STATISTICAL ANALYSIS OF THE LSND EVIDENCE AND THE KARMEN EXCLUSION FOR $\bar{\nu}_\mu \to \bar{\nu}_e$ OSCILLATIONS

Klaus Eitel
Forschungszentrum Karlsruhe, 76021 Karlsruhe, Germany

Abstract

A combined statistical analysis of the experimental results of the LSND and KARMEN $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillation search is presented. LSND has evidence for neutrino oscillations that is not confirmed by KARMEN. For both data sets, the analyses are based on likelihood functions. A frequentist approach is applied to deduce confidence regions for each experiment individually and for a combination of both. A detailed description of this work can be found in [1].

1. INTRODUCTION

The controversial results of the two experiments LSND (Liquid Scintillator Neutrino Detector [2] at LANSCE, Los Alamos, USA) and KARMEN (KArlsruhe Rutherford Medium Energy Neutrino experiment [3] at ISIS, Rutherford, UK) both searching for neutrino oscillations $\bar{\nu}_\mu \to \bar{\nu}_e$ have led to intense discussions. The two experiments are similar as they use $\bar{\nu}_\mu$ beams from the $\pi^+ - \mu^+$ decay at rest (DAR) chain $\pi^+ \to \mu^+ + \nu_\mu$ followed by $\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$ with energies up to 52 MeV. Furthermore, both experiments are looking for $\bar{\nu}_e$ from $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations via the reaction $p(\bar{\nu}_e, e^+)$ providing a spatially correlated delayed coincidence signature of a prompt e^+ and a subsequent neutron capture signal. LSND has observed a clear beam–on minus beam–off excess of events with $\bar{\nu}_e$ signature, i.e. (e^+, n) sequences. These have been interpreted as evidence for $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations [4]. On the other hand, KARMEN has found no excess events above the expected background.

The statistical analysis of the data has become a showcase of how to determine statistical significance and upper limits. KARMEN with no apparent $\bar{\nu}_\mu] \to \bar{\nu}_e$ signal and very low background has the problem of treating a result in a low statistics regime near the physical boundary $\sin^2(2\Theta) = 0$. In LSND, the maximum likelihood analysis of the data clearly indicates an oscillation signal. A problem arises when determining a region of correct confidence, i.e. statistical significance, in the $(\sin^2(2\Theta), \Delta m^2)$ plane having a likelihood function in two parameters, which shows a pathological behavior, namely an oscillatory dependence in Δm^2 with numerous local maxima. In 1998, the discussion was intensified by a paper of Feldman and Cousins [5], who described a method of dealing with the problems described above.

This report describes the individual evaluation of both data sets with maximum likelihood methods. The statistical interpretation of the likelihood functions and confidence regions is based on a frequentist approach and follows closely the analysis suggested by Feldman and Cousins. The main purpose of such an approach is to determine correct regions of confidence in $(\sin^2(2\Theta), \Delta m^2)$. A correct coverage is defined in terms of frequency, i.e. fraction of occurrence for future experiments. Probability or confidence in this context does not mean "degree of belief" as defined in a Bayesian statistics.

Although the central statements of LSND and KARMEN are contradicting there can be a region in the $(\sin^2(2\Theta), \Delta m^2)$ parameter space where the results are compatible. Combining the two experiments is done in different ways of constructing statistical distributions, pointing out that there is no unique way of determining regions of specific confidence. However, as we will see, the regions of compatibility in $(\sin^2(2\Theta), \Delta m^2)$ are very similar.

A statistical analysis combining two experimental results which apparently disagree is a delicate and controversial approach. It is not the task nor the purpose of this analysis to overcome this disagreement. However, assuming that there is no serious systematical error in either of the experiments and the

interpretation of their results with respect to oscillations $\bar{\nu}_\mu \to \bar{\nu}_e$, the question of statistical compatibility of the individual results is well justified and should be addressed quantitatively. This is the objective of the analysis presented in this paper.

2. DATA EVALUATION

2.1 KARMEN2 data

With an upgraded experimental configuration, KARMEN is running as KARMEN2 since February 1997. Starting as a simple counting experiment [6], the evaluation method was changed to a more sophisticated maximum likelihood analysis of the data set (Feb. 97 through Feb. 99), making use of detailed event information in energy, time and spatial position. After all cuts, 8 sequences remain. In total, the background expectation amounts to 7.8 ± 0.5 events. In order to extract more information from the 8 events about any potentially small oscillation signal a detailed maximum likelihood analysis was performed.

The likelihood function analyses 5 event parameters: the energies of the prompt signal, E_p, and the delayed event, E_d, the prompt time t_p and the delayed coincidence $\Delta t = t_d - t_p$ as well as the spatial correlation $\Delta \vec{x} = \vec{x}_d - \vec{x}_p$. The likelihood is calculated varying the oscillation signal r_{osc} as well as the background components relative to the overall data sample: r_{CC} for charged current events $^{12}\mathrm{C}(\nu_e, e^-)^{12}\mathrm{N}_{g.s.}$, r_{cos} for cosmic background, r_{ran} for random coincidences with a ν–induced prompt event and r_{con} for the intrinsic $\bar{\nu}_e$ contamination. With the condition $\sum_{j=1}^{5} r_j = 1$ and $\rho = (r_{osc}, r_{CC}, r_{cos}, r_{ran}, r_{con})$ the likelihood function for the $M = 8$ events can be written as

$$L(\rho) = \prod_{k=1}^{M} \{\sum_{j=1}^{5} r_j \cdot f_{j1}(E_p^k) \cdot f_{j2}(E_d^k) \cdot f_{j3}(t_p^k) \cdot f_{j4}(\Delta t^k) \cdot f_{j5}(\Delta \vec{x}^k)\} \times \prod_{j=2}^{5} P(r_j | r_j^{expected}) \quad (1)$$

