Investigation of Possible Biases in Tau Neutrino Mass Limits

Kyle Armour
Departments of Physics and Mathematics,
University of California, San Diego, La Jolla, CA 92093
(Dated: August 8, 2003)

We study the effectiveness of the maximum likelihood estimator when applied to the estimation of ν_τ mass limits. Monte Carlo methods are used to generate one-dimensional distributions whose features are similar to those of real data sets. These features include endpoints that depend on the parameter being estimated, sharply sloped cutoffs and the presence of background. Maximum likelihood techniques are used to estimate parameters in order to investigate the meaning of limits and the presence of bias in both high and low statistics cases. We demonstrate that this application of the maximum likelihood estimator is significantly biased and that care must be taken when determining limits on the tau neutrino mass.

I. INTRODUCTION

The accurate determination of the mass of the τ neutrino is a very important problem in particle physics, and has major implications in cosmology. Since the τ neutrino is weakly interacting, it cannot be directly observed and its properties must instead be inferred from τ decays in which it is produced. The most common way to set a limit on the ν_τ mass is to look at the endpoint of the decay of a τ to many pions and a neutrino[1]. In e+e- experiments, τ+τ- pairs are produced and the invariant mass-energies of the pions from one τ decay are measured. The endpoint and the shape of the one-dimensional mass distribution of the pions depend on the mass of the τ neutrino. In order to obtain a value for the mass from this distribution statistical methods must be applied, and the method that is universally used is that of the *maximum likelihood* (ML) estimator. Although, like any estimator, the ability of ML to produce accurate estimations depends on the particular problem to which it is applied, its effectiveness in the case of the ν_τ mass has been widely assumed, with little verification.

The principle of ML is to estimate the value of a parameter as that which makes the observed distribution of data the most probable. For a data sample $\{x_1, ..., x_N|a\}$, the value of the parameter a is found such that the likelihood:

$$\mathcal{L}(x_1, ..., x_N|a) = P(x_1|a)P(x_2|a)...P(x_N|a) = \Pi P(x_i|a) \tag{1}$$

is a maximum, where $P(x|a)$ is the distribution of the observed quantity x for the parameter a. In practice, it is easier to maximize the logarithm of the likelihood, which is equal to the sum of the logarithms of the probabilities.

When an estimator is applied to a data sample of size N, it will produce an estimate of the quantity being measured. This estimate will differ from the true value because of statistical fluctuations in the sample. However, if an estimator is *consistent* its estimations will tend to the true value as N increases to infinity. For a finite N it is unlikely that the estimate for a particular data sample will have the same value as the true value. If the estimator is *unbiased*, though, its estimates from various samples will average to this value. If the estimator is *biased* then its expectation value is not equal to the true value, and in order

to obtain an accurate estimate it is of crucial importance that the bias is understood and accounted for. Since an estimate will vary depending on the particular data sample used, it is important that the spread of these possible values be small. An estimator with a small spread, or variance, is said to be *efficient*. If an estimator has a small variance, then it is likely that an estimate is close to the mean value. The best estimator would be one that is consistent, unbiased and efficient. Unfortunately, for most applications, there are none that display all three of these characteristics and choices must be made between those that have varying degrees of bias and efficiency.

In some cases ML has significant advantages over other estimators. For large sample sizes, its estimated values have a distribution that is unbiased and normally distributed about the true value, and its variance is as small as possible, saturating what is known as the *minimum variance bound*[2][1]. Also, there is no loss of information due to binning. When estimating using large sample sizes, ML tends to be consistent, unbiased and efficient, making it a very effective estimator. It is important to note that there is no set number of points that constitute a large sample size, and this condition must be evaluated on a case by case basis. There are disadvantages to the ML method as well. For small samples, it is usually biased and has a variance that is somewhat greater than the minimum variance bound. Even though the sample sizes associated with the estimation of the τ neutrino mass are relatively small, these characteristics of ML are often ignored.

Another problem associated with the estimation of the ν_τ mass is that the endpoints of the probability distributions depend on the mass itself, while the standard proofs for the minimum variance bound assume endpoints that are independent of the parameter being estimated [2]. Thus, the concept of efficiency for ML, even for large N, has to be reconsidered when applied to distributions with this feature. Also, when these features are present there is a possibility that even for large sample sizes the ML estimator will be biased and produced a skewed distribution of estimated values.

Although there could be significant problems associated with the use of the ML method in determining the ν_τ mass limits, they have been ignored and unaccounted for. Most neutrino analyses do not acknowledge the introduction of bias and assume confidence in the validity of raw upper limits. In view of these possible problems, it is the purpose of this study to examine what inaccuracies might arise from the use of the maximum likelihood estimator in setting limits on the ν_τ mass.

II. METHOD

Monte Carlo methods are used to generate one-dimensional distributions whose features are similar to those of real distributions, without the effects of detector smearing. These features include sharp cutoffs, background and endpoints that depend on the parameter being estimated. The distributions are created by the acceptance and rejection of randomly generated points according to defined functions. The number of points in the signal region of the distribution (above the background) is kept constant for consistency.

We investigate the behavior of ML when used to estimate various parameters of these generated distributions. Studied first is the estimation of the mean of a Gaussian distribution. This serves as an example of a well-behaved ML estimator. We then study the

[1] These features do not hold in some cases, as will soon be shown.

estimation of the endpoints of: a) a tophat distribution, b) sloped tophat distribution, and c) a distribution that is similar to that of a real neutrino mass distribution. The tophat and sloped tophat distributions can be thought of as first and second order approximations, respectively, of real one-dimensional τ decay product mass distributions. Finally, we compare the various methods used to calculate limits on the ν_τ mass.

III. DATA AND ANALYSIS

A. Gaussian Distribution

Before the effects introduced by the use of the maximum likelihood estimator on tau neutrino mass data can be investigated, it is important to have an understanding of its use under optimal conditions. A Gaussian distribution is nicely behaved (no sharp edges) and its endpoints do not depend on the parameter being estimated.

Gaussian distributions of various sample sizes are generated at a given mean and fixed width and ML is used to estimate the mean of each sample. The general equation for the Gaussian used is:

$$G(x|\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-(x-\mu)^2/2\sigma^2\right), \tag{2}$$

where σ is set to one, and μ is zero.

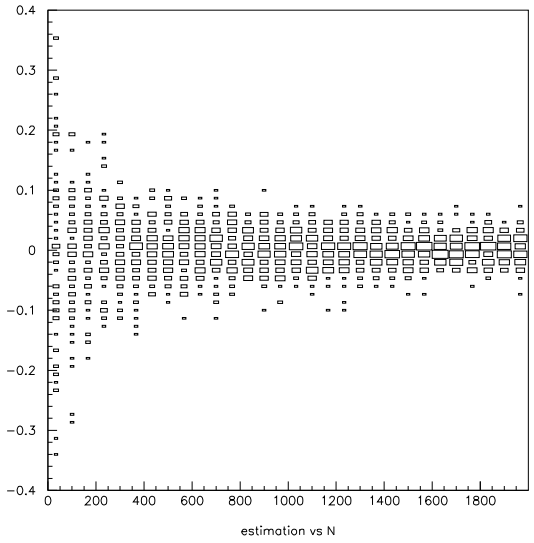

FIG. 1: Estimated mean values as a function of the number of points in a Gaussian distribution centered at zero.

Fig 1 shows the estimated values of the mean as a function of the number of data points, N, in a distribution centered at 0. While the spread of the distribution is initially quite large, it quickly becomes small, approaching the true value as N increases towards infinity. This is a clear indication that ML as applied to a Gaussian distribution is consistent. Also apparent is that there are approximately equal amounts of estimated values on either side of the true value, indicating that the estimator may be unbiased.

