

UNIVERSIDADE FEDERAL DE ALFENAS

Marllos Eustáquio Fonseca Fernandes

**Inferência bayesiana na determinação do ângulo de mistura  $\theta_{13}$ ,  
no experimento Double Chooz**

Poços de Caldas/MG  
2016

Marllos Eustáquio Fonseca Fernandes

**Inferência bayesiana na determinação do ângulo de  
mistura  $\theta_{13}$ , no experimento Double Chooz**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Física pela Universidade Federal de Alfenas. Área de concentração: Física das Partículas Elementares e Campos.

Orientador: Prof. Dr. Gustavo do Amaral Valdiviesso.

Poços de Caldas/MG

2016

F363i Fernandes, Marlos Eustáquio Fonseca.  
Inferência Bayesiana na determinação do ângulo de mistura  $\theta_{13}$ ,  
no experimento Double Chooz / Marlos Eustáquio Fonseca  
Fernandes. – Poços de Caldas, 2016.  
73 f. –

Orientador: Gustavo do Amaral Valdiviesso.  
Dissertação (Mestrado em Física) – Universidade Federal de  
Alfenas, Poços de Caldas, MG, 2016.  
Bibliografia.

1. Física. 2. Neutrinos. 3. Desintegração beta. I. Valdiviesso,  
Gustavo do Amaral. II. Título.

CDD: 530

Marllos Eustáquio Fonseca Fernandes

## **Inferência bayesiana na determinação do ângulo de mistura $\theta_{13}$ , no experimento Double Chooz**

A Banca examinadora abaixo assinada, aprova a dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Física pela Universidade Federal de Alfenas. Área de concentração: Física das Partículas Elementares e Campos.



---

**Prof. Dr. Gustavo do Amaral Valdiviesso**  
Universidade Federal de Alfenas  
Orientador



---

**Prof. Dr. Fernando Gonçalves Gardim**  
Universidade Federal de Alfenas



---

**Prof. Dr. Pedro Cunha de Holanda**  
Universidade Estadual de Campinas

Aprovada em 14/04/2016.

## AGRADECIMENTOS

“Pois, Nele, foram criadas todas as coisas, nos céus e sobre a terra, as visíveis e as invisíveis.”

*Colossenses 1.16a*

Agradeço ao meu Senhor e Salvador Jesus, o Filho Único de DEUS, que foi a minha confiança e a minha esperança em todas as conquistas na minha vida. Muito obrigado meu Deus!

Agradeço a minha esposa Carla Nunes Fonseca Fernandes, por esses dois anos que enfrentamos juntos e vencemos juntos. Minha companheira fiel e presente.

Agradeço ao meu orientador Prof. Dr. Gustavo do Amaral Valdiviesso, por tudo que me ensinou como físico e como pessoa. O mundo torna-se um lugar melhor quando as pessoas passam para frente o bem que aprenderam.

Agradeço ao Prof. Dr. Pietro Chimenti (UFABC) por ter aberto a oportunidade de trabalharmos juntos no seu projeto de Análise Bayesiana do Experimento de Double Chooz.

Agradeço a todos meus professores da UNIFAL-MG: Prof. Dr. Person Pereira Neves, Prof. Dr. Gustavo do Amaral Valdiviesso, Prof. Dr. Ihosvany Camps Rodriguez, Prof. Dr. Fernando Gonçalves Gardim e Prof. Dr. Cássius Anderson Miquele de Melo. Foi muito gratificante o convívio com todos eles. São grandes exemplos que espero poder seguir.

Agradeço aos meus irmãos da Igreja Batista Nacional de Alfenas, por todo amor, carinho e oração.

## RESUMO

O Experimento Double Chooz, como outros experimentos, utilizam detectores baseados na identificação de reações beta inversas. O nêutron e o pósitron gerados por tais reações geram radiação que são capturadas por fotomultiplicadoras o que permite recompor a energia do neutrino que desencadeou todo o processo. Desse modo é possível obter o espectro de energia dos neutrinos provenientes do reator. O mesmo espectro pode ser previsto com base em modelos de reações termonucleares que ocorrem no reator. A comparação desses dois espectros permite identificar a existência das oscilações de sabores dos neutrinos, além de podermos quantificar o principal parâmetro dessa oscilação, que é o ângulo de mistura  $\theta_{13}$ . Entretanto, existem processos físicos, como radioatividade natural presente no meio circuncidante, que geram sinais que não podem ser distinguidos dos sinais gerados por neutrinos. Esses processos físicos produzem backgrounds que precisam ser descontados na identificação dos neutrinos. Em nosso trabalho, é utilizada a análise bayesiana para inferir o ângulo de mistura  $\theta_{13}$ , levando-se em conta todos esses backgrounds. A característica diferenciada da análise bayesiana é a possibilidade de utilizar distribuições de probabilidades definidas com base na percepção subjetiva do pesquisador e nos dados previamente estabelecidos a respeito do objeto de estudo.

**Palavras-chave:** Ângulo de mistura  $\theta_{13}$ . Experimento de Double Chooz. Inferência bayesiana. Neutrinos de reatores.

