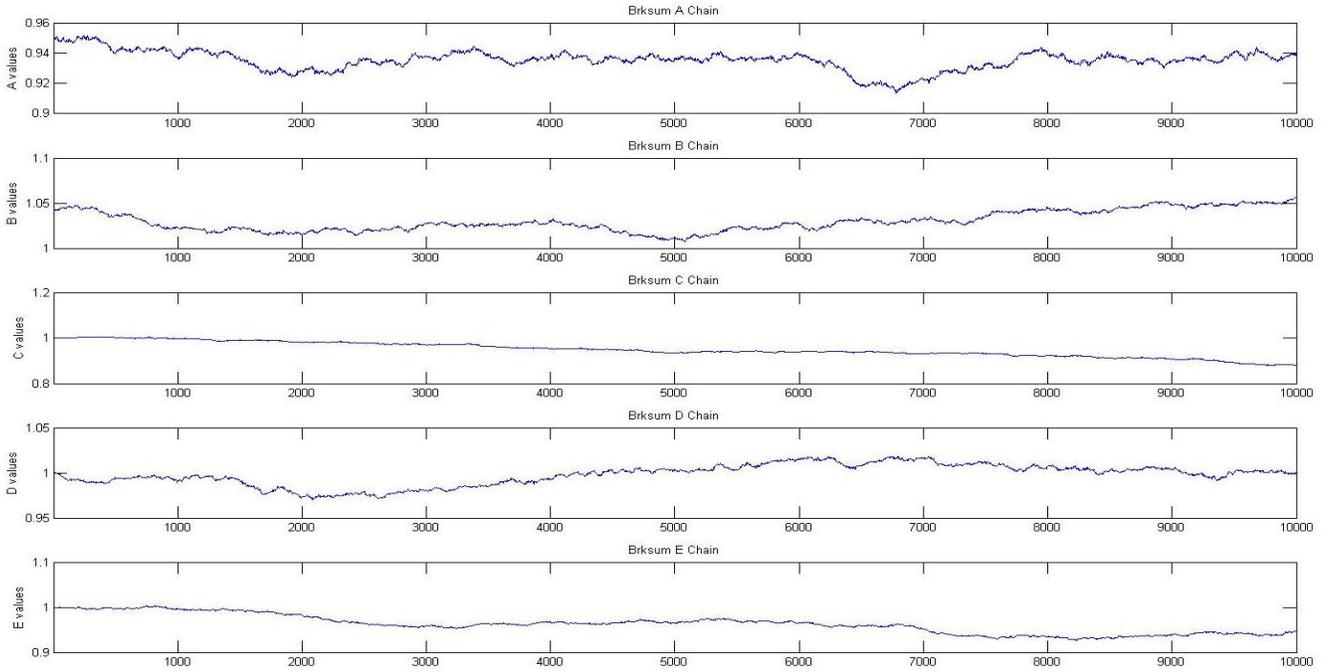
Figures 26 & 27: The figure to the left is the y and z pdfs for the brksumandholo model. In the top left of this figure is the y pdf, the bottom right is the z pdf, and the bottom left is a contour plot of the posterior probability over the y and z paramerers. The figure below displays the y,z and L (posterior probability) chains. The values of y and z which maximized the posterior probability were as follows: y=-18.0011 and z=1.3208.