To get a more quantitative understanding of these trends, many Monte Carlo experiments are done with the same number of data points. Fig 2 is a histogram of estimated values

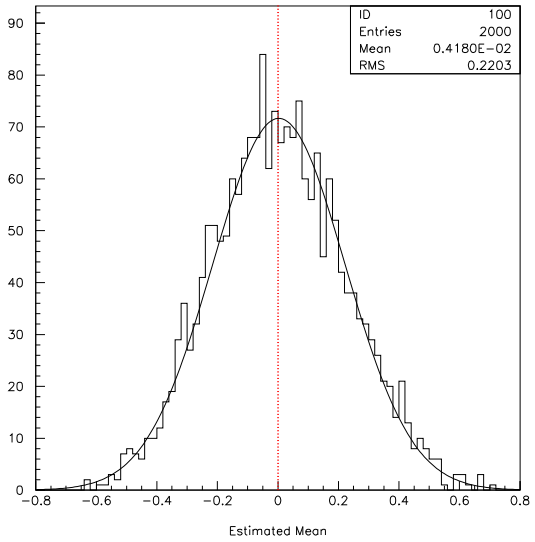

FIG. 2: Histogram of estimated mean values for many experiments done for the Gassian distribution (N=20). The dotted line shows the true value.

when N=20. This distribution is Gaussian in shape and its mean is slightly higher than the true value. To determine if this difference between mean and true value constitutes a bias, we have to look at what is called the *standard error on the mean*: $\frac{\sigma}{\sqrt{n}}$, where σ is the RMS deviation and n is the number of experiments in the distribution [2]. The standard error on the mean is defined for any distribution, not just Gaussian. In this case, the standard error on the mean is a larger value than the difference between the mean and the actual value of 0. If the mean is less than the value of the standard error away from the true value, bias is not observed, and the ML estimator can be said to be unbiased. Similar histograms with N=200 and N=2000 (not shown) are also unbiased, and verify that the ML estimator is unbiased and consistent when applied to this distribution.

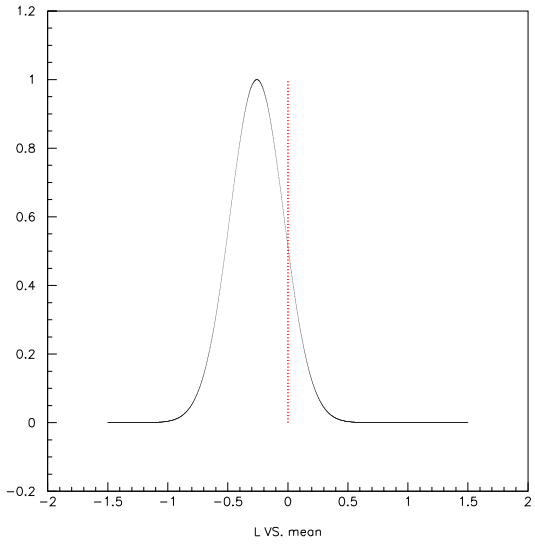

FIG. 3: Typical likelihood function for experiments done with the Gaussian distribution (N=20). The dotted line shows the true mean value.

The typical likelihood for a distribution of size N=20 is shown in Fig 3. The peak of this curve corresponds to the mean value that would make the observed distribution of data the most probable distribution. The curve is of a nicely behaved Gaussian shape, making it straightforward to interpret in a Bayesian approach (see section 4).

The ML estimator is known to be consistent and unbiased in the estimation of the mean of a Gaussian distribution. Also note that its variance is equal to the minimum variance bound in this case, making it efficient. These properties make the ML estimator a very good choice when estimating with functions such as this one, whose endpoints do not depend on the parameter being estimated. In the next sections we will consider the ML estimator when applied in a similar way, but to distributions that do not have such nice features, and compare those results to these.

B. Tophat Distribution

The first of the tophat distributions is generated in order to test the effects of estimating the endpoint of a distribution with a sharp cutoff. The general equation for the tophat distribution is:

$$f(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 + background \text{ if } 0 < x < ep \\ background \text{ if } x > ep. \end{cases} \tag{3}$$

In this case, *ep* (the endpoint of the distribution) is 0.7 and the background is 0. Care must be taken when applying ML without the presence of background. The value of the probability for certain points in the distribution is at times zero when scanning over possible parameter values, which leads to infinities when taking the log likelihood. In practice we get around these infinities by defining log(0) = -999.

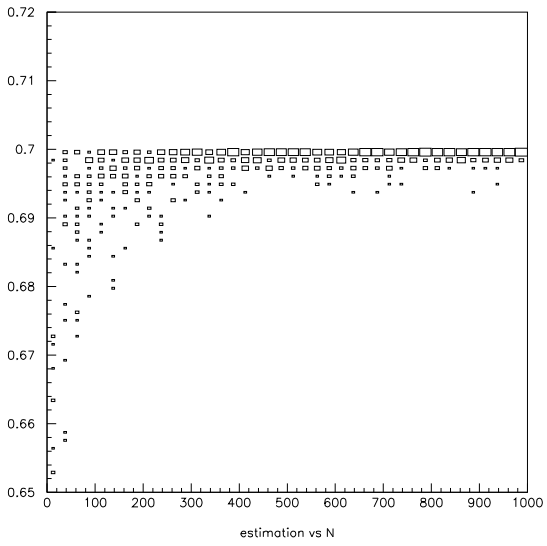

FIG. 4: Estimated endpoint values as a function of the number of points in the tophat distribution with a true endpoint value of 0.7.

Fig 4 shows the estimated values of the endpoint as a function of N for this distribution. Again, as N increases to infinity, the estimated values approach the true value, indicating

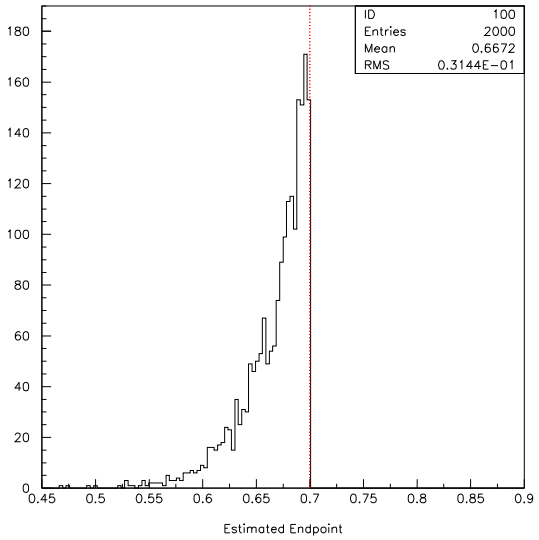

FIG. 5: Histogram of estimated endpoint values for many experiments done for the tophat distribution without background (N=20). The dotted line shows the true endpoint value.

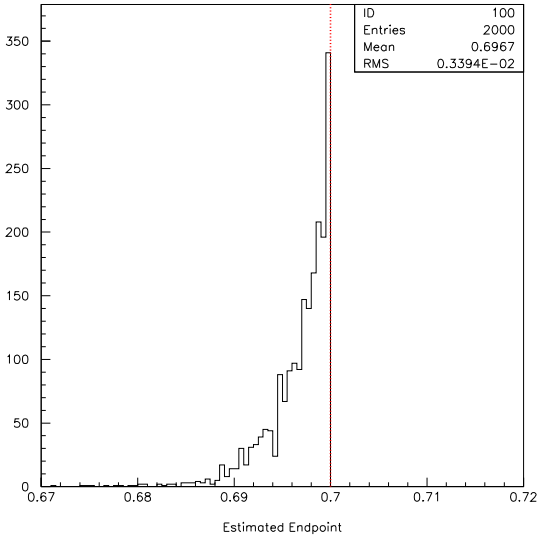

FIG. 6: Histogram of estimated endpoint values for many experiments done for the tophat distribution without background (N=200). The dotted line indicates the true endpoint value.

that the ML estimator is consistent. What is strikingly different from the equivalent figure for the Gaussian distribution is that all of the estimated values are below the true endpoint value, indicating that there is a bias. To determine the extent of this bias, many distributions are generated with a constant sample size, as before. The histogram of ML estimations for distributions with N=20 is shown in Fig 5. Note that this histogram is not of a nice Gaussian shape, and that there is a bias towards estimations that are lower than the true value. The standard error on the mean is less than the difference between the mean and the true value, which means that the observed bias is significant. The estimation histogram for N=200, shown in Fig 6, indicates that the same trends occur for higher statistics as well. The bias in this case is smaller than the bias for N=20, suggesting again that the estimator is

consistent, but is still significant.