The density functions f_{ji} contain the spectral information of all components, and as the positron energy spectrum depends on Δm^2, the dependence of L on Δm^2 enters via the density function f_{11}. The parameter $\sin^2(2\Theta)$ is determined by the ratio of oscillation events $N_{osc} = M \cdot r_{osc}$ divided by the expected number of events for maximal mixing $N_{exp}(\Delta m^2, \sin^2(2\Theta) = 1)$: $\sin^2(2\Theta) = N_{osc}/N_{exp}$. The second line in (1) is the combined Poisson probability $\prod P$ for the background contributions r_j calculated with the expectation values $r_j^{expected}$. For technical reasons, it is more convenient to optimize the logarithmic likelihood function lnL. Figure 1 shows lnL where the maximum in the physically allowed range $\sin^2(2\Theta) \geq 0$ has been renormalized to a value of $lnL(\sin^2(2\Theta) = 0, \Delta m^2) = 100$. From the likelihood function it is obvious that there is no oscillation signal in the data.
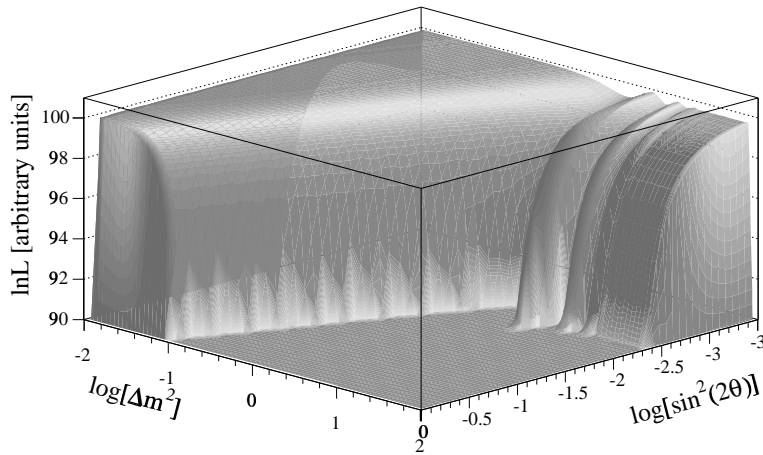

Fig. 1: Logarithmic likelihood function $lnL(\sin^2(2\Theta), \Delta m^2)$ for the 8 events of KARMEN2. The maximum in the physically allowed region $\sin^2(2\Theta) \geq 0$ is set to a value of 100, the minimum of this plot is set to 90.

2.2 LSND data

The LSND data analysed in this context have been reduced by requiring relatively loose criteria. Details of the event reconstruction and the definition of R can be found in [2] and [7]. To determine the oscillation parameters $\sin^2(2\Theta)$ and Δm^2, an event sample comprising 3049 beam-on events is used. Four variables are used to categorize the events: The energy of the primary electron, its spatial distribution in the detector expressed in distance L to the neutrino source and the angle $\cos\theta$ between the direction of the incident neutrino and the reconstructed electron path. The fourth variable is the likelihood ratio R for a (e^+,n) coincidence. The evaluation method uses these 4 correlated parameters to extract the oscillation signal, i.e. $\sin^2(2\Theta)$ and Δm^2, from beam related (BRB) and beam unrelated (BUB) background sources.

The likelihood function is the product of all M individual event likelihoods to fit a combination of 4-dim density distributions $f(E, R, L, \cos\theta)$ where the relative strengths r of the contributions are the parameters to be optimized with the side condition $\sum r_j = 1$. The likelihood function is defined as

$$L(r_{osc}, r_{brb}) = \prod_{k=1}^{M} \{r_{osc}f_{\Delta m^2}(E_k, R_k, L_k, \cos\theta_k) + r_{brb}f_{brb}(E_k, R_k, L_k, \cos\theta_k)$$

$$+(1 - r_{osc} - r_{brb})f_{bub}(E_k, R_k, L_k, \cos\theta_k)\} \cdot e^{-\frac{(r_{brb}M - N_{brb})^2}{2\sigma_{brb}^2}} \cdot e^{-\frac{(r_{bub}M - N_{bub})^2}{2\sigma_{bub}^2}} \quad (2)$$

There are effectively three free parameters: r_{osc} or $\sin^2(2\Theta)$, Δm^2 and r_{brb}. The Gaussian terms account for the background expectation values and their systematic and statistical uncertainties. The oscillation parameter $\sin^2(2\Theta)$ is determined as a function of Δm^2 according to $\sin^2(2\Theta) = r_{osc} \cdot M/N_{\Delta m^2}(\sin^2(2\Theta) = 1)$ where $N_{\Delta m^2}(\sin^2(2\Theta) = 1)$ indicates the number of oscillation events expected for a given Δm^2 and full mixing $\sin^2(2\Theta) = 1$ in the detector, taking all resolution functions and cuts into account. In a next step, the original likelihood function (2) is then integrated along the axis of the parameter r_{brb} which is of no further interest. The logarithmic likelihood lnL is therefore a function of the 2 free oscillation parameters $lnL(\sin^2(2\Theta), \Delta m^2)$ which is shown in Fig. 2. The exact position of the maximum in $(\sin^2(2\Theta), \Delta m^2)$ is not significant due to the flatness of the likelihood function along its 'ridge' for small values of Δm^2.
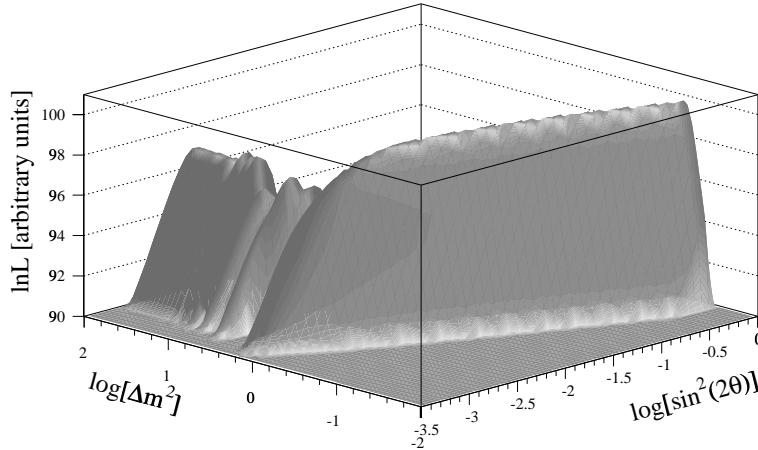

Fig. 2: Logarithmic likelihood function $lnL(\sin^2(2\Theta), \Delta m^2)$ for the LSND data 1993-1998 sample containg 3049 events. The maximum in the physically allowed region $\sin^2(2\Theta) \leq 1$ is set to a value of $lnL(\sin^2(2\Theta)_m, \Delta m_m^2) = 100$.

