Physics Letters B 773 (2017) 307–312

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Reactor antineutrino shoulder explained by energy scale nonlinearities?

G. Mention [a,*], M. Vivier [a], J. Gaffiot [a], T. Lasserre [a,b], A. Letourneau [a], T. Materna [a]

[a] IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France
[b] AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/DRF/IRFU, Observatoire de Paris, Sorbonne Paris Cité, 75205 Paris Cedex 13, France

ARTICLE INFO

Article history:
Received 26 May 2017
Accepted 16 August 2017
Available online 23 August 2017
Editor: H. Weerts

Keywords:
Reactor
Antineutrino
Spectra
Energy nonlinearity
Statistical analysis

ABSTRACT

The Daya Bay, Double Chooz and RENO experiments recently observed a significant distortion in their detected reactor antineutrino spectra, being at odds with the current predictions. Although such a result suggests to revisit the current reactor antineutrino spectra modeling, an alternative scenario, which could potentially explain this anomaly, is explored in this letter. Using an appropriate statistical method, a study of the Daya Bay experiment energy scale is performed. While still being in agreement with the γ calibration data and ^{12}B measured spectrum, it is shown that a $\mathcal{O}(1\%)$ deviation of the energy scale reproduces the distortion observed in the Daya Bay spectrum, remaining within the quoted calibration uncertainties. Potential origins of such a deviation, which challenge the energy calibration of these detectors, are finally discussed.

1. Introduction

Reactor antineutrino experiments have played a leading role in neutrino physics starting with the discovery of the electron antineutrino in 1956 [1,2], through the first observed oscillation pattern in KamLAND [11], up to recent high precision measurements on the θ_{13} mixing angle [3–5]. Future projects JUNO [12] and RENO50 [13] even aim at reaching sub-percent accuracy on θ_{12} on top of solving the neutrino mass hierarchy puzzle. However, two anomalies in the measured antineutrino spectra are being observed. The first is an overall rate deficit around 6% known as "The reactor antineutrino anomaly" [8]. The second one is a shape distortion in the 4–6 MeV region, often quoted as a "bump" or "shoulder" in the spectra. It should be particularly stressed out that the relation between these two anomalies is not straightforward since shape distortion does not necessarily imply a change in the total rate.

This letter focuses on the second anomaly. In Section 2, a quantitative comparison of four reactor antineutrino experiments (Bugey 3 [9], Daya Bay [4], Double Chooz [3] and RENO [5]) is performed to demonstrate their incompatibility, thus questioning nuclear effects as a common origin, as proposed in [10]. The next sections are dedicated to the study of an alternative scenario accounting for the observed distortion. Section 3 reviews the energy scale determination in such reactor antineutrino experiments. Section 4 introduces a combined analysis of the Daya Bay calibration and reactor antineutrino data. Results are presented in section 5 and show that a 1% unaccounted break at 4 MeV in the energy scale can reproduce the observed antineutrino spectrum and still comply with calibration data within uncertainties. Section 6 discusses possible origins of such an energy nonlinearity and especially questions calibration of such detectors.

2. Reactor spectra comparison

2.1. On statistical compatibility of reactor spectra

Among all existing reactor antineutrino experiments, four of them give precise reactor spectra shape information. The Bugey 3 experiment (B3) [9] has until recently provided the finest reactor antineutrino spectrum. The B3 measurement was in very good agreement with previous predictions [14–16]. The comparison is here updated to the most recent predictions [6,7], after correcting for the rate anomaly [8]. As indicated on Fig. 1, the net effect is an additional 1%/MeV decrease through the full energy range. This update is still compatible with prediction within the 2% linear spectral uncertainty envelop quoted in [9]. New measurements have been provided by three experiments: Double Chooz (DC) [3], Daya Bay (DB) [4] and RENO (RN) [5]. Their ratios to the state of the art prediction [6,7], also corrected for the rate anomaly [8], are depicted on Fig. 1 and exhibit a significant deviation from unity around 5 MeV. At first glance they clearly show a common feature

* Corresponding author.
E-mail address: guillaume.mention@cea.fr (G. Mention).

http://dx.doi.org/10.1016/j.physletb.2017.08.035

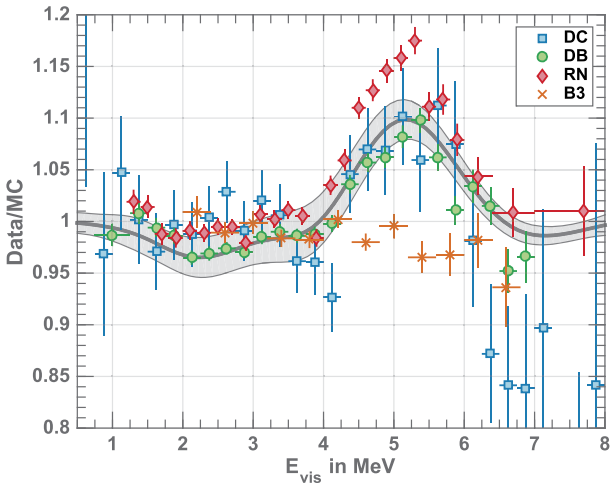

Fig. 1. Ratios of observed reactor antineutrino spectra to current best predictions [6,7], all corrected for the rate anomaly [8]. Despite similar fuel compositions, Double Chooz (DC), Daya Bay (DB) and RENO (RN) display significant deviations around 5 MeV, while Bugey 3 (B3) does not. The global best fit is indicated by the gray curve. A dedicated Monte Carlo simulation shows the compatibility p-value is below 10^{-5} at 99% CL. The uncertainty on the best fit is illustrated by the gray band. These results are therefore not compatible. B3 and coeval experiments before CHOOZ were all provided in units of kinetic positron energy, missing the two annihilation γ. We chose to use the visible energy and therefore shifted up B3 spectra by 1.022 MeV for comparison.

which is described as a bump in the 4 to 6 MeV region. Nevertheless, to our knowledge, no quantitative comparison is available in the literature.