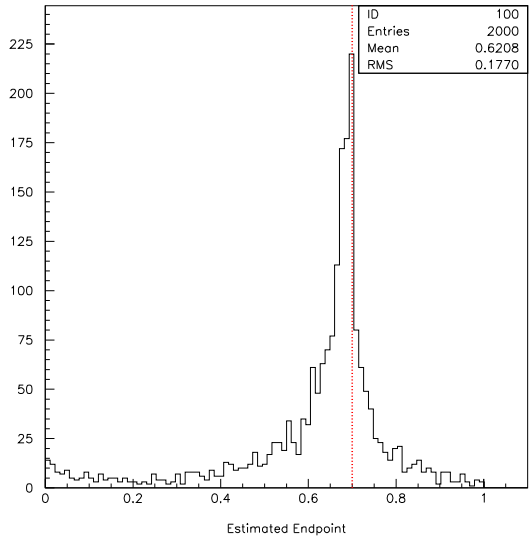

FIG. 7: Histogram of estimated endpoint values for many experiments done for the tophat distribution with background (N=20). The dotted line shows the true endpoint value.

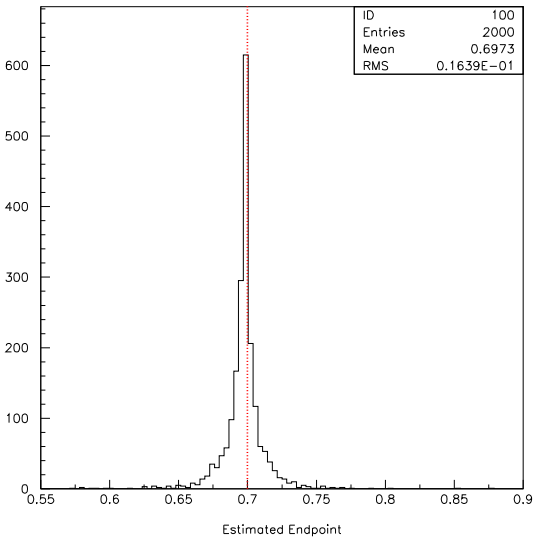

FIG. 8: Histogram of estimated endpoint values for many experiments done for the tophat distribution with background (N=200). The dotted line shows the true endpoint value.

In order to understand its effects, background is added to the tophat distribution (background = .4). The addition of background alleviates the problems associated with taking the logarithm of zero, but has other effects as well. The histogram of estimated endpoints for N=20, Fig 7, shows that an increase in background results in a more normal distribution, estimating points both above and below the true value. While the shape of the distribution is more normal with background, there is a greater variance, and comparing the difference between mean and true value to the standard error on the mean reveals that the ML estimations are even more biased. Even for estimations done with a larger sample size, as shown

in Fig 8 (N=200), these trends continue.

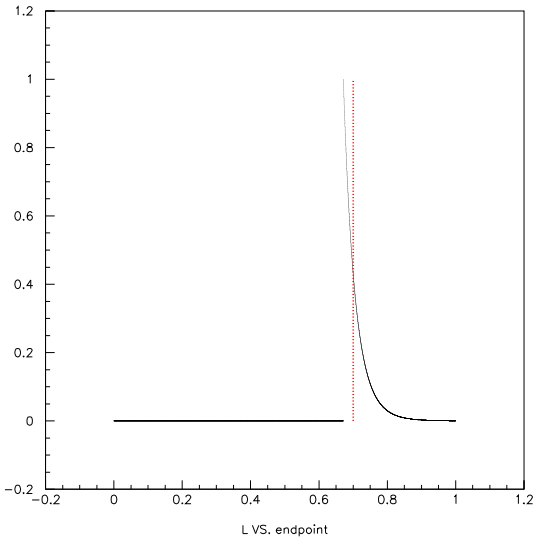

FIG. 9: A typical likelihood curve for experiments done for the tophat distribution without background (N=20). The dotted line shows the true endpoint value.

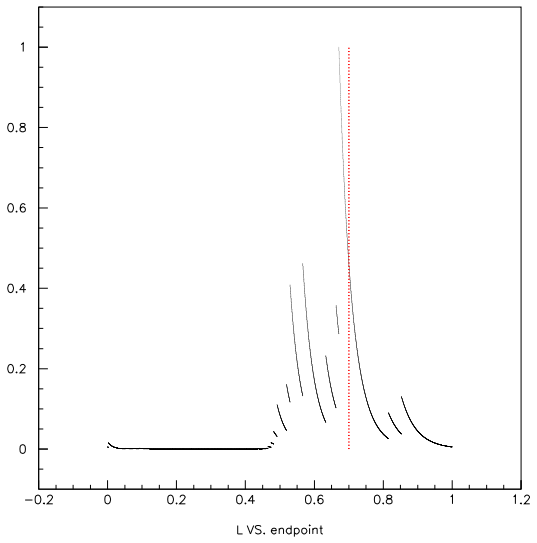

FIG. 10: A typical likelihood curve for experiments done for the tophat distribution with background (N=20). The dotted line shows the true endpoint value.

A typical likelihood curve for a tophat distribution (N=20) without background is shown in Fig 9. It is evident that this curve is not of the nice Gaussian shape that it was for the Gaussian distribution. These strange, asymmetric likelihood shapes are due to the lack of background, which leads to infinities when the logarithm of 0 is taken during the application of ML, as discussed above. Looking at a likelihood curve in the presence of background, as shown in Fig 10 (N=20), we see that it too has discontinuities, but is beginning to look more symmetric.

We have observed a bias in the application of the ML estimator to the tophat distribution in both low and high statistics cases. This is an ideal example of how indefinite the concept

of a large sample size really is. In this case, only an infinite number of points in the observed distribution would produce unbiased ML estimations of the endpoint. The addition of background, while making the ML method mathematically more reasonable, leads to estimations that are more biased and more likely to be far away from the true value. These results show that when the ML estimator is applied to a poorly shaped distribution whose endpoints depend on the parameter being estimated, there is the possibility of significant bias.

C. Sloped Tophat Distribution

The next distribution considered is that of the sloped tophat, defined by:

$$f(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 + background \text{ if } 0 < x < 0.3 \\ 1/(ep - 0.3) * x + background + ep/(ep - 0.3) \text{ if } 0.3 < x < ep \\ background \text{ if } ep < x < 1 \\ 0 \text{ if } x > 1 \end{cases} \tag{4}$$

By changing the slope in the generated distribution (ep=0.7, with no background), we can get some information about how its shape near the endpoint affects ML estimation. With the tophat function in the previous section, we saw that there is significant bias associated with estimating the endpoint of a distribution with a sharp cutoff. Now we explore the estimation of the endpoint of a distribution that is more nicely shaped, and compare the results.