3. CONSTRUCTION OF CONFIDENCE REGIONS

The basic idea of getting correct confidence regions using the logarithmic likelihood function lnL is to create a statistic of an appropriate estimator based on a frequentist approach. A high number of

event samples is created by Monte Carlo using all experimental information on the event parameters. Different hypotheses are tested by including in the generated event samples oscillation events according to the oscillation parameters $(\sin^2(2\Theta), \Delta m^2)$. In these proceedings we will describe this method for the LSND experiment only and then show the results for both KARMEN and LSND.

For a preselected $\bar{\nu} \to \bar{\nu}_e$ oscillation hypothesis H with oscillation parameters $(\sin^2(2\Theta)_H, \Delta m_H^2)$ the creation of a LSND-like event sample is done in two steps. First, the number of oscillation events, BRB and BUB are thrown on the basis of the corresponding expectation values. In a second step, for each event, parameters (E,R,L,$\cos\theta$) are generated from the density functions $f_j(E, R, L, \cos\theta)$. The index j stands for the 3 different contributions. After an event sample is generated, the sample is analysed in exactly the same way as the experimental sample, i.e. the logarithm of the likelihood function (2) is calculated as a function of $(\sin^2(2\Theta), \Delta m^2)$.

In the following we demonstrate such a procedure on a specific example of an oscillation hypothesis H with parameters $(\sin^2(2\Theta)_H, \Delta m_H^2) = (4.2 \cdot 10^{-3}, \Delta m^2 = 1 eV^2)$ for which 1000 samples are generated by MC. To construct confidence regions, the distribution shown in Fig. 3 is the central estimator distribution suggested by [5] and should be read in the following way: To include the oscillation hypothesis $(\sin^2(2\Theta)_H, \Delta m_H^2)$ with a probability (frequency of occurrence) of 90%, the area in $(\sin^2(2\Theta), \Delta m^2)$ has to be defined by cutting lnL at a value of $\Delta lnL(90\%) = 3.25$ below the maximum for each individual likelihood function. This statistic as a function of ΔlnL shows the spreading of the maximal value of lnL compared to a given pair of oscillation parameters. If, for a given experiment, the value ΔlnL^{exp} is smaller than ΔlnL obtained for a specific hypothesis, such a parameter combination $(\sin^2(2\Theta)_H, \Delta m_H^2)$ would be included in the region of 90% confidence. For the LSND experimental result, the difference of the logarithmic likelihood function is 1.4, clearly within the 90% confidence region of the LSND experimental result.
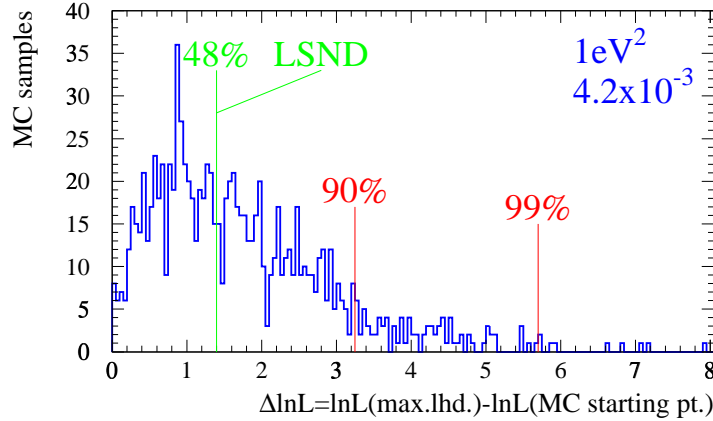

Fig. 3: Differences in lnL between the actual maxima and the values at the MC starting point. Also indicated is the difference of the logarithmic likelihood function $\Delta lnL = lnL(\sin^2(2\Theta)_m, \Delta m_m^2) - lnL(4.2 \cdot 10^{-3}, 1 eV^2) = 1.4$ for the LSND sample.

As $\Delta lnL(90\%)$ is itself a function of the parameters $(\sin^2(2\Theta)_H, \Delta m_H^2)$, the generation of MC samples has to be repeated for a grid of possible parameter combinations $(\sin^2(2\Theta), \Delta m^2)$ under consideration. The normalized distribution in Fig. 3 is named $C'(\Delta lnL)$ and the variable

$$lnL(\sin^2(2\Theta)_m, \Delta m_m^2) - lnL(\sin^2(2\Theta)_H, \Delta m_H^2) = \Delta lnL \equiv \Delta \quad . \tag{3}$$

Plotting the normalized integration of C' as function of Δ defined as

$$C(\Delta) = \frac{\int_0^\Delta C'(x)dx}{\int_0^\infty C'(x)dx} \tag{4}$$

allows an easy extraction of the 90% confidence value Δ^{90} for which $C(\Delta^{90}) = 0.9$. Shown in Fig. 4 are some distributions $C(\Delta_L)$ including the one for $(\sin^2(2\Theta)_H = 4.2 \cdot 10^{-3}, \Delta m_H^2 = 1 eV^2)$ for the LSND analysis. Note that these C distributions could be quite different.
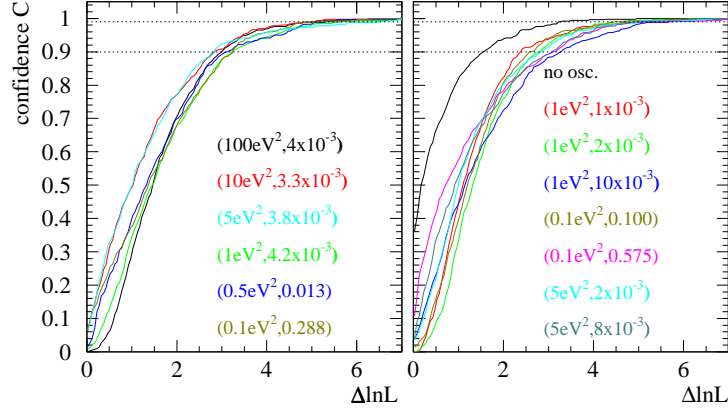