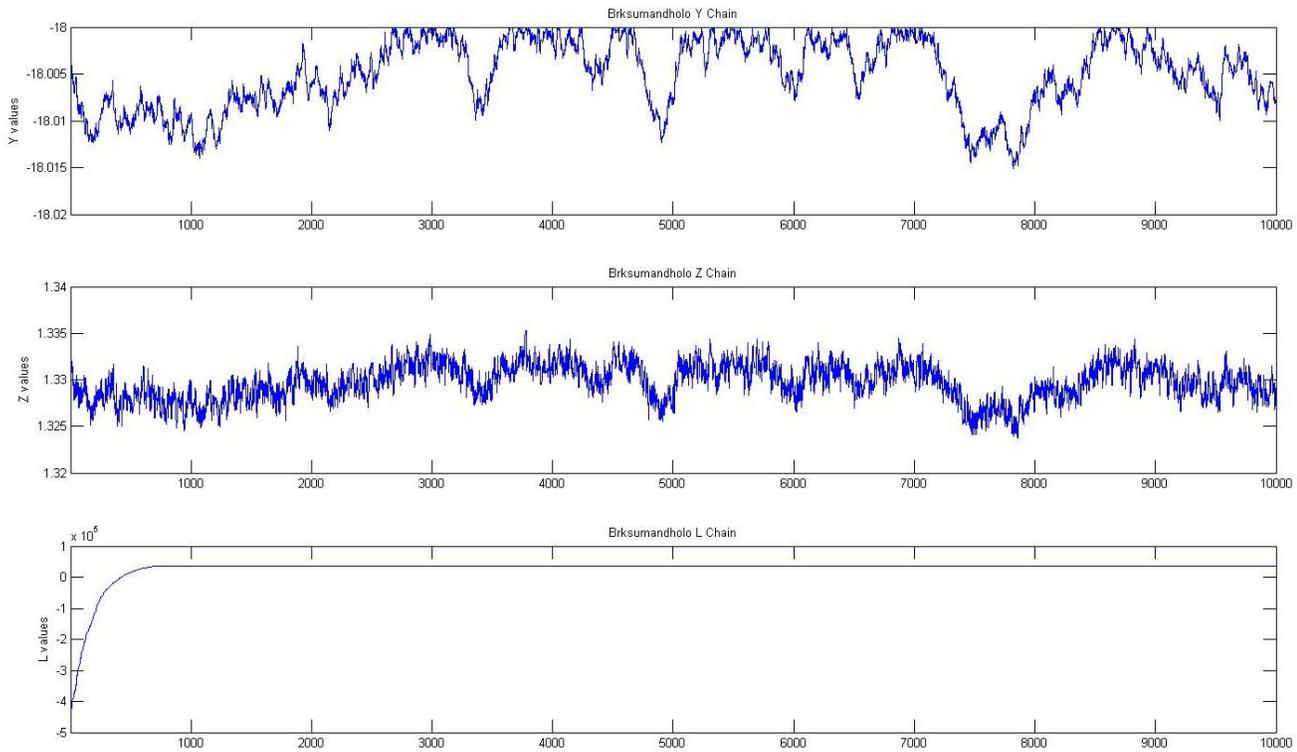
Figures 28 & 29 : These figures show the results of the MCMC for the model known as brksumandholo. In the above graph the noise projections have been multiplied by their corresponding coefficients determined by the MCMC to maximize the posterior probability. In the figure below the noise curves have not been multiplied by