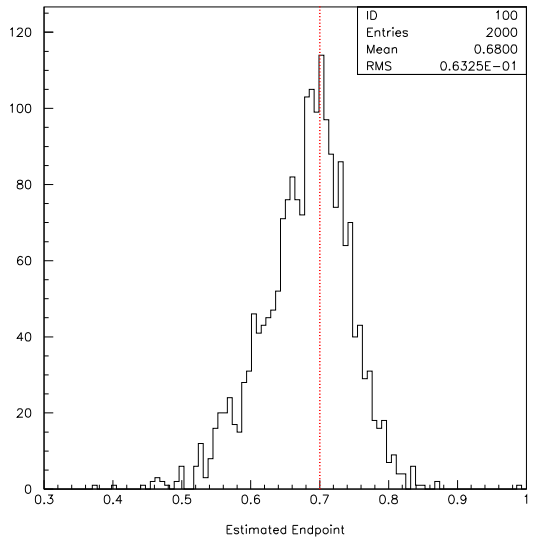

FIG. 11: Histogram of estimated endpoint values for many experiments done for the sloped tophat distribution without background (N=20). The dotted line indicates the true endpoint value.

The histogram of estimated endpoint values for N=20, shown in Fig 11, is almost free of the asymmetry that dominated the histograms of the previous section. Its shape is more similar to those seen when estimating the mean of the Gaussian distribution. The sloped

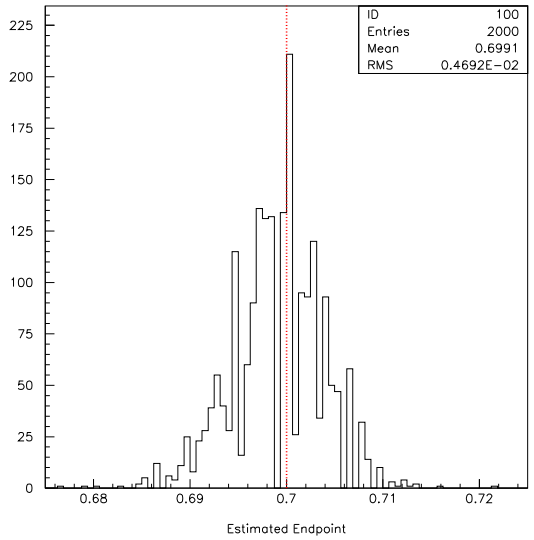

FIG. 12: Histogram of estimated endpoint values for many experiments done for the sloped tophat distribution without background (N=2000). The dotted line shows the true endpoint value.

tophat histogram of estimated endpoints is still biased, however, but to a lesser extent than that of the tophat. Therefore, it is mainly the sharp cutoff near the endpoint of an observed distribution that leads to estimation histograms that are poorly shaped. And bias, while affected by this sharp cutoff to some extent, seems to stem from having endpoints that are dependent on the parameter being estimated. Even for the large sample size of N=2000, as seen in Fig 12, the ML estimator remains biased. The average of the estimations does approach the true value as N increases, though, once again showing that the ML estimator is consistent.

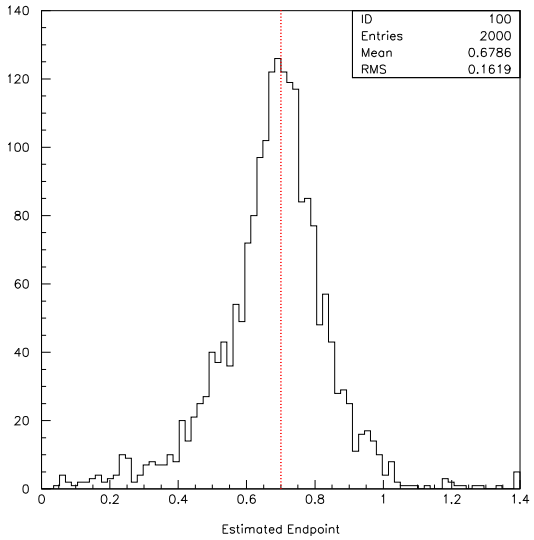

FIG. 13: Histogram of estimated endpoint values for many experiments done for the sloped tophat distribution with background (n=20).

The presence of background (background=.4) in the sloped tophat distribution shows the same effects as those observed in the case of the tophat with background: an increase in

bias and spreading in estimated values. This is shown for N=20 in Fig 13, and applies to larger sample sizes as well. Therefore, the ML method estimates most accurately when the observed data distribution is nicely shaped and free of background.

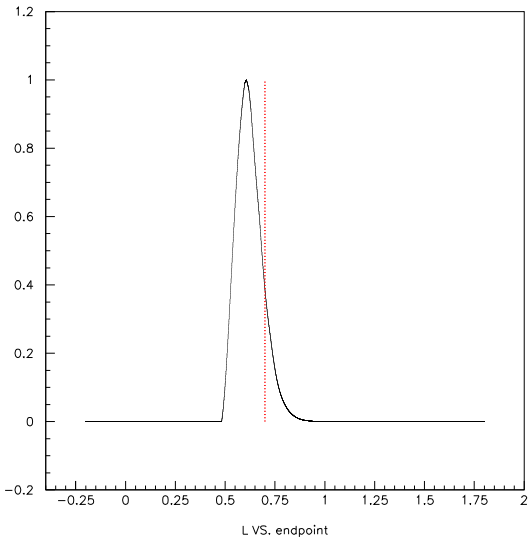

FIG. 14: A characteristic likelihood curve for the estimation of the endpoint of the sloped tophat distribution without background (N=20). The dotted line shows the true endpoint value.

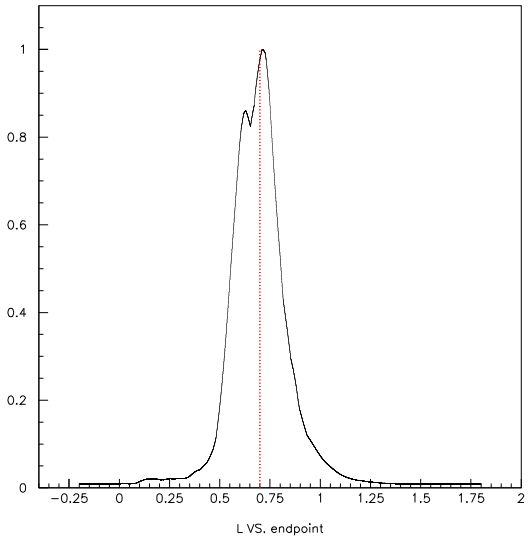

FIG. 15: A typical likelihood curve for the estimation of the endpoint of the sloped tophat distribution with background (N=20). The dotted line shows the true endpoint value.

A typical likelihood curve for the sloped tophat without background is shown in Fig 14. While it no longer has discontinuities, it still displays the asymmetry that is associated with a lack of background in the tophat case. In Fig 15, a characteristic likelihood curve for the sloped tophat with background is shown. Almost Gaussian in shape, it demonstrates once again that the presence of background helps to make the likelihood curve more symmetric.

For the sloped tophat distribution, we find trends that are similar to those of the tophat. The persistence of bias in this case indicates that even for nicely behaved distributions, the

process of estimating the endpoint with ML induces a bias. In the next section we see to what extent these findings hold for distributions that are similar in shape to real neutrino mass distributions.