Fig. 4: Cumulative distributions or confidence $C(\Delta ln L_L)$ for various starting points $(\sin^2(2\Theta)_H, \Delta m_H^2)$. The left plot shows the distributions C for hypotheses with high likelihood for the LSND sample whereas the right figure is based on 'unlikely' starting hypotheses. The intersection of C with the dotted lines can be used to extract the Δ^{90} and Δ^{99} values.

On the basis of the distributions $C(\Delta)$ the values Δ^{CL} for a given confidence level CL are given for the calculated $(\sin^2(2\Theta)_H, \Delta m_H^2)$. The corresponding confidence regions for both experiments were then obtained by cutting the logarithmic likelihood function $lnL(\sin^2(2\Theta), \Delta m^2)$ at values of $\Delta^{CL}(\sin^2(2\Theta), \Delta m^2)$ below the absolute maximum of lnL. At 90% CL, each individual experimental outcome was compared with other experiments. Figure 5 shows the oscillation parameters inside the 90% CL LSND region and the 90% CL limits from KARMEN2 and other experiments. Notice that the limits of the Bugey $\bar{\nu}_e \rightarrow \bar{\nu}_x$ search [8], the CCFR combined $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ search [9] and the preliminary results from the NOMAD $\nu_\mu \rightarrow \nu_e$ search [10] are not based on this unified frequentist approach by Feldman and Cousins.
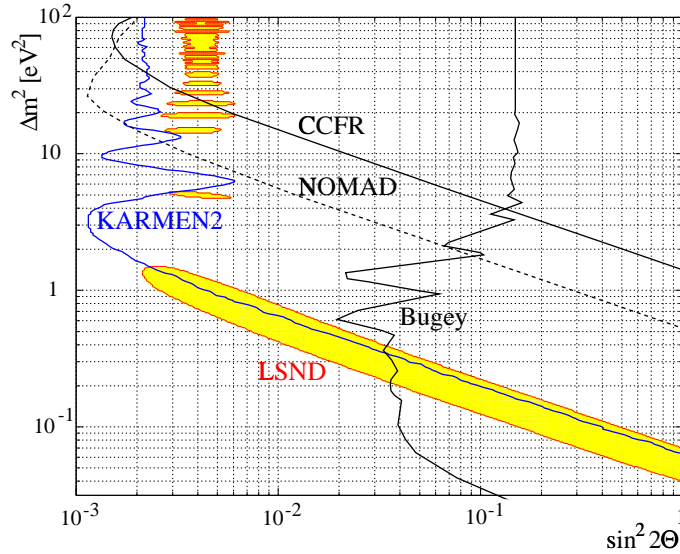

Fig. 5: LSND 90% CL region in comparison with other 90% CL exclusion curves in the corresponding $(\sin^2(2\Theta), \Delta m^2)$ area. The extraction of the 90% CL curves for NOMAD, CCFR and Bugey are not based on the frequentist approach used for KARMEN and LSND.

One of the most misleading but nevertheless very frequently used interpretation of the LSND and KARMEN results is to take the LSND region left of the KARMEN exclusion curve as area of $(\sin^2(2\Theta), \Delta m^2)$ 'left over'. Such an interpretation, though appealingly straight forward, completely ignores the information of both likelihood functions and reduces them to two discrete levels of individual 90% confidence. To be able to combine the two experimental results and extract combined confidence regions, we have to go some steps back to the original information of the distributions $C'_K(\sin^2(2\Theta), \Delta m^2)$ for KARMEN and $C'_L(\sin^2(2\Theta), \Delta m^2)$ for LSND.

4. COMBINING EXPERIMENTAL RESULTS

4.1 Likelihood functions

It is a well known procedure to multiply the likelihood functions of two independent experiments in order to combine the experimental results. Instead of multiplying the likelihood functions, an equivalent way is to add the logarithms. As already indicated in Figs. 1 and 2, there is some freedom in choosing the absolute scale of lnL. A convenient presentation of lnL is to normalize the individual functions lnL_K and lnL_L to a point in $(\sin^2(2\Theta), \Delta m^2)$ where they are equally sensitive to a potential signal. In our case of the oscillation search this corresponds to values of $\sin^2(2\Theta) = 0$. A stringent exclusion would then lead to only negative values of lnL whereas a strong signal leads to a significant maximum with a positive value of lnL. Hence, the combined logarithmic likelihood function can be expressed as

$$\begin{aligned} lnL(\sin^2(2\Theta), \Delta m^2) &= \{lnL_K(\sin^2(2\Theta), \Delta m^2) - lnL_K(\sin^2(2\Theta) = 0)\} \\ &+ \{lnL_L(\sin^2(2\Theta), \Delta m^2) - lnL_L(\sin^2(2\Theta) = 0)\} \end{aligned} \tag{5}$$