their coefficients. Also in each figure the observed strain is the darker blue curve and the best fit which the model yielded is the red curve.

Model using a sum value that is broken down into dark,shot,srfp,lan bsar, and lan mic-vis noise, as well as a power law variable and a holographic noise component (reffered to as brksumand holo model). Best fit, using gaussian distribution Values of parameters: y = -18.0011, z = 1.3308, a = 0.9364, b = 1.0228, c = 0.8114, d = 1.0024, e = 0.9640 Noise curves plotted without their respective coefficients
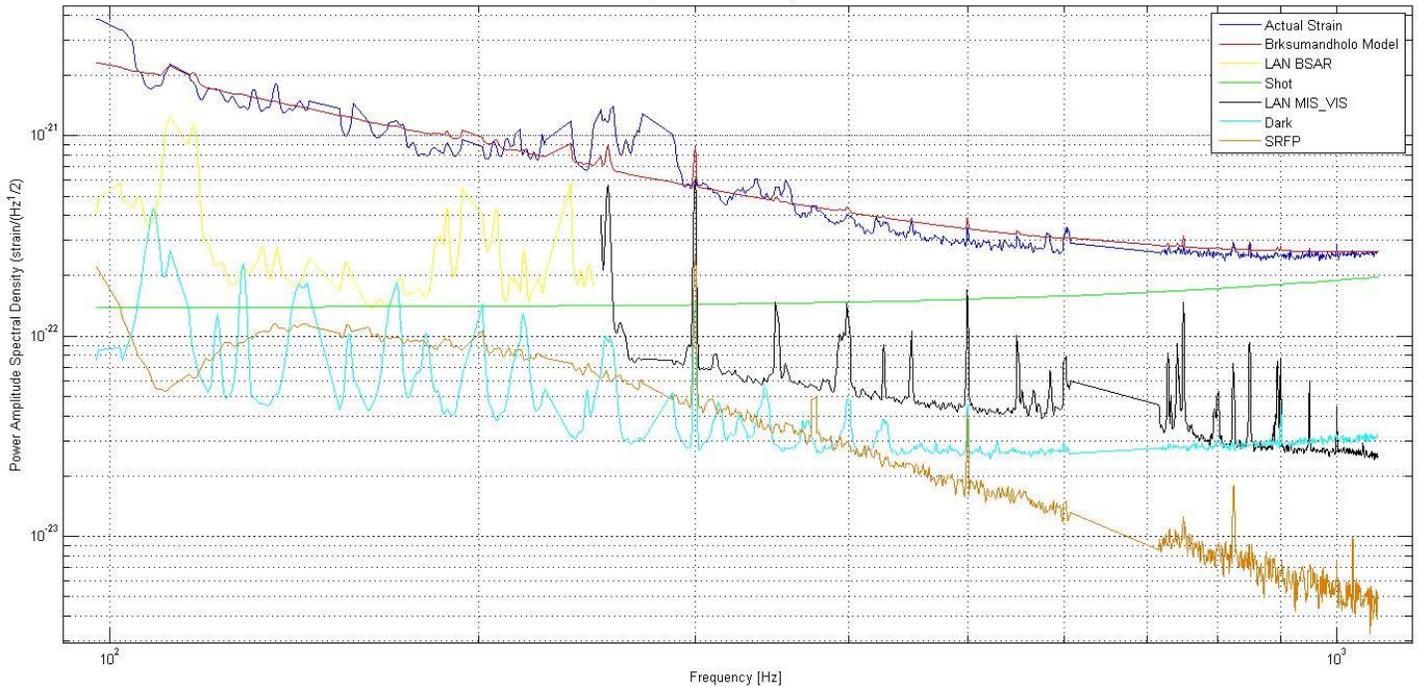
The final model tested by the MCMC method was one which combined the power law model with the broken down SUM and a flat noise parameter, this model was called brksumandflat, and it took the form:

$$brksumandflat$$
$$= \sqrt{\begin{array}{c}(a * lanb)^2 + (b * lanm)^2 + (c * shot)^2 + (d * dark)^2 + (e * srfp)^2 + \\ \left(\dfrac{1 \times 10^Y}{frequency^z}\right)^2 + (flat)^2\end{array}}$$