D. Neutrino Mass Distribution

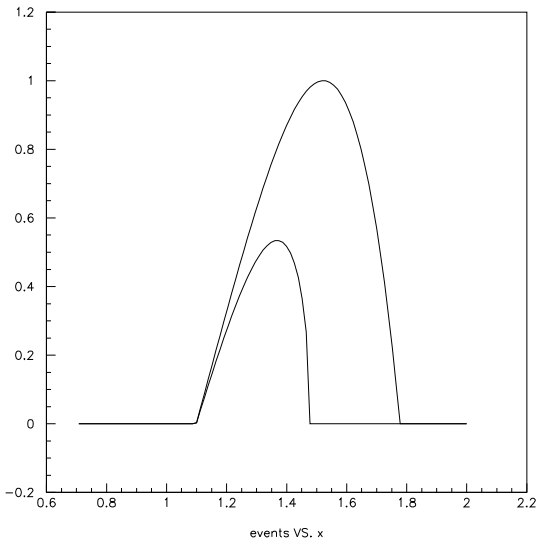

FIG. 16: Example of the shapes of the neutrino mass distributions used in this analysis (based upon real mass distributions). The top curve corresponds to a distribution for a 0MeV neutrino, while the bottom curve corresponds to that of a 300MeV neutrino.

We now study a distribution that is similar to that of real detector data[1]. The shapes of real distributions are quite different from those previously studied, but they have similar features that can affect the ability of the ML estimator to make accurate estimates. These features include sharp cutoffs, background, and endpoints that depend on the ν mass, the parameter to be estimated by ML. Fig 16 shows the shapes of the distributions for 0MeV (top) and 300MeV neutrinos (bottom). The number of observed data points in a real distribution decreases with increasing ν mass due to fewer decays[2]. Therefore, both the shape and the endpoint of these distributions depend on the neutrino mass. Note that in the following we neglect the important effects of detector smearing.

The first mass distribution is generated with a 200MeV input neutrino, without background, and the mass of the ν is estimated using ML. The histogram of estimated mass values for this distribution (N=20), as shown in Fig 17, has an asymmetry similar to those seen in the case of the tophat distribution[3]. This occurs because of the sharp cutoff near the endpoint of the distribution that is common through all real mass distributions. It also shows that the ML estimator is biased towards values that are larger than the true value. The histogram of estimated values for a larger sample size, as shown in Fig 18 (N=200),

[2] In a real experiment, only a fixed number of points are observed.

[3] The neutrino mass distribution is scaled in such a way that an estimation of ".2" for the neutrino mass corresponds to an estimation of 200MeV from a real distribution.

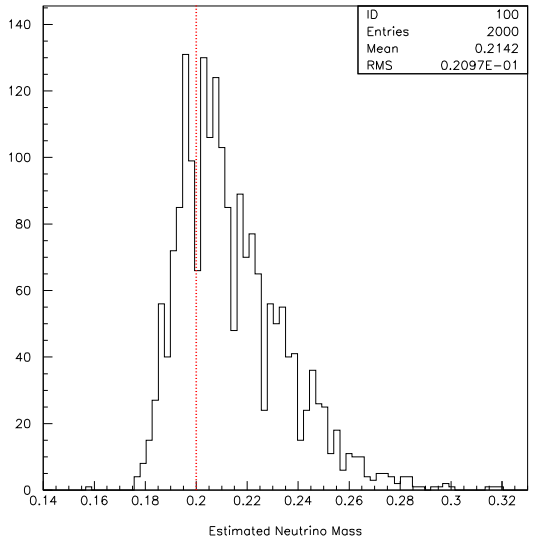

FIG. 17: Histogram of estimated neutrino mass values for many experiments done for the neutrino mass distribution without background (N=20). The dotted line shows the actual mass value.

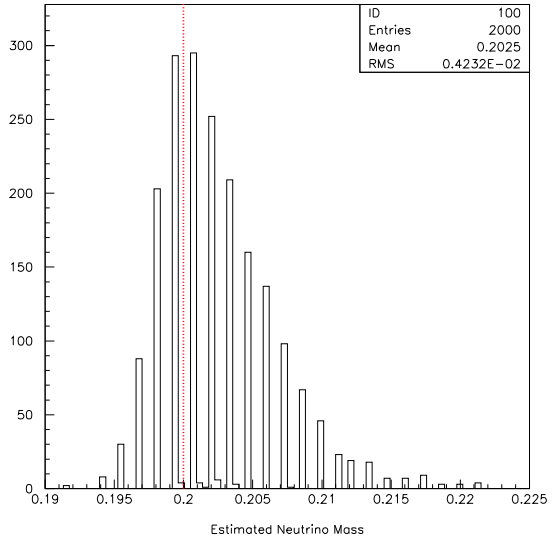

FIG. 18: Histogram of estimated mass values for many experiments done for the neutrino mass distribution without background (N=200). The dotted line corresponds to the true mass value.

shows that the ML estimator is consistent and that bias is observed in sample sizes that are comparable to those of real data. This indicates that real data sets are of small enough size that the ML estimator has the possibility of bias in neutrino mass estimation. It is likely that, like the tophat distribution, the observed distribution of real neutrino mass data would have to have a nearly infinite size for the ML method to be unbiased. Note that typical detector smearing is on the order of 10MeV, consistent with the scale at which we see bias. Thus firm conclusions on the behavior of the ML estimator will require examining smeared distributions.

The addition of background to the neutrino mass distribution actually decreases the bias of the ML estimator in some cases. This effect is not fully understood, and it did not occur

for any other distribution studied here. If the background is large enough, however, the ML estimator behaves exactly as it did in those other cases, producing estimates with a greater bias and variance.

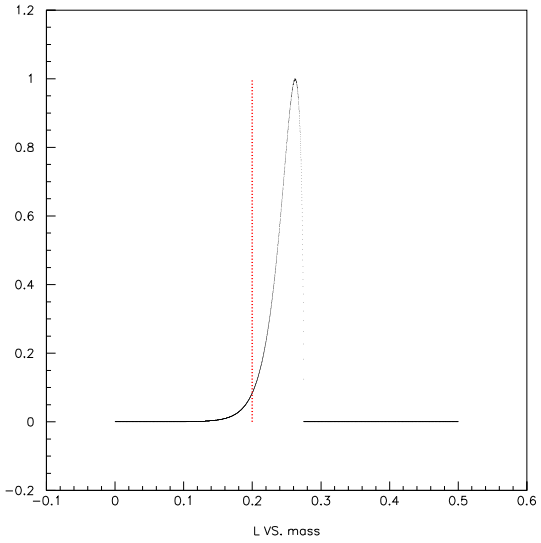

FIG. 19: A typical likelihood curve for experiments done for the neutrino mass distribution without background (N=20). The dotted line indicates the true mass value.

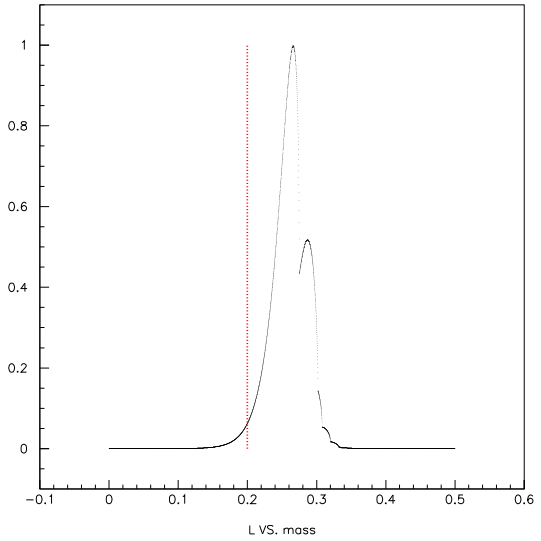

FIG. 20: A characteristic likelihood curve for experiments done for the neutrino mass distribution with background (N=20). The dotted line indicates the true mass value.

A typical likelihood curve for the neutrino mass distribution (N=20), without background, is shown in Fig 19. This curve is much like those likelihood curves seen in the tophat and sloped tophat distributions without background; it is highly asymmetric. With the addition of background to the ν mass distribution, ML produces likelihood curves such as shown in Fig 20. As background is increased, these curves become increasingly symmetric and Gaussian in shape.