Figure 6 shows the combined function $lnL(\sin^2(2\Theta), \Delta m^2)$ with its maximum on a long flat 'ridge' of low Δm^2 values. Figure 7 shows slices for some values of Δm^2 for the three normalized functions lnL_K (leftmost or green curves), lnL_L (rightmost or blue curves) and lnL as defined in Eq. 5.
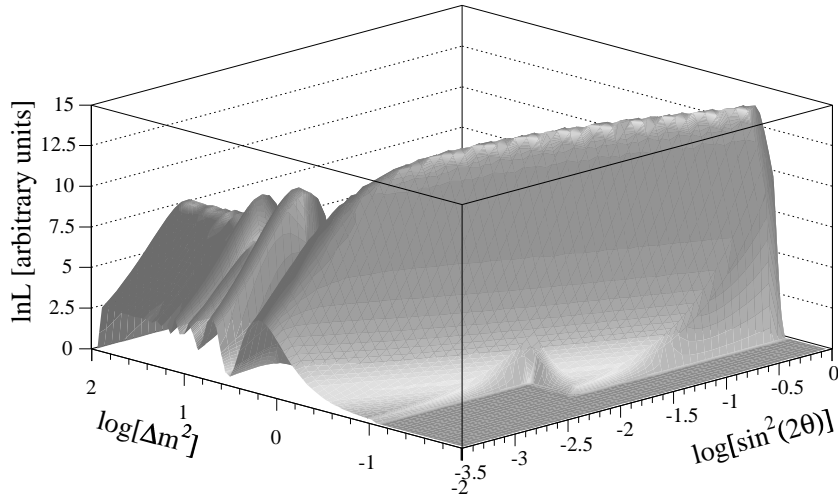

Fig. 6: Combined logarithmic likelihood function $lnL(\sin^2(2\Theta), \Delta m^2)$ as defined in Eq. (5).

The function $lnL(\sin^2(2\Theta), \Delta m^2)$ allows a direct qualitative interpretation of the experiments: There is a clear maximum of the combined likelihood function with a positive value of lnL favoring overall the evidence for oscillations given by LSND. On the other hand, compared to the individual LSND maximum, lnL_L, the negative KARMEN result reduces the maximal value by 1.6 units (see Fig. 7 for $\Delta m^2 = 0.1 \, \text{eV}^2$) which corresponds to a reduction to only 20% of the original maximal likelihood. This reduction of the global maximum is a direct reflection of the general disagreement of the two

experimental results. From Fig. 7 it is seen that for low Δm^2 the position in $\sin^2(2\Theta)$ of the maximum is not substantially shifted. In contrast, for larger Δm^2 the negative influence of the KARMEN result clearly shifts the maximum in $\sin^2(2\Theta)$ and strongly reduces the LSND likelihood value. It also increases the difference ΔlnL to the global maximum which is an important fact in terms of the statistics $C'(\Delta)$ and demonstrates that values of $\Delta m^2 > 2\,\text{eV}^2$ have a much smaller likelihood than some combinations $(\sin^2(2\Theta),\Delta m^2)$ in the low Δm^2 region. Although these observations help in assessing the combination of the two experiments, probability statements in a frequentist manner cannot be deduced from the above arguments. However, an evaluation of quantitative confidence regions can be based on the distributions $C'(\Delta)$, which is shown below.
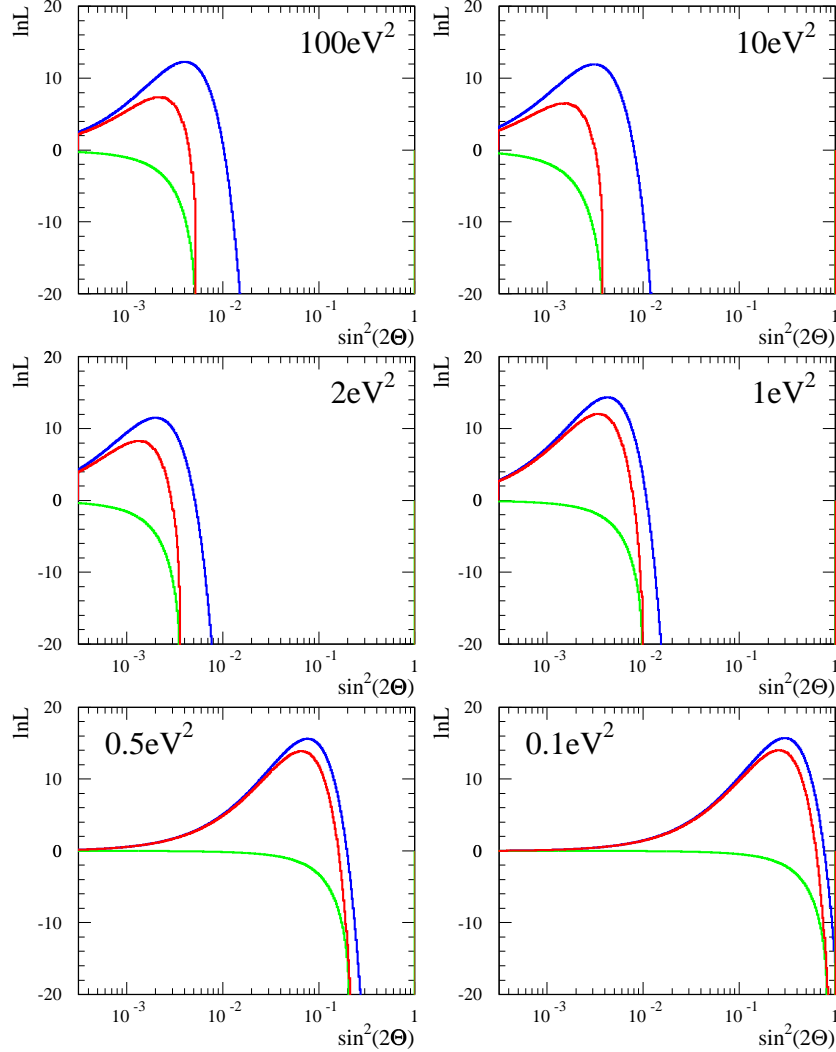

Fig. 7: Slices of constant Δm^2 of the logarithmic likelihood functions for KARMEN (leftmost or green), LSND (rightmost or blue) and the combination (middle or red). For definition of lnL see text.