The procedure for the MCMC was the same as the previous models, a logarithmic Gaussian likelihood was used, and the number of parameters was expanded. The scaling factors chosen to control the size of the Cholesky decomposition were chosen to be the same as the previous models; y:0.0005999, z:0.001555, a, b, c, d, and e: 0.0009, and flat:0.009. The y parameter was again permitted to vary between -18.6 and -18, z from 0.9 to 1.5, flat between 10^{-23} and 10^{-21}, and the error parameters from 0.8 to 1.2. The acceptance ratio during the burn in was 0.53 and that of the chain was 0.54. The following figures show the result of this method.
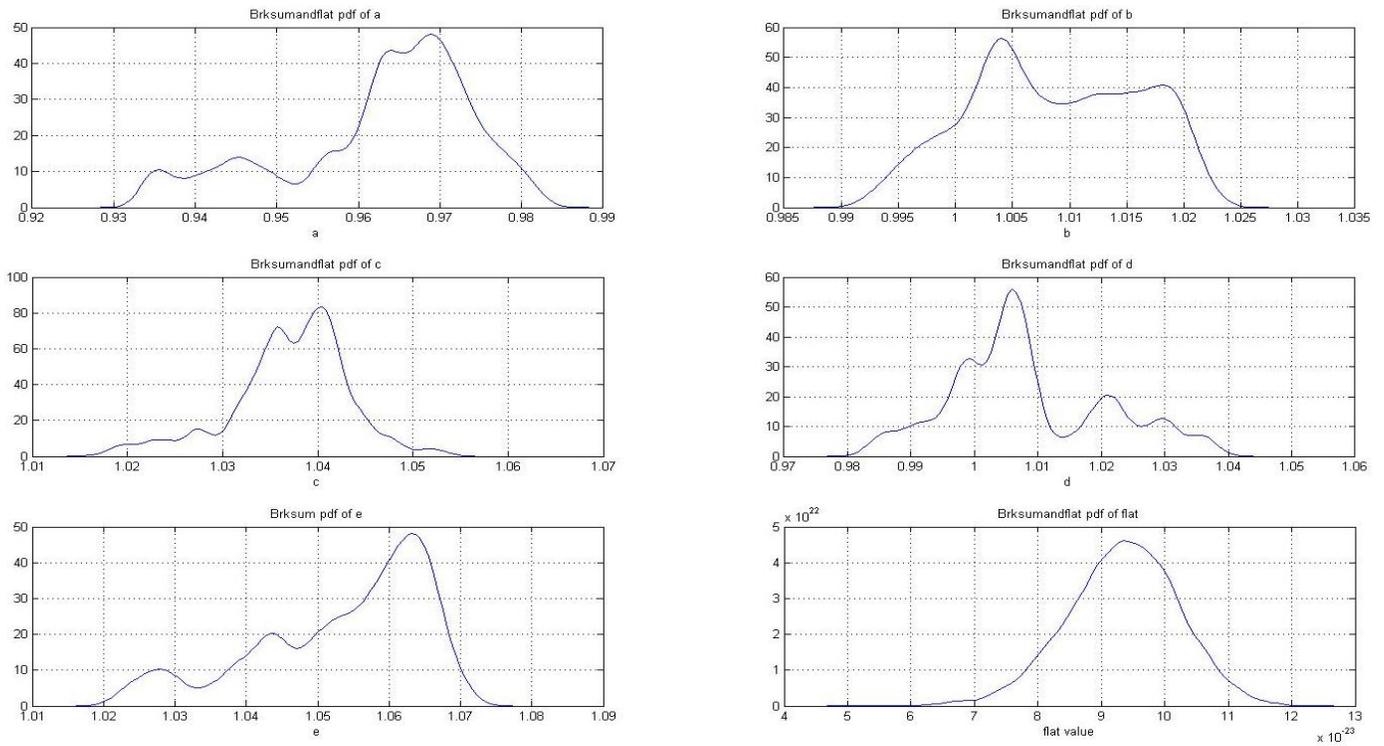