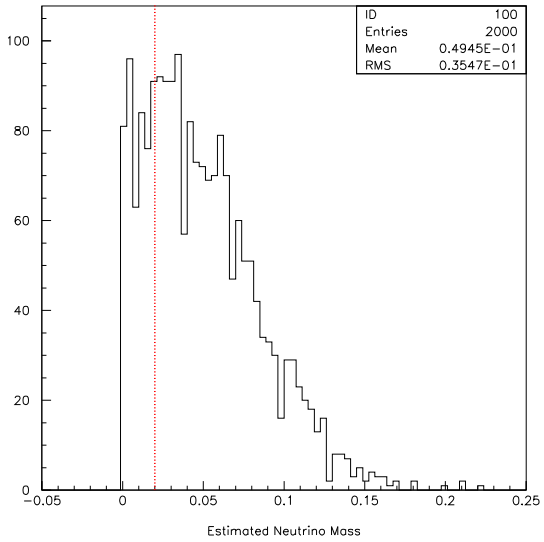

FIG. 21: Histogram of estimated mass values for many experiments done for the neutrino mass distribution generated with a 20MeV input mass. Note the strange pile-up of estimated values at zero. The dotted line shows the true mass value.

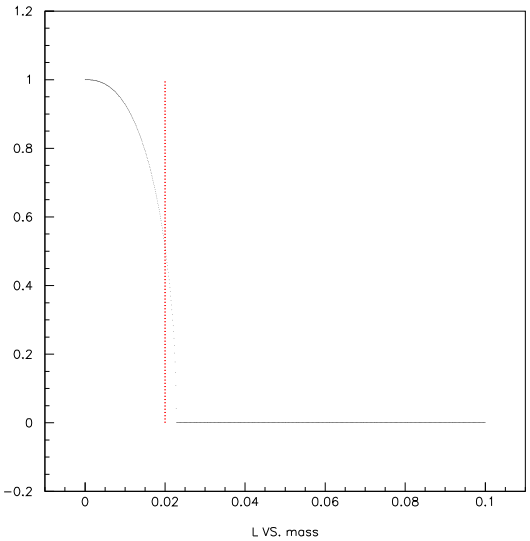

FIG. 22: A typical likelihood curve for those experiments that estimate a 0MeV neutrino for the input mass of the neutrino mass distribution generated with a 20MeV neutrino.

The analysis of the neutrino mass distribution above is done with a 200MeV input neutrino, and its results are similar to those of previously investigated distributions. An interesting effect occurs for a lower neutrino mass, though, as is shown in Fig 21, the histogram of estimated mass values for a 20MeV input neutrino (N=20). The strange pile-up of estimations at 0MeV is due to the inability of the ML estimator to scan below this value (the equations defining the probability distribution used for ML cannot accommodate a neutrino with a negative mass). Likelihood curves that would have peaked at a mass lower than zero due to an unlucky distribution instead report 0MeV as their maximum value. This effect is illustrated in the likelihood curve shown in Fig 22 for a 20MeV input neutrino

(N=20). This effect must be taken into account when estimating the neutrino mass from a real distribution.

IV. LIMITS

A. Definitions of Limits

One possible way to define limits on the ν_τ mass is the Bayesian method. The Bayesian method considers the likelihood curve generated by ML to be a probability distribution for the mass of the neutrino. In order to set an upper limit on the mass, the likelihood function is integrated from 0MeV to the point at which the integral is a certain percent of the total integral. If the curve is integrated to the 90th percentile, this point is reported as the upper limit with a 90% confidence level. This is the conventional definition used in neutrino mass analysis.

Another possible definition of neutrino mass limits is the frequentist limit, which requires the generation of many Monte Carlo experiments done at each possible neutrino mass value. The peaks of the generated likelihood curves are plotted against input parameter values, creating a look-up table of observed versus generated values. If a 90% confidence level on an upper limit is desired, each generated distribution is integrated vertically to the 10th percentile. The measured neutrino value is then found on a curve connecting these points, and its corresponding input neutrino value is reported as the upper limit with a 90% confidence level.

Unlike the Bayesian definition of limits, the frequentist approach does not create the concept of a probability distribution for the mass of the neutrino. It instead states that if the true value is greater than this set limit, then the probability of getting a smaller measurement is less than 10% (or that 90% of the time the upper limit is the largest possible value that could have produced the observed distribution). While any single statement of a limit is either right or wrong, by construction if you take a large number of these statements 90% of them will be valid[2]. The properties of the Bayesian and frequentist definitions when applied to the Gaussian and neutrino mass distributions are compared here.

B. Limits on the Mean of the Gaussian Distribution

The Bayesian and frequentist methods of obtaining limits are first studied for the Gaussian distribution (μ=0), where ML estimations are free of biases and the effects of being unable to scan below 0. According to Bayesian techniques, in order to set an upper limit with a 90% confidence level on the mean of an observed distribution, the likelihood curve is integrated to the 90th percentile, as shown in Fig 23. This typical likelihood curve for a Gaussian distribution is nicely behaved, making it straightforward to interpret according to the Bayesian approach. A plot showing the many distributions of peak likelihood values generated by Monte Carlo for the frequentist method is shown in Fig 24. The line across the distributions is that connecting the 10th percentile point in each distribution that is used to set an upper limit on the mean of the observed Gaussian distribution. Note that the slope of this line is constant and close to one over the entire range of possible mean values. As is indicated, when a measurement is obtained from a distribution according to the ML

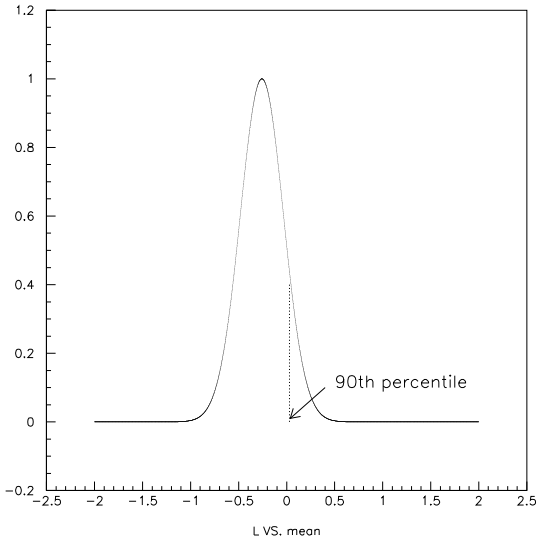

FIG. 23: A typical likelihood curve for experiments done with the Gaussian distribution centered at zero. This illustrates the Bayesian definition of limits, which requires integrating to the 90th percentile as shown by the dotted line.

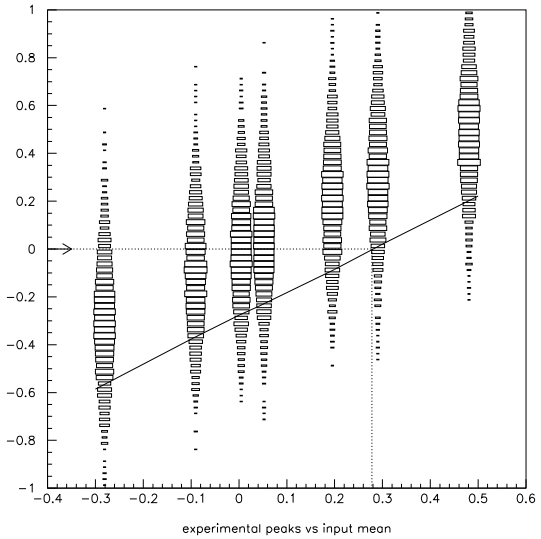

FIG. 24: A look-up table used to find upper limits on the mean of the Gaussian distribution (N=20) according to the frequentist definition of limits. The solid line connects the 10th percentile point in each distribution, and the dotted line indicates the method used to set an upper limit.

method, its value is found on the vertical axis and, moving across to the line and down to horizontal axis, its corresponding input mean value is found and reported as the upper limit.