4.2 Frequentist approach

In this section we describe 4 different methods to extract areas in $(\sin^2(2\Theta),\Delta m^2)$ of a certain confidence level CL. Though they can be derived analytically we follow a more phenomenological approach. The methods are based on different ways of ordering in a two dimensional space created by the individual statistics of the two experiments, C'_L and C'_K. The assumption that the two experiments LSND and KARMEN are independent is well justified. Therefore, a two dimensional distribution $C'(\Delta_L, \Delta_K)$ can

be constructed from the one dimensional normalized distributions $C'(\Delta_L)$ and $C'(\Delta_K)$ by an inverse projection. A box plot of $C'(\Delta_L, \Delta_K)$ and its original functions C' are shown in Fig. 8 for an example of a chosen parameter combination of $(\sin^2(2\Theta) = 4 \cdot 10^{-3}, \Delta m^2 = 2eV^2)$. The different lines in figure 8 correspond to the limits for 90% CL of the different methods described below.
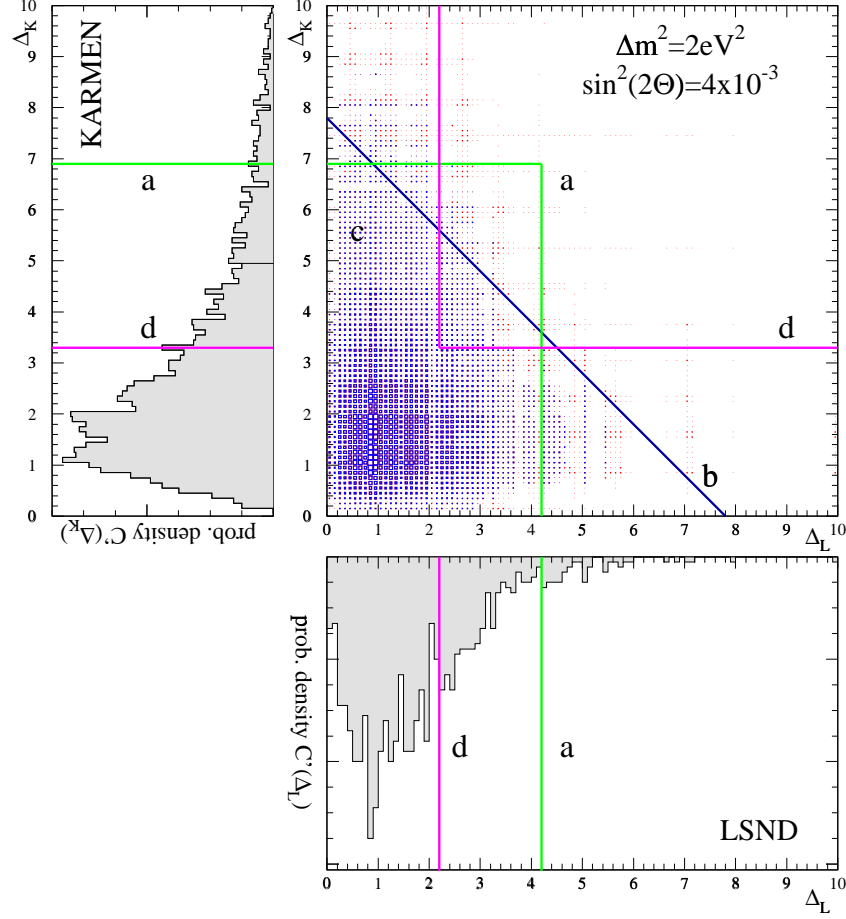

Fig. 8: Box plot of the two dimensional distribution $C'(\Delta_L, \Delta_K)$ for a given oscillation parameter combination $(\sin^2(2\Theta) = 4 \cdot 10^{-3}, \Delta m^2 = 2eV^2)$ and its projections for the individual experiments. The different combining methods indicated (a) through (d) are described in the text.

Method (a) combines LSND and KARMEN by integrating the distributions for both experiments $i = K, L$ individually. This corresponds to a rectangle in (Δ_L, Δ_K) defined by the side lengths Δ_K^{CL} and Δ_L^{CL}. The combined confidence is then $CL_{comb} = (CL)^2$. To obtain a confidence level of $CL_{comb} = 0.9$ we therefore have to determine Δ_i^{95}. The lines in Fig. 8 labeled (a) show these values Δ_i^{95} and the resulting rectangle in (Δ_L, Δ_K). If the experimental value $(\Delta_L^{exp}, \Delta_K^{exp})$ lies within this rectangle the parameter combination $(4 \cdot 10^{-3}, 2eV^2)$ is included in the combined 90% CL region. This method can be expressed also by taking the overlap of the $\sqrt{CL}$ confidence regions of both experiments to deduce the combined CL confidence region.

Method (b) is based on the combined statistic $C'(\Delta)$ with $\Delta = \Delta_L + \Delta_K$ defined as the convolution of the individual ones

$$C'(\Delta) = \int_0^\Delta C'_L(\Delta_L) \cdot C'_K(\Delta - \Delta_L) d\Delta_L \quad . \tag{6}$$

The confidence value Δ^{CL} is then defined by integration of C':

$$\int_0^{\Delta^{CL}} C'(\Delta)d\Delta = CL \quad . \tag{7}$$

For a given CL, the limit corresponds to a diagonal line in Fig. 8, where (b) indicates Δ^{90} for this specific $(\sin^2(2\Theta),\Delta m^2)$. The value $\Delta^{exp} = \Delta_L^{exp} + \Delta_K^{exp}$ is then compared with this Δ^{90}. If $\Delta^{exp} \leq \Delta^{90}$ the combination $(4 \cdot 10^{-3}, 2eV^2)$ is accepted at a 90% confidence level. Such an approach in (Δ_L, Δ_K) corresponds to an ordering along lines of constant combined likelihood, Δ below the two maxima of the likelihood functions.

Method (c) is based on an ordering principle of the elements $C'(\Delta_L, \Delta_K)$, i.e. the frequency or probability of occurrence of (Δ_L, Δ_K). This differs to integrating starting at $\Delta = 0$ as it is done in the previously described approaches. For a given confidence level CL, combinations (Δ_L, Δ_K) are added up in descending order starting with the highest probability of occurrence C' until a fraction of CL of the total $\int C'(\Delta_L, \Delta_K)d\Delta_L d\Delta_K$ is reached. In Fig. 8 this subset S of all (Δ_L, Δ_K) is shown in blue. If $(\Delta_L^{exp}, \Delta_K^{exp}) \in S$, the combination $(\sin^2(2\Theta),\Delta m^2)$ under consideration is included in the confidence region.