## ABSTRACT

The Experiment Double Chooz, as other experiments, use detectors based on the identification of inverse beta reactions. The neutron and positron generated by such reactions generate radiation that are catches by photomultiplier that allows restore the energy of the neutrino that triggered the whole process. In this way you can get the energy spectrum of the neutrinos from the reactor. The same spectrum can be predicted based on models of thermonuclear reactions that occur in the reactor. The comparison of these two spectra allows to identify the existence of oscillation neutrino flavors, besides being able to quantify the main parameter of this oscillation, which is the mixing angle  $\theta_{13}$ . However, there are physical processes, such as natural radioactivity present in the surrounding medium that generate signals that can not be distinguished from signals generated by neutrino. These physical processes produce backgrounds that need to be discounted in the identification of neutrinos. In our work, the Bayesian analysis is used to infer the mixing angle  $\theta_{13}$ , taking into account all these backgrounds. The distinctive feature of the Bayesian analysis is the possibility of using probability distributions defined based on the subjective perception of the researcher and the data previously established about the object of study.

**Keywords:** Mixing angle  $\theta_{13}$ . Double Chooz Experiment. Bayesian inference. Neutrinos reactors.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Família dos léptons . . . . .	13
Figura 2 – Mudança de base . . . . .	14
Figura 3 – Oscilação do sabor eletrônico em função da distância . . . . .	18
Figura 4 – Oscilação do sabor eletrônico em função da energia . . . . .	19
Figura 5 – Aproximação em duas famílias: $\nu_e, \nu_\mu$ , em função da distância . . . . .	22
Figura 6 – Aproximação em duas famílias: $\nu_e, \nu_\mu$ , em função da energia . . . . .	23
Figura 7 – Aproximação em duas famílias: $\nu_e, \nu_\tau$ , em função da distância . . . . .	24
Figura 8 – Oscilação de neutrinos e a posição do detector . . . . .	25
Figura 9 – Oscilação de neutrinos, visão geral . . . . .	25
Figura 10 – Probabilidade condicional . . . . .	26
Figura 11 – Partição do espaço amostral . . . . .	27
Figura 12 – Simulação do lançamento de uma moeda . . . . .	30
Figura 13 – Ruído gaussiano . . . . .	36
Figura 14 – Ajuste polinomial do terceiro grau . . . . .	44
Figura 15 – Exemplo de p.d.f. gerada pelo BAT . . . . .	45
Figura 16 – Valores de $\sin^2 2\theta_{13}$ ao longo do tempo . . . . .	47
Figura 17 – Decaimento beta inverso . . . . .	48
Figura 18 – Detector Double Chooz . . . . .	50
Figura 19 – Eficiência com Cf . . . . .	54
Figura 20 – Calibração com Cf . . . . .	54
Figura 21 – Ensaio de linearidade . . . . .	56
Figura 22 – Medidas de correlacionados . . . . .	57
Figura 23 – Medidas de acidentais . . . . .	58
Figura 24 – Medidas Gd-III . . . . .	58
Figura 25 – Dados do DCIII e melhor ajuste . . . . .	68
Figura 26 – P.d.f de $\sin^2 2\theta_{13}$ . . . . .	68
Figura 27 – Correlação de $\sin^2 2\theta_{13}$ e $\Delta m_{13}^2$ . . . . .	69



## LISTA DE TABELAS

Tabela 1 – Ajuste polinomial - ROOT . . . . .	44
Tabela 2 – Ajuste polinomial - BAT . . . . .	44
Tabela 3 – Intervalos do histograma para o sinal de prompt . . . . .	61
Tabela 4 – Valores iniciais para os parâmetros. . . . .	66
Tabela 5 – Valores iniciais e resultado final. . . . .	67
Tabela 6 – Matriz de covariância entre os parâmetros. . . . .	67
Tabela 7 – Matriz de correlação entre os parâmetros. . . . .	67

## SUMÁRIO

1	<b>OSCILAÇÃO DE NEUTRINOS</b>	12
1.1	Introdução	12
1.2	A partícula neutrino	13
1.3	Mistura massa e sabor	14
1.4	O caso de duas famílias	15
1.5	O caso de três famílias	17
2	<b>ANÁLISE BAYESIANA</b>	26
2.1	Teorema de Bayes	26
2.2	O problema da honestidade de uma moeda	28
2.3	Estimativa com dois parâmetros - ruído gaussiano	32
2.4	Estimativa com dois parâmetros, caso geral	36
2.5	Estimativa com M parâmetros	39
2.6	Método dos mínimos quadrados	41
3	<b>O EXPERIMENTO DOUBLE CHOOZ</b>	46
3.1	Neutrinos de reatores	46
3.2	Mecanismo de detecção	48
3.3	Estrutura do detector	48
3.3.1	Sinal imediato	50
3.3.2	Sinal atrasado	51
3.3.3	Eficiência na identificação de IBD	53
3.3.4	Identificações espúrias	55
3.3.4.1	Cosmogênicos	55
3.3.4.2	Correlacionados	56
3.3.4.3	Acidentais	57
3.4	Resultados	57
4	<b>DETERMINANDO <math>\theta_{13}</math></b>	59
4.1	Os parâmetros	59
4.2	Mínimos quadrados	59
4.3	Definição do $\chi^2$ do DCIII	60
4.4	Termos do $\chi^2$	61
4.4.1	Intervalos do histograma	61
4.4.2	Neutrinos previstos	61
4.4.3	Matriz de covariância	61