To compare these two methods, we look at the histograms of upper limits values as defined by each technique for many different experiments (N=20). The distribution of upper limit values as defined according to the Bayesian approach is shown in Fig 25. Note that the number of upper limit values that lie below the true value is consistent with 10%. Although this is not required by the Bayesian definition, it comforting from a frequentist perspective that over many experiments producing upper limits with a 90% confidence level, 90% of

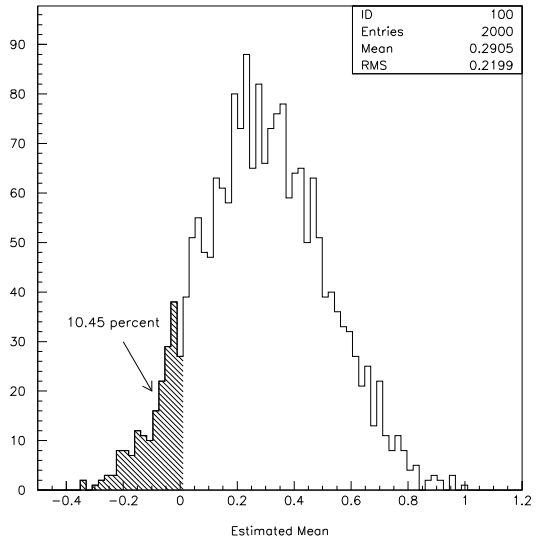

FIG. 25: Histogram of upper limits on the mean of the Gaussian distribution (N=20) according to the Bayesian method. The shaded region indicates those upper limits that are below the true mean value.

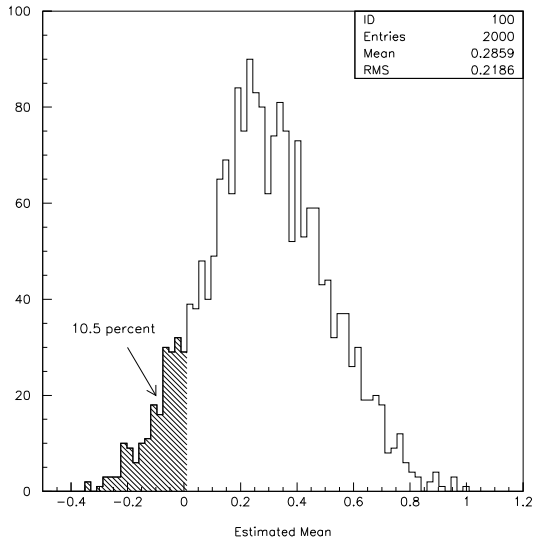

FIG. 26: Histogram of upper limits on the mean of the Gaussian distribution (N=20) according the the frequentist method. The shaded region indicates those upper limit values that are below the true mean value.

these limits are valid. The histogram of upper limits according to the frequentist method, for the same group of experiments, is shown in Fig 26. Note that by construction, around 10% of the upper limit values are below the true value. Also apparent is that these upper limits histograms for each method are similar in shape and mean value. Although not shown here, these definitions are found to remain consistent throughout the entire range of possible mean values for the Gaussian distribution.

C. Limits on the Mass for the Neutrino Mass Distribution

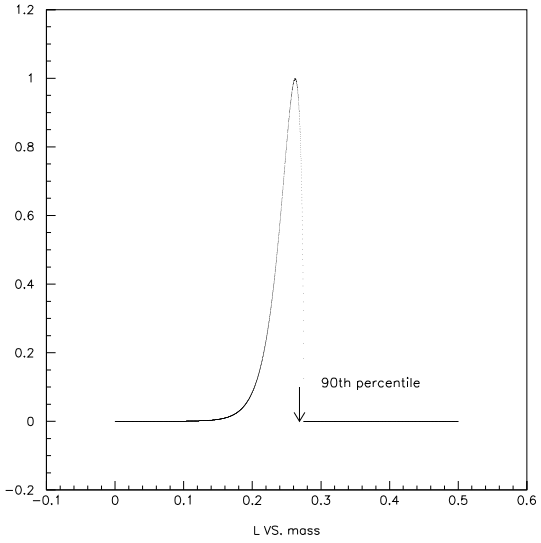

FIG. 27: A typical likelihood curve for an experiment done for the neutrino mass distribution without background (N=20), illustrating the Bayesian definition of upper limits on the mass. The curve is integrated to the 90th percentile, as indicated.

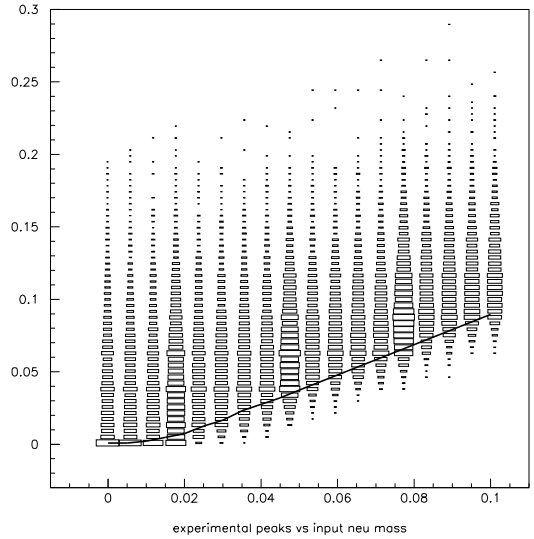

FIG. 28: A look-up table used to find upper limits on the mass for the neutrino mass distribution without background (N=20) according to the frequentist definition of limits. The solid line connects the 10th percentile point in each distribution.

We now compare these two different definitions of limits as applied to the neutrino mass distribution (N=20, without background). A typical likelihood curve for the distribution for a 200MeV neutrino is integrated to the 90th percentile according to the Bayesian definition, as is shown in Fig 27. As is characteristic of the likelihood curves for this distribution it is asymmetric, possibly making it more difficult to interpret in the Bayesian approach. In

Fig 28 the plot used to define upper limits according the frequentist technique is shown. The line through the vertical distributions is again at the 10th percentile for use in setting upper limits. What is important to note from this plot is the behavior of this line. In the high input neutrino mass region (above about 50MeV) it is fairly linear, with a slope of nearly one. What is interesting is its behavior in the lower input mass region, where it changes slope. This effect is a direct result of the inability to scan below zero while applying the ML estimator and the pile-up of estimation values at 0MeV. This information about the ML method, while obvious from this plot, is not revealed by the Bayesian method.

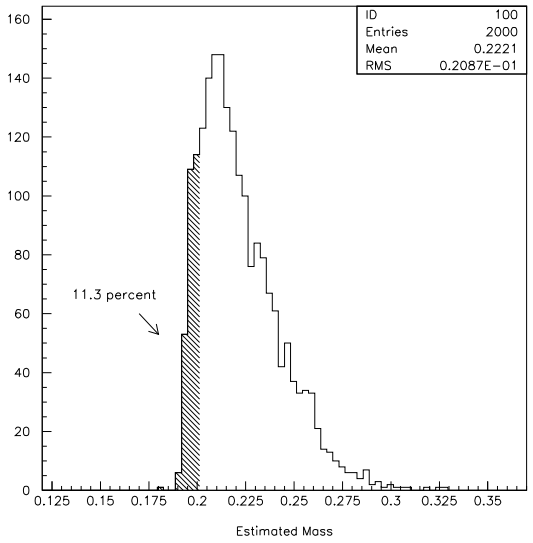

FIG. 29: Histogram of upper limit values as defined by the Bayesian technique for many experiments done for the neutrino mass distribution (N=20) generated for a 200MeV input mass. The shaded region indicates those upper limit values that are below the true mass value.