Method (d) results in a confidence region dramatically different to those obtained by all other methods. Instead of taking the overlap of two regions of $\sqrt{CL}$ confidence, the individual regions of $1 - (1 - CL)^2$ confidence are added to form the combined region of CL confidence. For a 90% CL this means adding (mathematically building the .OR. of) the regions of 68.4% individual confidence. In a graphical view, this is demonstrated by the line labelled (d) in Fig. 8.

It is instructive to discuss the differences of the methods by comparing the corresponding areas of the (Δ_L, Δ_K) plane (see Fig. 8) by each method. The triangle defined by (b) and the rectangle defined by (a) have almost the same area. In their corners with high values of Δ_i they allow experimental outcomes which are very unlikely, at least for one experiment. This drawback is overcome by the method (c) of ordering along probability of occurrence which has the disadvantage of principally disfavoring the unlikely, but very best fits of very small Δ_i. On the other hand, the convolution method integrates along contours of constant likelihood for the combined likelihood function which is a very plausible procedure.

The combined regions of 90% and 95% confidence are shown in Fig. 9 as green and yellow areas in $(\sin^2(2\Theta),\Delta m^2)$. The Figs. (a) through (d) correspond to the methods (a) through (d) described in section 4.2. Also shown for comparison are the individual experimental results: The KARMEN 90% CL exclusion curve (K) and the LSND 90% CL region (L) according to the frequentist approach (see Fig. 5) as well as the exclusion curves of the two experiments Bugey $\bar{\nu}_e \to \nu_x$ (B) and NOMAD $\nu_\mu \to \nu_e$ (N).

Comparing the results of methods (a) to (c), the confidence regions have only minor differences. High Δm^2 solutions are not excluded at 95% confidence, although the convolution and ordering methods clearly favor $\Delta m^2 < 10\,\mathrm{eV}^2$. The confidence region for $\Delta m^2 < 2\,\mathrm{eV}^2$ is almost identical for all combinations. At first sight, these regions are even similar to the 90% CL region of LSND only (see lines indicated with L in Fig. 9), however the combined 90% CL region extends to smaller values of $\sin^2(2\Theta)$ in the low Δm^2 region. For large Δm^2, the combined region is reduced and shifted to smaller mixing values. Although there are regions at $\Delta m^2 > 2\,\mathrm{eV}^2$ within a 90% CL these solutions have considerably smaller likelihood than along the 'ridge' at low Δm^2, as was discussed in section 4.1. This argument is underlined if regarding regions of combined confidence at an 80% confidence level. At such a level, none of the methods (a) through (c) include solutions above $\Delta m^2 = 2\,\mathrm{eV}^2$. Figure 9(d) shows a very distinct region of 90% confidence. It was chosen in this context only to demonstrate how regions of correct statistical confidence might differ and will not be discussed further.
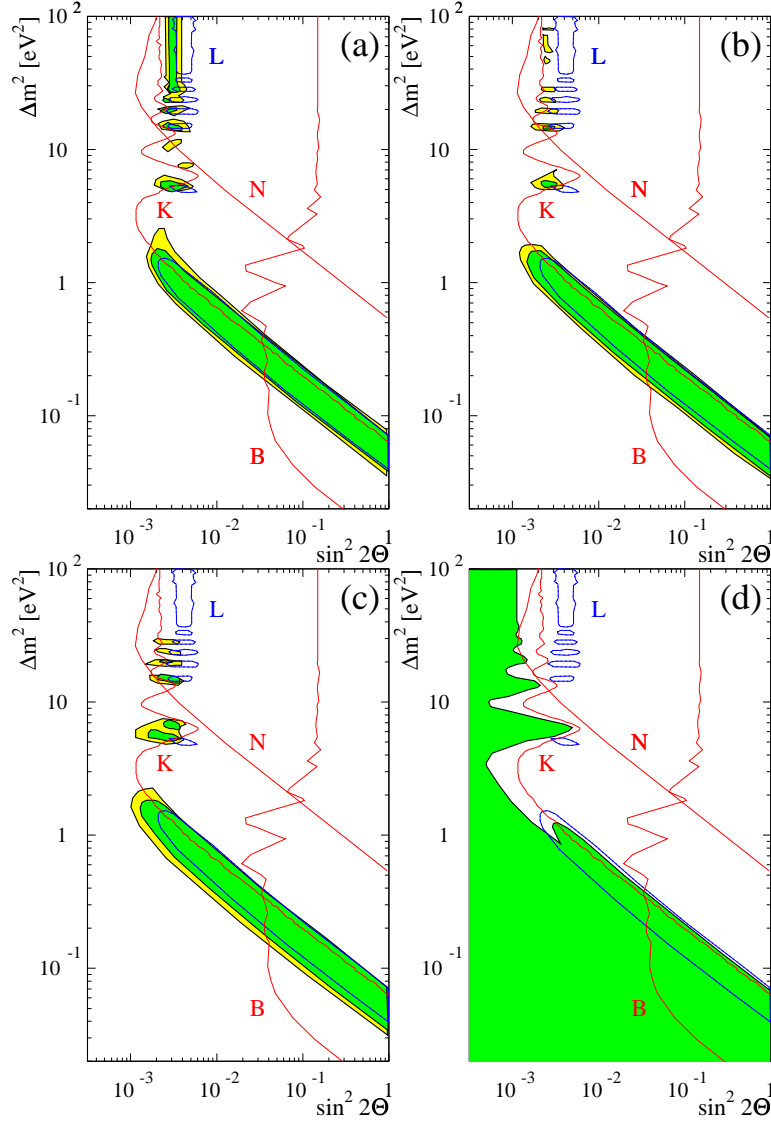
Fig. 9: Regions of 90% and 95% confidence for KARMEN and LSND combined as well as individual results of different experiments. See text for further explanations.