Figures 30 & 31: The above figure shows the pdfs of a, b, c, d, e, and flat, and the figure below shows each of their chains, as well as the chain of the posterior probability. The values of each parameter that maximized the posterior probability were as follows: a=0.9689, b=1.0041, c=1.0405, d=1.0060, e=1.0631, and flat=9.3480×10^{-23}. This value of flat is 50.5% smaller than Hogan's proposed value of 1.85×10^{-22}.
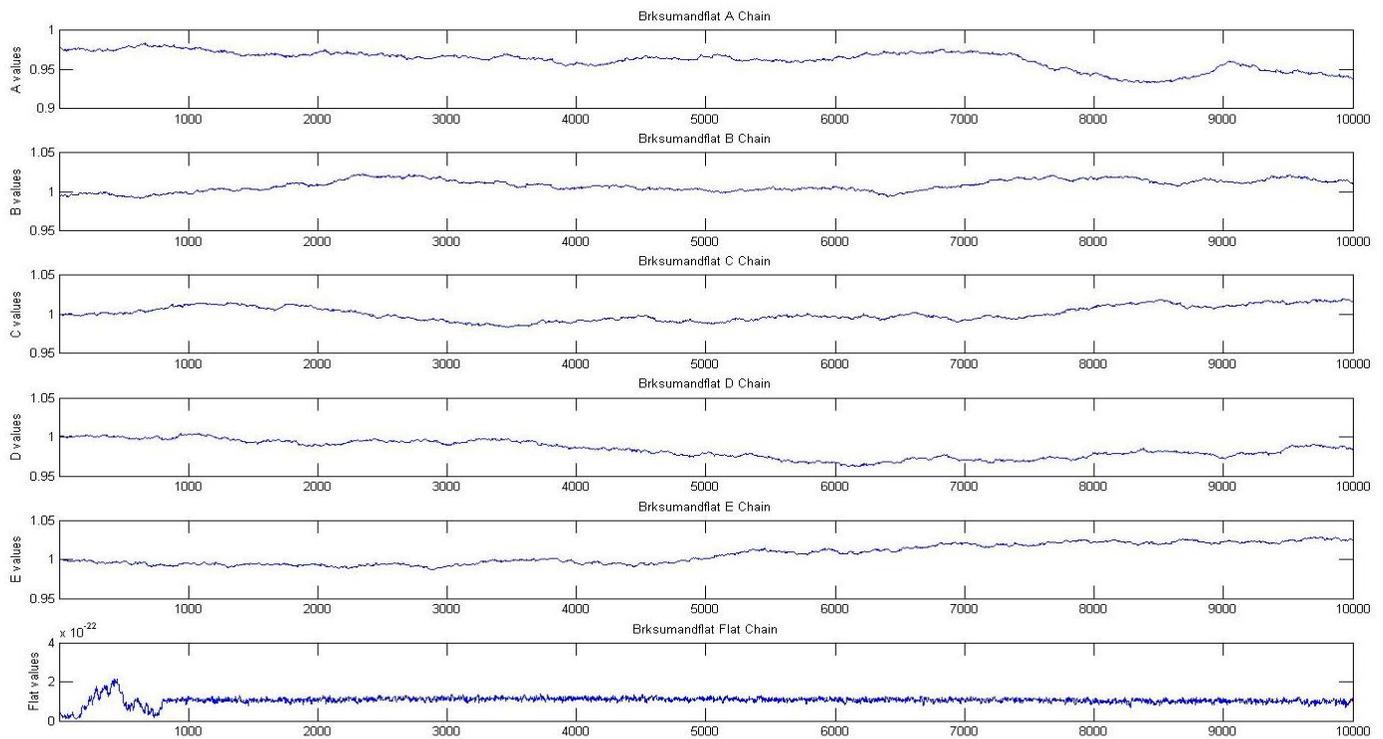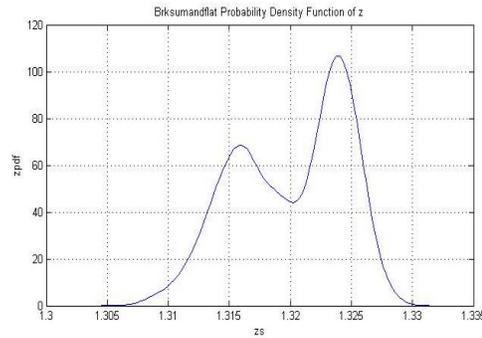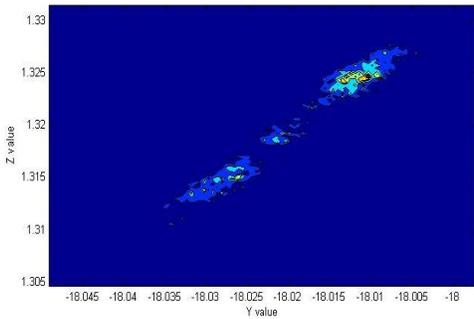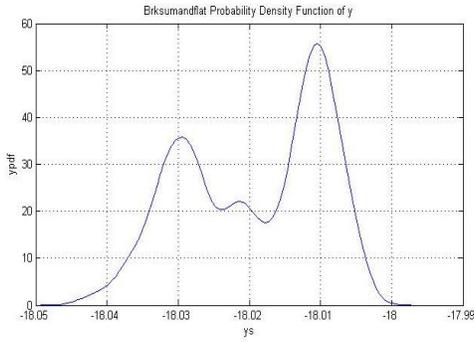
Figures 32 & 33: The figure to the left is the y and z pdfs for the brksumandflat model. In the top left of this figure is the y pdf, the bottom right is the z pdf, and the bottom left is a contour plot of the posterior probability over the y and z paramerers. The figure below displays the y,z and L (posterior probability) chains. The