To gain quantitative insights on their compatibility, each spectrum having different bin centers and widths, a direct χ^2 comparison is not possible. A bespoke statistical test was constructed to assess if all the observed spectra could come from a common unique distribution. This shared distribution was estimated thanks to a χ^2 approach using a Gaussian mixture model [21]. Such an approach offers the advantages of being reactor modeling independent and flexible enough to accurately fit each single data set. To do so, a large number ($K = 42$) of knots $(x_1, \ldots, x_K)$ were used, evenly distributed between 0 and 10 MeV. The spectral density, $f(x) = \frac{1}{h}\sum_{k=1}^{K} w_k \phi\left(\frac{x-x_k}{h}\right)$, with weight parameters w_k, a bandwidth fixed to the inter-knot distance $h = x_{k+1} - x_k$, ϕ being the standard normal probability distribution function, was then integrated over each bin width to properly model their content. On top of statistical uncertainties, normalization and linear energy scale uncertainties were included in covariance matrices. For DC a 1% in normalization and scale was used, for DB a 2.1% and 1%, for RN a 1.5% and 1% and for B3 a 4% and 1.4% as published by the collaborations [3–5,9]. The χ^2 was expanded with a quadratic term penalizing higher local curvatures for smoothness: $\lambda \int f''(x)^2\,dx$. The amplitude of the positive parameter λ controls the regularization strength or equivalently the smoothness of the fitting function. It was automatically determined from data through the generalized cross validation method [20]. This procedure is nearly equivalent to minimizing the model predictive error. The optimization of the global χ^2 was iteratively performed until reaching convergence, each step alternating between the χ^2 function minimization with respect to the w_k parameters and the generalized cross validation criterion minimization with respect to λ. The model was fitted to all possible subset of experiments. While our method was able to individually reproduce each spectrum accurately adjusting differently the w_k and λ coefficients, it prominently stressed out their inconsistency when combined. A parametric bootstrap procedure was used to estimate the χ^2_{min} distribution. For each experiment, the combined best fit model (gray curve on Fig. 1) was used to draw 10^4 Monte Carlo simulations using their respective covariance matrix. Each of these 10^4 data sets was fitted with the same global framework and the residual sum of squares was computed. Their distribution was found to follow with a good accuracy a Gamma probability density function with a shape parameter of $33\,(1 \pm 9\%)$ and a scale parameter of $2.8\,(1 \pm 9\%)$. The associated p-value was estimated using another 10^6 Monte Carlo simulations, because of the uncertainties on the Gamma PDF fitted parameters. The 99th percentile of the p-value distribution was found below 10^{-5} while the median was located around 10^{-8}. As a prominent conclusion, the observed spectra cannot come from a unique distribution with a high statistical significance (more than 4.4σ at 99% CL). This result is mostly driven by the mismatch between the DB, RN and B3 spectra since they have by far the highest statistics. Any combination of two of these three experiments is also not consistent. Note that DB and RN chose a different ad hoc normalization. A quick integration of their respective spectra seems to indicate a 2.9% offset in their respective normalization. Our aforementioned procedure outputs respectively a median 3.4 σ (4σ) incompatibility with (without) this normalization adjustment. A normalization free fit (increasing normalization uncertainties to infinity in both DB and RN covariance matrices), yields a consistent result with a 3.4 σ incompatibility (p-value below $7\,10^{-4}$). Except for DC whose tests are compatible with all other experiments within 1.6 σ, the lowest rejection significance between pairs of experiments was found for DB and B3, with a 2.5 σ significance. A free norm fit still yielded a 2.3 σ significance rejection of compatibility hypothesis.

Table 1
Fission fractions in % of Bugey 3, Double Chooz, Daya Bay and RENO experiments.

	^{235}U	^{239}Pu	^{238}U	^{241}Pu
Bugey 3	53.8	32.8	7.8	5.6
Double Chooz	49.6	35.1	8.7	6.6
Daya Bay	58.6	28.8	7.6	5.0
RENO	56.9	30.1	7.3	5.6

2.2. On isotopic compatibility of reactor spectra

On the antineutrino production side, all reactors are similar pressurized water reactors. Seeking for differences among experiments, the core isotopic composition is the main differing component which could play a major role in the antineutrino spectral shape. For a complete fuel burning (a 60–70 GWd/t burnup), the reactor antineutrino rates are known to roughly decrease by −10% and the spectra to tilt by approximately −4%/MeV between the fresh beginning and the far end of nuclear fuel burning. This effect is reduced by a factor 3 to 4 taking into account reactor operations where fuel assemblies are refreshed by thirds or quarters of the whole core. As indicated in Table 1, the average isotopic compositions over the data taking periods are very close from each other [3–5,9]. On the one hand, DB and RN have isotopic compositions which differ upmost by 1.7% (mixture of 6 cores with comparable fuels), but disagree on the distortion amplitude by 60%. On the other hand DC has the most burnt makeup among the four experiments and displays a distortion amplitude comparable to DB, while B3 reactor composition is half-way between DC and DB but does not observe any distortion in the spectrum. There is no simple coherent pattern explaining these variations with isotopic compositions. In a more complex scenario if the antineutrino spectra were extremely sensitive to fission fractions from each reactor, the near/far relative measurement strategy for the θ_{13} quest would have failed. With many cores and the disparity of solid angle exposures among each detector, the near/far ratios of DB and