5. CONCLUSION AND OUTLOOK

The data sets of both the LSND and KARMEN experiment were analysed with a maximum likelihood method. For the first time, a frequentist approach based on [5] was applied to determine confidence regions of correct coverage for the LSND experiment. It is shown that in the case of a likelihood function depending on the oscillation parameters $\sin^2(2\Theta)$ and Δm^2, the approach assuming a two dimensional Gaussian likelihood function is only a rough approximation and does not lead to correct coverage. As both the KARMEN and LSND experimental data were analysed with a likelihood function and the statistics to deduce confidence regions were built in the same manner, it is possible to combine the likelihood functions and extract combined confidence regions based on a combination of the individual statistics created by Monte Carlo procedures. These regions are regions of correct coverage in terms of a frequentist approach.

This paper describes a statistical analysis combining both the LSND and KARMEN experimental outcomes and shows the feasability and results of such a method. As there are other experiments like NOMAD, CCFR and Bugey sensitive in part to the confidence region in $(\sin^2(2\Theta), \Delta m^2)$, a complete

analysis should also include these results on the basis of the same statistical analysis. This implies, however, the detailed knowledge of experimental data of these experiments not accessible to the author. In addition, the exclusion curve from the Bugey experiment is based on the disappearance search $\bar{\nu}_e \rightarrow \bar{\nu}_x$. Combining this experiment correctly with the appearance results of $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ or $\nu_\mu \rightarrow \nu_e$ in terms of mixing angles would therefore also require a full three or four dimensional (with a sterile neutrino) mixing scheme.

References

[1] K. Eitel, New Journal of Physics, **2** (2000) 1.1-1.25.

[2] C. Athanassopoulos et al., Nucl. Instr. and Meth. **A388** (1997) 149.

[3] G. Drexlin et al., Nucl. Instr. and Meth. **A289** (1990) 490.

[4] C. Athanassopoulos et al., Phys. Rev. Lett. **77** (1996) 3082; R. Tayloe et al., (1999) *proceedings of the Lake Louise Winter Institute 1999*, to be published by World Scientific Publishing Co.

[5] G. Feldman and R. Cousins R Phys. Rev. **D57** (1998) 3873.

[6] K. Eitel and B. Zeitnitz (1998) *hep-ex/9809007*.

[7] C. Athanassopoulos et al., Phys. Rev. **C54** (1996) 2685.

[8] B. Achkar et al., Nucl. Phys. **B434** (1995) 503.

[9] A. Romosan et al., Phys. Rev. Lett. **78** (1997) 2912.

[10] M. Mezzetto et al., Nucl. Phys. **B70** (1999) *(Proc. Suppl.)* 214.

Discussion after talk of Klaus Eitel. Chairman: Peter Igo-Kemenes.

R. Nolty

Was this work done in cooperation with LSND and how would you describe that cooperation?

K. Eitel

I spent a year at LSND so it was very nice, I really enjoyed it. You cannot do such an anlaysis if you don't have the full experimental information. Really you have to have all the information to create these likelihood functions and the Feldman/Cousins estimator distributions, in particular.

Bill Murray

As you say, you have all the information to create the likelihood functions and their distributions. I wasn't sure why you didn't do the full Cousins and Feldman analysis to the combined data set rather than coming up with some other prescription to combine the likelihood functions.

K. Eitel

First of all, if you combine the likelihood functions in themselves it's hard because you have to find a good way to normalize the likelihood function in a way that you weight both experiments in the same way. You see, the actual value of the likelihood function is completely different because you analyze different parameters in both experiments, so you have to think up how you really want to combine them at the likelihood level. In that plot where I showed the added likelihood function, it's easy at that stage because you just normalize it here [points to screen], both to zero so that you can combine them.

W. Murray

But in the Cousins/Feldman you normalize to the minimum so you have a defined normalization point.

K. Eitel

I think it's not so easy. Maybe we should really talk about that later in detail.

The technical problem is that the likelihood maximum of the combined likelihood function is different from the maxima of both individual likelihood functions. Therefore, one has to store (for each of thousands of MC samples!) the whole combined likelihood function, or do the Feldman/Cousins analysis simultaneously. I admit I didn't realize this in the beginning, and later the CPU time consumption didn't allow restarting the whole procedure. But for the final analysis of KARMEN2 and LSND, this will definitely be done.

C. Giunti

In your final result you have this big allowed region at low δm^2, and then you have some islands allowed at high δm^2. Can you say something on the credibility of these islands? For example, if you change the method, if you use B instead of A, what happens?

K. Eitel

If you really discuss these very small regions - I just want to show one point - this distribution (of the number versus the change in log-likelihood) is based on priors and samples, and you can imagine the

fluctuations here, so it's very hard to make anything at a percent level in confidence - you better do this at a level of at least 10 000 Monte Carlo samples, but that was actually too CPU time consuming. So, if you go into discussing these little islands, the small differences of the methods (a) through (c) are also due to statistical fluctuations of the samples. On the other hand, in terms of credibility, if you reduce your confidence level (i.e. become more stringent), if you don't look at 90% or 95% but let's say 80% , then all these areas at high δm^2 vanish.

G. D'Agostini

Just a comment. I have already discussed with the speaker about how misleading these kinds of results can be. For example, if somebody doesn't know the details of the analysis and looks only at this final 2-D plot, he understands that there is a region and some disconnected islands where these parameters could be, and the rest is excluded. If you see the 3-dimensional plots you understand much better what is going on. The left side is not excluded because in these kinds of problems, in which we are interested to give limits, the likelihood never goes to zero at the edge of the space of the parameters. This is shown very clearly in a 3-dimensional plot like this [shows transparency]. So there is certainly a strong region of evidence - you have very strong evidence - but obviously you cannot say that from a logical point of view, ("mathematically"), this is incompatible with background. You know, in fact, the likelihood never goes to zero.

K. Eitel

By construction I remained here in my normalization at zero, in yours this would be one. The only thing is that of course I compare this maximum with that value here [Points to transparency]