4.4.4	Pull terms . . . . .	62
4.4.5	Termos de reator desligado . . . . .	64
4.5	Resumo das entradas e melhor ajuste . . . . .	65
4.5.1	Escala de energia . . . . .	65
4.5.2	Backgrounds . . . . .	65
4.5.3	Reator desligado . . . . .	65
4.5.4	Resumo das entradas . . . . .	65
4.5.5	Resultados . . . . .	65
5	CONCLUSÃO . . . . .	70
	REFERÊNCIAS . . . . .	72

# 1 OSCILAÇÃO DE NEUTRINOS

## 1.1 Introdução

Em 1930, uma nova partícula é sugerida pelo físico austríaco Wolfgang Pauli(1). A ideia era que o princípio da conservação da energia pudesse ser mantido nos processos de decaimento beta. Essa partícula, mais tarde, recebeu o nome de neutrino. Com isso teve início a tarefa sempre difícil da sua detecção. Finalmente, em 1956, 18 anos depois, Frederick Reines e Clyde Cowan detectaram o neutrino pela primeira vez. Ele deixa de ser apenas uma partícula hipotética, cujo único propósito seria de salvar o princípio da conservação da energia.

Rapidamente as técnicas são melhoradas sendo possível detectar neutrinos provenientes de diversas fontes, como reatores nucleares, estrelas, atmosfera terrestre. Os neutrinos também ganharam *sabores*, que são estados quânticos que os classificam em função da partícula que também é criada junto com ele. Se o neutrino é criado junto com o elétron, temos o neutrino do elétron. Se é criado junto com o múon, temos o neutrino do múon. Se é criado junto com o tau, temos o neutrino do tau. Os sabores sendo respectivamente eletrônico, muônico e tauônico.

Na década de 60 surge um grande problema: os neutrinos provenientes do Sol não são observados na Terra na mesma taxa prevista pelo modelo solar padrão (MSP)(2). Vários experimentos comprovam essa observação, entre eles: Homestake(3), SAGE(4) e Gallex(5). Em geral podemos falar numa taxa em torno de 50% do fluxo esperado(6). A solução desse problema veio através da hipótese da oscilação de neutrinos. A base dessa hipótese é a seguinte: o que entendemos como sabor do neutrino é na verdade, uma superposição de estados quânticos de massas diferentes. Como veremos na seção 1.3, o hamiltoniano de propagação dos neutrinos fará com que o estado quântico de sabor oscile entre seus valores permitidos, durante a propagação, com uma probabilidade que depende a distância percorrida. Isto quer dizer que, um neutrino que é produzido no Sol, com o sabor eletrônico, pode atingir a Terra com o sabor muônico, por exemplo. Os experimentos mencionados acima, só podiam detectar um tipo de sabor, por isso os neutrinos que oscilaram no caminho eram invisíveis na detecção. Essa hipótese foi comprovada pelo experimento SNO(7) por ser ele capaz de detectar todos os sabores de neutrinos. Então, o fluxo de sabor eletrônico, produzido no Sol foi comparado com o fluxo total, incluindo todos os sabores, que atingem a Terra. A concordância entre esses fluxos comprovam a hipótese da oscilação.

## 1.2 A partícula neutrino

Neutrinos são partículas elementares, léptons, sem carga elétrica e de spin  $\frac{1}{2}$ . Embora o modelo padrão preveja para ele massa nula, sabe-se que ele possui massa muito pequena comparada com as outras partículas elementares. O limite superior da sua massa é cerca de quatro milionésimos da massa do elétron(8). Isto faz do neutrino uma partícula ultrarrelativística, que sempre se propaga com velocidade muito próxima da velocidade da luz(9).

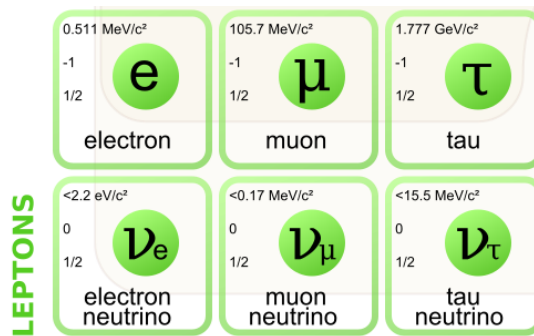


Figura 1 – Família dos léptons composta pelo elétron, múon, tau, com seus respectivos neutrinos. Os valores indicados verticalmente são: massa, carga elétrica, spin.

Fonte: Wikipédia (10).