values of y and z which maximize the posterior probabiliy are as follows: y=-18.0105 and z=1.3041
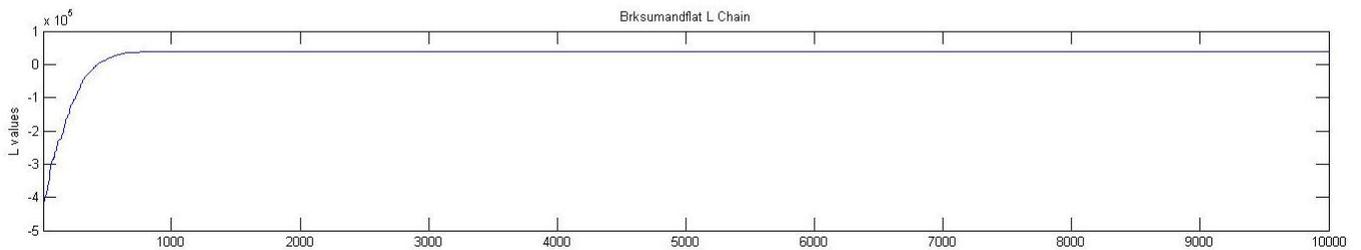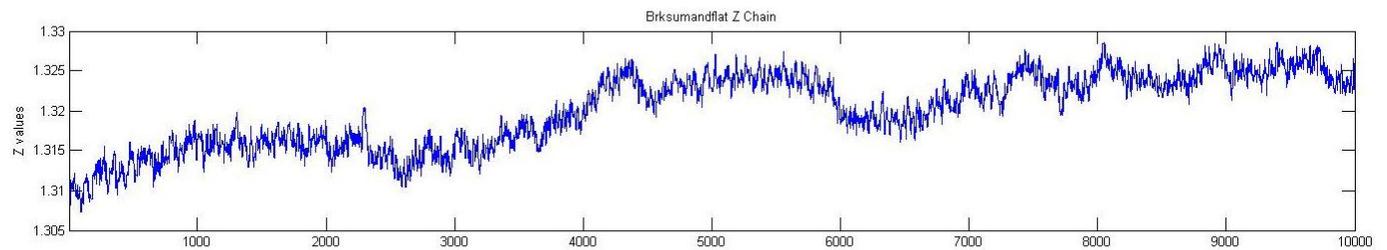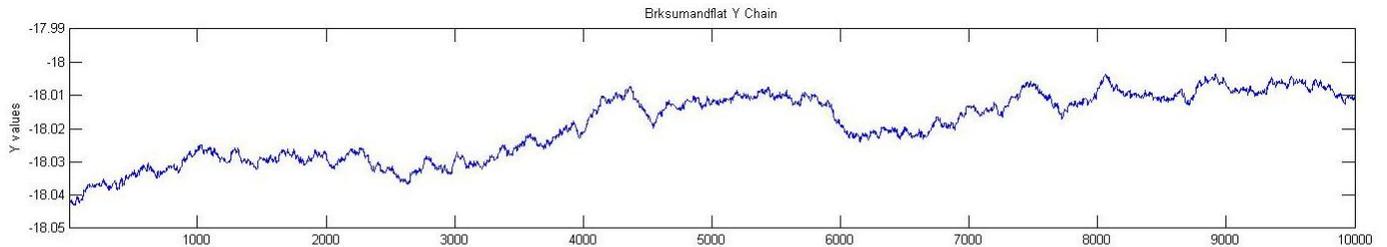
Model using a sum value that is broken down into dark, shot, lan bsar, srfp, and lan mic-vis noise, as well as a power law variable and varying flat noise (reffered to as brksumandflat model).
Best fit, using gaussian distribution Values of parameters: y = -18.0105, z = 1.3238, a = 0.9689, b = 1.0041, c = 1.0405, d = 0.0060, e = 1.0631, flat = 9.3480e-23