To compare these definitions, we look at the histograms of upper limit values for both high and low input neutrino mass distributions. The histogram of upper limits according to the Bayesian definition for many distributions (N=20) based upon a 200MeV neutrino mass is shown in Fig 29. Again, the number of upper limit values below the true value is consistent with 10%. The histogram according to the frequentist method, using the same experiments, is shown in Fig 30. The number of upper limits below the true value is around 10%, as it should be for this definition, and all other features are comparable to the histogram for the Bayesian approach. In the high neutrino mass region, where the line connecting the 10th percentile points on the frequentist plot is linear, these two definitions are found to be nearly equivalent.

The region of interest in an actual neutrino mass analysis is well below 50 MeV, so distributions in this region are studied here. The histogram of upper limits according to the Bayesian method for many distributions (N=20) generated for a 30MeV neutrino mass is shown in Fig 31. Note that the number of upper limits below the true value for this distribution is much less than 10%, indicating that the Bayesian approach is producing conservative upper limits. The histogram of upper limits according to the frequentist definition is shown in Fig 32. The number of upper limits below the true value for this method is still consistent with 10%, as it should be by definition. Also, the mean of this distribution is somewhat closer to the true value than the mean of the Bayesian limits, indicating that this technique

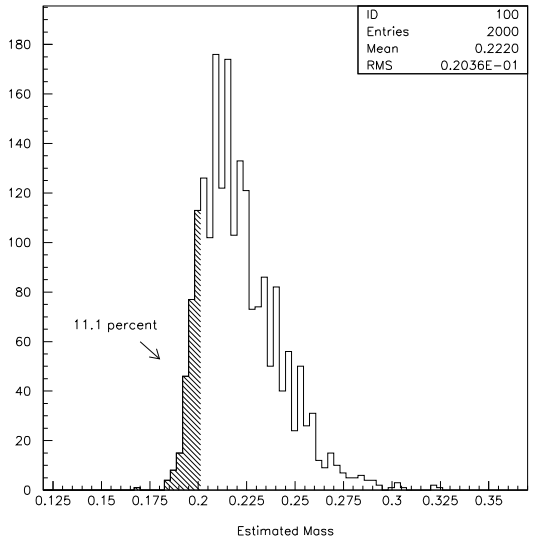

FIG. 30: Histogram of upper limit values as defined by the frequentist method for many experiments done for the neutrino mass distribution (N=20) generated for a 200MeV input mass. The shaded region indicates those upper limit values that are below the true mass value.

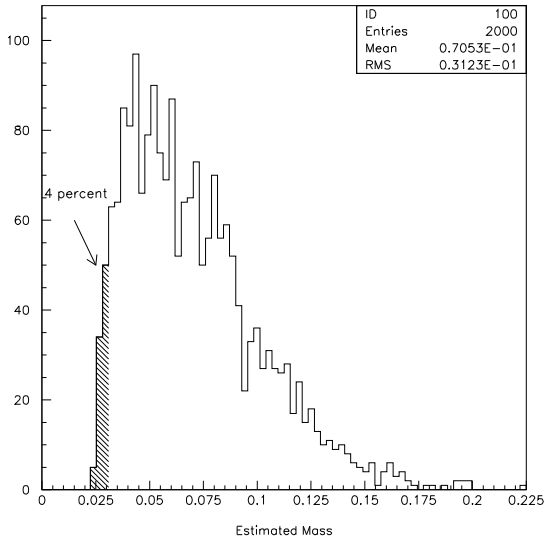

FIG. 31: Histogram of upper limit values as defined by the Bayesian technique for many experiments done for the neutrino mass distribution (N=20) generated for a 30MeV input mass. The shaded region indicates those upper limit values that are below the true mass value.

may correct for bias in the ML estimator. It is likely that the frequentist approach, serving as a calibration of all possible values, produces limits that take bias into account. These results indicate that the Bayesian and frequentist definitions may not produce similar upper limits in the neutrino mass region that is currently being studied, so one must be careful in choosing a definition of limits.

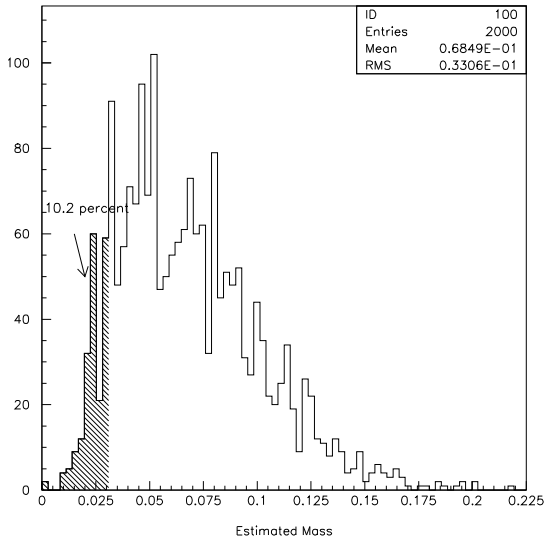

FIG. 32: Histogram of upper limit values as defined by the frequentist method for many experiments done for the neutrino mass distribution (N=20) generated for a 30MeV input mass. The shaded region indicates those upper limit values that are below the true mass value.

D. Alternate Definitions

There are other methods of calculating limits on the neutrino mass that, like the Bayesian and frequentist approaches, each answer slightly different questions. One of these methods uses distributions constructed similarly to those in the frequentist method, but integrates horizontally across the distributions as opposed to vertically to set limits. Although this method is not studied here, it should be investigated as another possible definition that may take biases into account.

Another method that has been used in neutrino mass analysis is an approach that is similar to the Bayesian definition[3]. It again treats the likelihood function as a probability distribution, but makes the additional assumption that it is Gaussian in shape and applies equations to obtain an upper limit accordingly. As is shown in all cases studied here, when estimating the endpoint of a distribution, or the neutrino mass, the typical likelihood curve is in general not a Gaussian. In addition, a lack of background and detector smearing results in likelihood functions that are asymmetrical and poorly behaved. Therefore, this definition seems to lead to the paradoxical situation in which it is dependent on these adverse effects to remain valid. Clearly this method of calculating limits is not viable in the case of the tophat, sloped tophat, and neutrino mass distributions presented here.

V. CONCLUSIONS

We observe bias in the application of the ML method to the estimation of the endpoints of the tophat, sloped tophat, and tau neutrino mass distributions. This suggests that bias may originate from the use of the ML estimator when the endpoints of the distributions involved are dependent on the parameter being estimated, and may be present in the estimation of the tau neutrino mass from real data.

A lack of background and distributions with sharp slopes near their endpoints are found

to cause poorly shaped likelihood curves that can be discontinuous in some cases. Therefore, likelihood distributions are in general not Gaussian, and their interpretation according to a Bayesian approach my not be straightforward. Also, the Bayesian definition of limits provides little information other than that of a single limit value, while the frequentist approach may correct for bias, and provide a calibration of all possible results. These two methods have very different behaviors in the neutrino mass region that is of interest in real studies. For these reasons we urge caution in setting limits on the tau neutrino mass.

VI. ACKNOWLEDGMENTS

I would like to thank Jean Duboscq of Cornell University for proposing this project and directing my research. He has made this a truly wonderful experience.

This work was supported by the National Science Foundation REU grant PHY-0243687 and research co-operative agreement PHY-9809799.

---

[1] R. Ammar *et al.,* (CLEO Collaboration) Phys. Lett. **B431**, 209 (1998).

[2] Barlow, R.J. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences* John Wiley & Sons, West Sussex, England:1989.

[3] D. Buskulic *et al.,* (ALEPH Collaboration) Phys. Lett. **B349**, 585 (1995).