A descrição quântica do neutrino necessita da definição de um observável chamado de sabor, indicado por  $F$  (do inglês *flavor*). Os neutrinos do elétron, do múon e do tau, são então autoestados desse observável. Sua equação de autovalor é escrita como

$$F |\nu_l\rangle = l |\nu_l\rangle, \quad (1.1)$$

onde  $l = (e, \mu, \tau)$  indica o autovalor do observável sabor. Imediatamente após criado, o neutrino fica sujeito a um novo hamiltoniano

$$H = \text{diag}\{E_i\} \quad (1.2)$$

que vai governar seu estado de propagação no espaço. Aqui temos um ponto singular para entender a oscilação do neutrino: a hipótese de que o hamiltoniano de propagação  $H$  não comuta com o observável de sabor  $F$ , isto é  $[H, F] \neq 0$ . Como consequência, os autoestados de  $F$ , nos quais os neutrinos são criados, não são autoestados de  $H$ , nos quais os neutrinos se propagam pelo espaço. Por isso, a equação de Schrodinger vai evoluir o estado inicial de sabor durante a propagação. Esta evolução, se dará, como veremos, em forma de oscilação entre os estados de  $F$ . Se os observáveis  $F$  e  $H$  comutassem, então o autoestado inicial dado por  $F$  também seria autoestado de  $H$ , e como sabemos, os autoestados do hamiltoniano são *estados estacionários*, portanto não oscilam. A hipótese de que os observáveis  $F$  e  $H$  não comutam é conhecida como a *hipótese da mistura massa e sabor*.

A equação de autovalores do hamiltoniano (1.2) é escrita como

$$H |\nu_i\rangle = E_i |\nu_i\rangle, \quad (1.3)$$

onde  $i = 1, 2, 3$ . O autoestados de  $H$  são chamados de estados de propagação ou estados de massa. Enquanto que os autoestados de sabores são chamados de estados de interação.

### 1.3 Mistura massa e sabor

Como o hamiltoniano  $H$  é um observável, seus autoestados ou autovetores  $\{|\nu_i\rangle\}$ , formam uma base para o espaço de estados do neutrino. O mesmo ocorre com os autoestados  $\{|\nu_l\rangle\}$  de  $F$ . Vamos reservar os índices  $i, j, k$  para os vetores da base de massa e  $l, m, n$  para os vetores da base de sabor. Os autovalores desses observáveis não podem ser medidos diretamente, como medimos a energia ou o momento de um elétron, por exemplo. O que podemos fazer é pensar na mudança entre as bases de massa e de sabor, e então medir os parâmetros associados com essa mudança. As bases são ortonormais, por isso teremos uma transformação unitária, que indicaremos por  $U$ . Numa visão geométrica, tal mudança de base equivale a uma rotação dos eixos definidos pelos vetores das bases, como mostra a figura (2) para o caso de duas famílias.

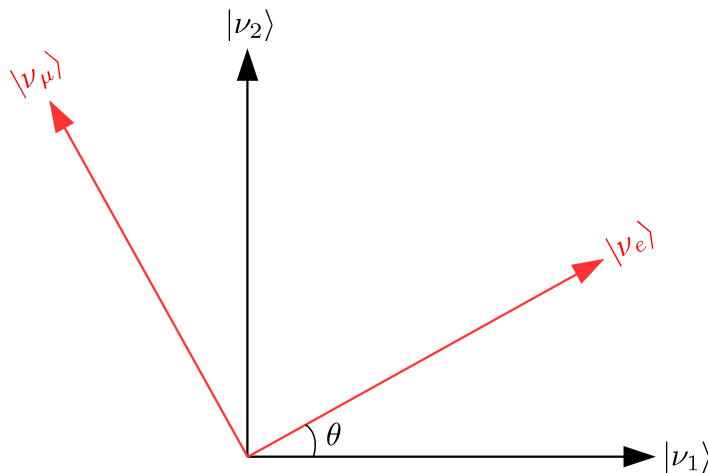


Figura 2 – Transformação ortogonal da base  $\{|\nu_i\rangle\}$  para a base  $\{|\nu_l\rangle\}$ . Equivale a uma rotação de ângulo  $\theta$  entre os eixos.

Fonte: Do autor.

Esta rotação pode ser representada pela seguinte equação

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \end{pmatrix}. \quad (1.4)$$

No caso de três famílias usa-se o que se chama de parametrização PMNS (Pontecovo, Maki, Nakagawa, Sakata). Ela determina o modo como os ângulos de rotação são definidos. Em PMNS, a rotação em torno do eixo  $k$ , normal ao plano  $i, j$ , é indicada por um operador

de rotação  $R_{ij}$  e o ângulo dessa rotação, é indicado por  $\theta_{ij}$ . Vamos abreviar  $\sin \theta_{ij}$ ,  $\cos \theta_{ij}$  como  $s_{ij}$ ,  $c_{ij}$  respectivamente. Então, podemos representar a matriz unitária  $U$  da rotação em três famílias como

$$U = R_{23}R_{13}R_{12} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23} & c_{12}c_{23} - s_{12}s_{13}s_{23} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23} & -c_{12}s_{23} - s_{12}s_{13}c_{23} & c_{13}c_{23} \end{pmatrix}. \quad (1.5)$$

Utilizando os dados mais recentes(8) sobre os parâmetros de (1.5) temos

$$U = \begin{pmatrix} 0,825 \pm 0,017 & 0,545 \pm 0,016 & 0,148 \pm 0,058 \\ -0,473 \pm 0,014 & 0,523 \pm 0,011 & 0,709 \pm 0,012 \\ 0,309 \pm 0,015 & -0,655 \pm 0,012 & 0,689 \pm 0,012 \end{pmatrix}. \quad (1.6)$$