Noise curves plotted with their respective coefficients
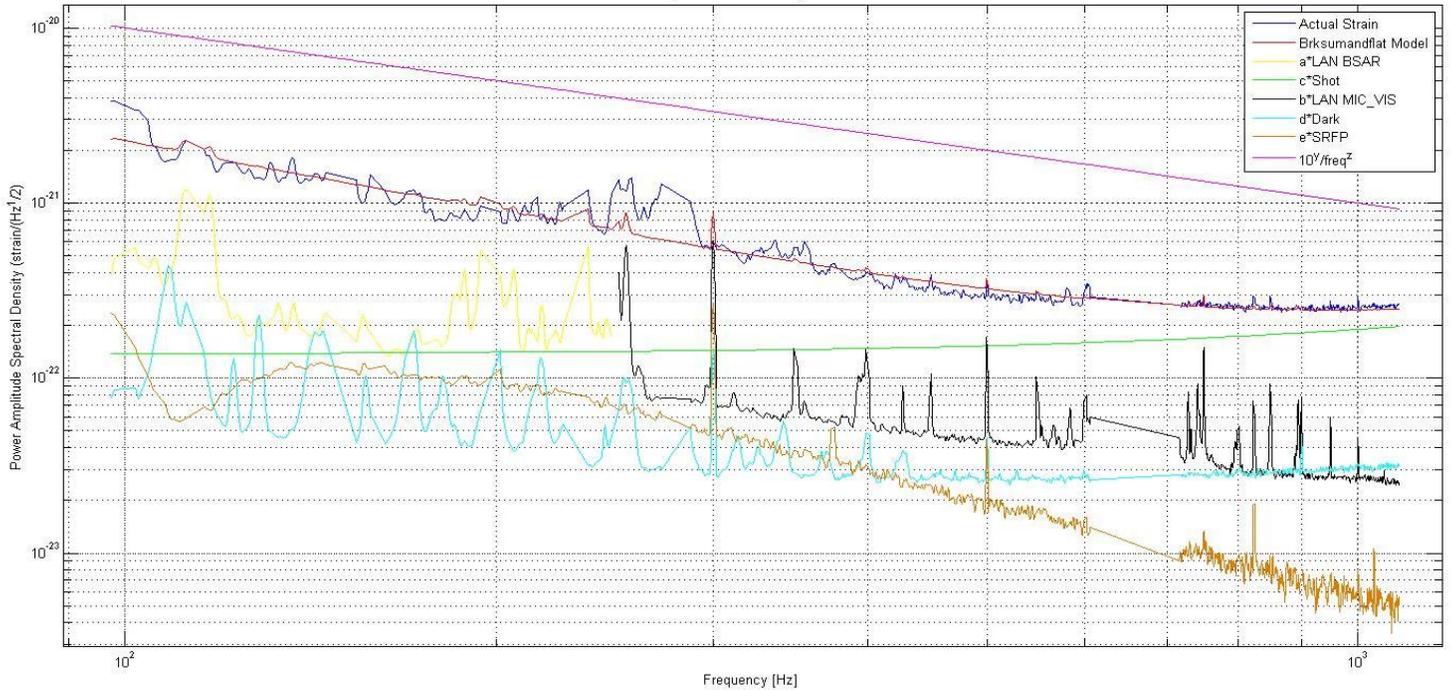
Figures 34 & 35: These figures show the results of the MCMC for the model known as brksumandflat. In the above graph the noise projections have been multiplied by their corresponding coefficients determined by the MCMC to maximize the posterior probability. In the figure below the noise curves have not been multiplied by their coefficients. Also in each figure the observed strain is the darker blue curve and the best fit which the model yielded is the red curve.