#### 1.4 O caso de duas famílias

Vamos começar com o caso mais simples que é a oscilação em duas famílias. Por exemplo a família do elétron e a família do múon, ou a família do elétron e a família do tau. A segunda família podemos representar genericamente por  $\alpha$ . Então, em duas famílias os autovalores do observável sabor são  $l = (e, \alpha)$ . Suponha que o um neutrino tinha sido criado com sabor eletrônico no instante  $t = 0$ . A evolução temporal desse estado de sabor é dada pela equação de Schrodinger

$$i \frac{d}{dt} |\nu(t)\rangle = H |\nu(t)\rangle, \quad (1.7)$$

onde utilizou-se o sistema natural de unidades<sup>1</sup>. A solução de (1.7) é dada por

$$|\nu(t)\rangle = e^{-iHt} |\nu_e\rangle. \quad (1.8)$$

Para resolver a equação (1.8), afim de encontrar o estado do neutrino no instante  $t$ , temos duas opções: 1) expressamos o estado  $|\nu_e\rangle$  na base  $\{|\nu_i\rangle\}$  onde  $H$  é diagonal, ou 2) expressamos o hamiltoniano  $H$  na base interação  $\{|\nu_l\rangle\}$ . No caso mais simples, onde a propagação ocorre no vácuo, a primeira opção é a mais fácil e direta. Usando a equação (1.4) podemos expressar o estado inicial  $|\nu_e\rangle$  (base sabor) na base de massa

$$|\nu_e\rangle = \cos \theta |\nu_1\rangle + \sin \theta |\nu_2\rangle \quad (1.9a)$$

$$\langle \nu_e | = \cos \theta \langle \nu_1 | + \sin \theta \langle \nu_2 |. \quad (1.9b)$$

O estado do neutrino na base de massa fica como

$$\begin{aligned} |\nu(t)\rangle &= e^{-iHt} |\nu_e\rangle \\ &= e^{-iHt} (\cos \theta |\nu_1\rangle + \sin \theta |\nu_2\rangle) \end{aligned}$$

<sup>1</sup> O sistema natural de unidades é definido de forma que  $\hbar = 1$ ,  $c = 1$  e as grandezas físicas tenham suas unidades expressas como funções da energia. Para saber mais veja(6).

$$= e^{-iE_1 t} \cos \theta |\nu_1\rangle + e^{-iE_2 t} \sin \theta |\nu_2\rangle. \quad (1.10)$$

A probabilidade do neutrino permanecer no sabor eletrônico, no instante  $t$ , é chamada de probabilidade de sobrevivência  $P_{ee}(t)$ , e é dada por

$$P_{ee}(t) = |\langle \nu_e | \nu(t) \rangle|^2. \quad (1.11)$$

Fazendo os cálculos e simplificações, o resultado de (1.11) é

$$P_{ee}(t) = 1 - \sin^2(2\theta) \sin^2\left(\frac{\Delta E}{2} t\right), \quad (1.12)$$

onde  $\Delta E = E_1 - E_2$  é a diferença dos autovalores do hamiltoniano  $H$ . Esta equação já nos mostra que o estado de sabor do neutrino oscila no tempo, com período dado por

$$T = \frac{\pi}{\Delta E/2} = \frac{2\pi}{\Delta E}. \quad (1.13)$$

Agora é necessário desenvolver um pouco o termo  $\Delta E$ . Como o neutrino é uma partícula relativística, os valores de  $E_i$  são dados por

$$E_i = (m_i^2 + p_i^2)^{1/2}. \quad (1.14)$$

A massa do neutrino é da ordem de poucos  $\text{eV}^2$ , que é um valor muito pequeno comparado com a sua energia total, que podemos considerar como da ordem de  $1 \text{ MeV}$  (11). Por isso, dada uma energia  $E_i$ , o valor do momento  $p_i$  praticamente não altera, quando usamos em (1.14), valores de  $m_1$  e  $m_2$ . Então podemos deixar de distinguir os momentos  $p_i$  e aproximar a equação como

$$\begin{aligned} E_i &= (m_i^2 + p^2)^{1/2} \\ &= p \left(1 + \frac{m_i^2}{p^2}\right)^{1/2}. \end{aligned} \quad (1.15)$$

Dado que  $m_i^2 \ll p^2$ , podemos fazer a expansão em série de Taylor

$$\begin{aligned} E_i &\approx p \left(1 + \frac{m_i^2}{2p^2}\right) \\ &\approx p + \frac{m_i^2}{2p}. \end{aligned} \quad (1.16)$$

Na equação (1.16) vamos trocar  $p$  por  $E$  que é a energia associada ao valor de  $p$ , que também é aproximadamente a energia do neutrino.

$$E_i \approx E + \frac{m_i^2}{2E}. \quad (1.17)$$

O valor de  $\Delta E$  será

$$\Delta E = \frac{m_1^2 - m_2^2}{2E}$$



$$= \frac{\Delta m_{12}^2}{2E}. \quad (1.18)$$

A probabilidade de sobrevivência como  $\nu_e$ , fica então como

$$P_{ee}(t) = 1 - \sin^2(2\theta_{12}) \sin^2\left(\frac{\Delta m_{12}^2 t}{4E}\right). \quad (1.19)$$

Os detectores de neutrinos tem localização fixa, com distância  $x = L$  em relação a fonte. Assim é mais prático expressar  $P_{ee}$  em relação a dimensão espacial  $x$ , do que em relação a  $t$ . Como os neutrinos se propagam com velocidade muito próxima da velocidade da luz, podemos substituir  $t$  por  $x$  em  $P_{ee}(t)$

$$P_{ee}(x) = 1 - \sin^2(2\theta_{12}) \sin^2\left(\frac{\Delta m_{12}^2 x}{4E}\right). \quad (1.20)$$

Nos experimentos com oscilação de neutrinos, é comum que as medidas das distâncias  $x$  estejam em quilômetros (km), as diferenças dos quadrados das massas estejam em elétron-volt ao quadrado ( $\text{eV}^2$ ) e as energias em mega elétron-volt (MeV). Para utilizar essas unidades em (1.20), a expressão no argumento do segundo  $\sin^2$  deve ser alterada com o fator

$$1\,266,932\,73(43) \frac{\text{MeV}}{\text{eV}^2 \text{km}}. \quad (1.21)$$

O que podemos arredondar para 1267. Nessas unidades, a equação (1.20) fica como

$$P_{ee}(x) = 1 - \sin^2(2\theta_{12}) \sin^2\left(1267 \frac{\Delta m_{12}^2 \text{MeV}}{\text{eV}^2} \frac{x}{E \text{ km}}\right). \quad (1.22)$$

Os valores experimentais mais recentes para os parâmetros da (1.22) são(8)

$$\sin^2(2\theta_{12}) = 0,846 \pm 0,021 \quad (1.23a)$$

$$\Delta m_{12}^2 = (7,53 \pm 0,18) \times 10^{-5} \text{ eV}^2. \quad (1.23b)$$

Utilizando esses valores podemos fazer os gráficos mostrando a oscilação de neutrinos, de acordo com a equação (1.22). No caso da figura 3, a oscilação ocorre em função da distância entre neutrino e a fonte. Os neutrinos possuem energia de 1 MeV o que leva ao período de oscilação espacial de

$$L_0 = \frac{\pi E}{1267 \Delta m_{12}^2} = 32,9 \text{ km}. \quad (1.24)$$

## 1.5 O caso de três famílias

O próximo passo é considerar o caso mais geral, onde existem três famílias de neutrinos. A equação (1.4) é generalizada como

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \\ |\nu_\tau\rangle \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix}, \quad (1.25)$$

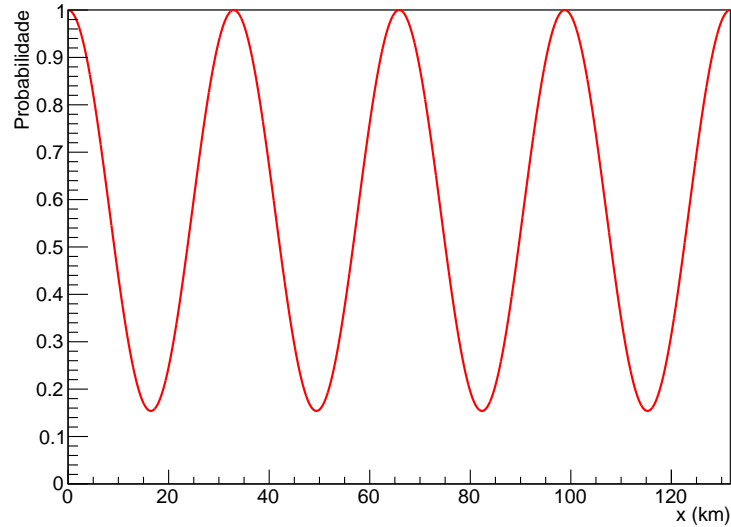


Figura 3 – Resultado da equação (1.22), em função da distância. Um neutrino  $\nu_e$  é produzido num reator, com energia  $E = 1$  MeV. O gráfico mostra a probabilidade desse neutrino sobreviver como  $\nu_e$  em função da distância  $x$  entre o neutrino e a fonte.

Fonte: Do autor.

o que pode ser escrito de modo mais compacto como

$$|\nu_l\rangle = \sum_i U_{li} |\nu_i\rangle, \quad (1.26a)$$

$$\langle \nu_l | = \sum_i U_{li}^* \langle \nu_i |. \quad (1.26b)$$

A solução equação de Schrodinger fornecendo a evolução do estado inicial é

$$\begin{aligned} |\nu(t)\rangle &= e^{-iHt} |\nu_l\rangle \\ &= e^{-iHt} \sum_i U_{li} |\nu_i\rangle \\ &= \sum_i e^{-iE_i t} U_{li} |\nu_i\rangle. \end{aligned} \quad (1.27)$$

Vamos calcular a probabilidade  $P_{lm}$  do neutrino, inicialmente no sabor  $l$ , de assumir o estado de sabor  $m$ , num instante posterior  $t$ . Primeiro temos que calcular o termo  $\langle \nu_m | \nu(t) \rangle$

$$\begin{aligned} \langle \nu_m | \nu(t) \rangle &= \sum_i U_{mi}^* \langle \nu_i | \sum_j e^{-iE_j t} U_{lj} |\nu_j\rangle \\ &= \sum_{ij} e^{-iE_j t} U_{mi}^* U_{lj} \langle \nu_i | \nu_j \rangle \\ &= \sum_i e^{-iE_i t} U_{mi}^* U_{li}. \end{aligned} \quad (1.28)$$