Model using a sum value that is broken down into dark, shot, lan bsar, srfp, and lan mic-vis noise, as well as a power law variable and varying flat noise (reffered to as brksumandflat model).
Best fit, using gaussian distribution Values of parameters: y = -18.0105, z = 1.3238, a = 0.9689, b = 1.0041, c = 1.0405, d = 0.0060, e = 1.0631, flat = 9.3480e-23
Noise curves plotted without their respective coefficients

30

3 Conclusions

Figure 36.

In the above plot the best fits that each model created are graphed, along with the strain observed by GEO600. This image clearly indicates that the model which includes only a sum of currently explained noise projections (brksumnopl model, light blue), yields an insufficient explanation of the strain which GEO600 is currently measuring. Also clear from this image is that a model including a sum of currently explained noise projections along with a flat noise component (brksumflatnopl, magenta) is also insufficient for explaining the current strain observed at GEO600. Also clear from the graph is that the other three models tested are very similar in their fits to the actual strain. Given that there is very little obvious difference between the results of the brksum, brksumandflat, and brksumandholo models, more analysis would need to be done to distinguish which of these three models would be most significant.

The methods presented in the previous section provide a starting point for this sort of analysis. These results could also be used in the construction of odds ratios, which could be used to determine the relative validities of two hypotheses compared to one another. Graham Woan gave a brief outline of a means of creating an odds ratio which would give a statement of the likelihood that the GEO noise model should contain a holographic component [7], this is one example where this sort of analysis would be both simple and useful. The methods and results gathered in this report could easily be used towards these sorts of means.

References

[1] D. Sivia and J. Skilling. *Data Analysis: A Bayesian Tutorial.* Oxford University Press Inc., New York. 2006.

[2] Smith, J. R. 2006. Formulation of Instrument Noise Analysis Techniques and Their Use in the Commissioning of the Gravitational Wave Observatory GEO600. PhD Thesis, Gottfried Wilhelm Leibniz Universität Hannover, Hannover.

[3] H. Lück et al. 2010. The Upgrade of GEO 600. arXiv:1004.0339v1 [gr-qc].

[4] Chown, M. 2009, January 15. Our World May Be a Giant Hologram. *New Scientist, issue 2691*.[Retrieved]6 June 2012, http://www.newscientist.com/article/mg20126911.300-our-world-may-be-a-giant-hologram.html?full=true&print=true.

[5] Hogan, C. J. 2012. Interferometers as Probes of Planckian Quantum Geometry. arXiv:1002.4880v27.

[6] Hogan. C. 2009. Holographic Noise. http://elmer.tapir.caltech.edu/cajagwr/pdf/hogan.pdf.

[7] Woan. G. 2009. Do We See Noise in GEO Data?. (publishing information unavailable).

[8] C. R Röver et al. 2009. Bayesian reconstruction of gravitational wave burst signals from simulations of rotating stellar core collapse and bounce. arxiv:909.1093v3 [gr-qc].