A probabilidade  $P_{lm}$  é dada por

$$P_{lm}(t) = \langle \nu_m | \nu(t) \rangle^2$$

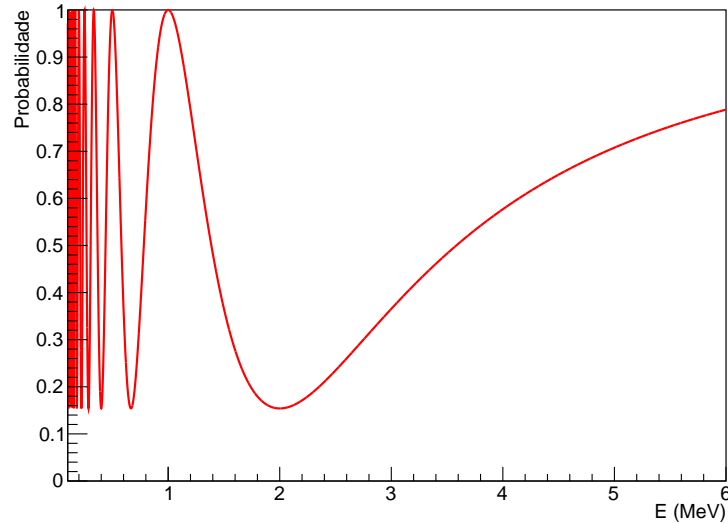


Figura 4 – Resultado da equação (1.22), em função da energia. Neutrinos  $\nu_e$  são produzidos num reator com diversas energias. Um detector está localizado à 33 km do reator. O gráfico mostra a probabilidade de sobrevivência do neutrino  $\nu_e$ , em função da sua energia, na posição fixa do detector.

Fonte: Do autor.

$$\begin{aligned}
 &= \langle \nu_m | \nu(t) \rangle \langle \nu_m | \nu(t) \rangle^* \\
 &= \sum_i e^{-iE_i t} U_{mi}^* U_{li} \sum_j e^{iE_j t} U_{mj} U_{lj}^* \\
 &= \sum_{ij} e^{-(E_i - E_j)t} U_{mi}^* U_{li} U_{mj} U_{lj}^*. \tag{1.29}
 \end{aligned}$$

Para trabalhar mais facilmente as próximas passagens, vamos definir

$$U_{mi}^* U_{li} U_{mj} U_{lj}^* = \rho_{ijlm} e^{i\varphi_{ijlm}}, \tag{1.30}$$

onde

$$\rho_{ijlm} = |U_{mi}^* U_{li} U_{mj} U_{lj}^*| \tag{1.31}$$

$$\varphi_{ijlm} = \arg(U_{mi}^* U_{li} U_{mj} U_{lj}^*). \tag{1.32}$$

Invertendo os índices  $i, j$  na equação (1.31)

$$\begin{aligned}
 \rho_{jilm} &= |U_{mj}^* U_{lj} U_{mi} U_{li}^*| \\
 &= |(U_{mj}^* U_{lj} U_{mi} U_{li}^*)^*| \\
 &= |U_{mj} U_{lj}^* U_{mi}^* U_{li}| \\
 &= |U_{mi}^* U_{li} U_{mj} U_{lj}^*| \\
 &= \rho_{ijlm}. \tag{1.33}
 \end{aligned}$$

O que mostra que o elemento  $\rho_{ijlm}$  é simétrico em relação aos índices  $i, j$ . Invertendo os índices  $i, j$  na equação (1.32)

$$\varphi_{jilm} = \arg(U_{mj}^* U_{lj} U_{mi} U_{li}^*)$$

$$\begin{aligned}
&= -\arg \left[ \left( U_{mj}^* U_{lj} U_{mi} U_{li}^* \right)^* \right] \\
&= -\arg \left( U_{li} U_{mi}^* U_{lj}^* U_{mj} \right) \\
&= -\arg \left( U_{mi}^* U_{li} U_{mj} U_{lj}^* \right) \\
&= -\varphi_{ijlm}.
\end{aligned} \tag{1.34}$$

Agora vamos mostrar uma propriedade das somatórias duplas.

$$\begin{aligned}
\sum_{ij} x_{ij} &= \sum_{ji} x_{ji} \text{ liberdade de escolher o nome dos índices} \\
&= \sum_{ij} x_{ji} \text{ inversão da ordem das somatórias} \\
\sum_{ij} x_{ij} + \sum_{ij} x_{ji} &= 2 \sum_{ij} x_{ij} \\
\sum_{ij} x_{ij} &= \frac{1}{2} \sum_{ij} (x_{ij} + x_{ji}).
\end{aligned} \tag{1.35}$$

Aplicando na na equação (1.29)

$$\begin{aligned}
P_{lm}(t) &= \sum_{ij} e^{-(E_i - E_j)t} U_{mi}^* U_{li} U_{mj} U_{lj}^* \\
&= \sum_{ij} e^{-(E_i - E_j)t} \rho_{ijlm} e^{i\varphi_{ijlm}} \\
&= \sum_{ij} \rho_{ijlm} e^{-[(E_i - E_j)t - i\varphi_{ijlm}]} \\
&= \frac{1}{2} \sum_{ij} \left( \rho_{ijlm} e^{-[(E_i - E_j)t - i\varphi_{ijlm}]} + \rho_{jilm} e^{-[(E_j - E_i)t - i\varphi_{jilm}]} \right) \\
&= \frac{1}{2} \sum_{ij} \left( \rho_{ijlm} e^{-[(E_i - E_j)t - i\varphi_{ijlm}]} + \rho_{ijlm} e^{-[(E_j - E_i)t + i\varphi_{ijlm}]} \right) \\
&= \sum_{ij} \rho_{ijlm} \frac{1}{2} \left( e^{-[(E_i - E_j)t - i\varphi_{ijlm}]} + e^{[(E_i - E_j)t - i\varphi_{ijlm}]} \right) \\
&= \sum_{ij} \rho_{ijlm} \cos [(E_i - E_j)t - \varphi_{ijlm}].
\end{aligned} \tag{1.36}$$

Quando não existe a violação de CP, os elementos  $U_{ij}$  são reais, o que leva em  $\varphi_{ij} = 0$ . Utilizando novamente a aproximação

$$E_i \approx E + \frac{m_i^2}{2E} \tag{1.37}$$

ou

$$E_i - E_j \approx \frac{\Delta m_{ij}^2}{2E}. \tag{1.38}$$

E considerando  $t = \frac{x}{c} = x$ , temos

$$P_{lm}(x) = \sum_{ij} \rho_{ijlm} \cos \left( \frac{\Delta m_{ij}^2}{2E} x \right). \tag{1.39}$$

Vamos agora calcular a probabilidade  $P_{ll}$  de sobrevivência de um estado  $l$  do neutrino

$$\rho_{ijll} = \left| U_{li}^* U_{li} U_{lj} U_{lj}^* \right| = U_{li}^2 U_{lj}^2, \quad (1.40)$$

e para simplificar o desenvolvimento, vamos definir

$$d_{ij} = \cos \left( \frac{\Delta m_{ij}^2}{2E} x \right), \quad (1.41)$$

então

$$\begin{aligned} P_{ll}(x) &= \sum_{ij} U_{li}^2 U_{lj}^2 d_{ij} \\ &= \sum_i \left( U_{li}^2 U_{li}^2 d_{ii} + \sum_{j \neq i} U_{li}^2 U_{lj}^2 d_{ij} \right) \\ &= \sum_i U_{li}^4 + \sum_{i,j \neq i} U_{li}^2 U_{lj}^2 d_{ij} \quad (d_{ii} = 1) \\ &= U_{l1}^4 + U_{l2}^4 + U_{l3}^4 + U_{l1}^2 U_{l2}^2 d_{12} + U_{l1}^2 U_{l3}^2 d_{13} + U_{l2}^2 U_{l1}^2 d_{21} + \\ &\quad U_{l2}^2 U_{l3}^2 d_{23} + U_{l3}^2 U_{l1}^2 d_{31} + U_{l3}^2 U_{l2}^2 d_{32}. \end{aligned} \quad (1.42)$$

Considerando que  $d_{ij} = \cos\left(\frac{\Delta m_{ij}^2}{2E} x\right) = \cos\left(-\frac{\Delta m_{ji}^2}{2E} x\right) = d_{ji}$

$$P_{ll}(x) = U_{l1}^4 + U_{l2}^4 + U_{l3}^4 + 2U_{l1}^2 U_{l2}^2 d_{12} + 2U_{l1}^2 U_{l3}^2 d_{13} + 2U_{l2}^2 U_{l3}^2 d_{23}. \quad (1.43)$$

Vamos considerar o caso particular, onde o sabor é  $l = e$

$$P_{ee}(x) = U_{e1}^4 + U_{e2}^4 + U_{e3}^4 + 2U_{e1}^2 U_{e2}^2 d_{12} + 2U_{e1}^2 U_{e3}^2 d_{13} + 2U_{e2}^2 U_{e3}^2 d_{23}, \quad (1.44)$$

que é a equação de sobrevivência do sabor eletrônico em três famílias.

Agora vamos fazer a aproximação em duas famílias ( $\nu_e, \nu_\mu$ ) da equação (1.44). Para isso, considere os valores de  $U_{ei}^2$ , obtidos a partir da primeira linha de (1.6)

$$(0,681 \pm 0,023 \quad 0,297 \pm 0,017 \quad 0,022 \pm 0,017). \quad (1.45)$$

Vemos que  $U_{e3}^2$  é pelo menos 13 vezes menor do que os outros  $U_{ei}$  que aparecem em (1.44), então essa equação pode ser aproximada por

$$\begin{aligned} P_{ee}(x) &= U_{e1}^4 + U_{e2}^4 + (U_{e3}^4 \approx 0) + 2U_{e1}^2 U_{e2}^2 d_{12} + (U_{e3}^2 \approx 0) (2U_{e1}^2 d_{13} + 2U_{e2}^2 d_{23}) \\ &\approx U_{e1}^4 + U_{e2}^4 + 2U_{e1}^2 U_{e2}^2 d_{12}. \end{aligned} \quad (1.46)$$

Utilizando (1.5)

$$\begin{aligned} P_{ee}(x) &= c_{12}^4 c_{13}^4 + s_{12}^4 c_{13}^4 + 2c_{12}^2 c_{13}^2 s_{12}^2 c_{13}^2 d_{12} \\ &= (c_{12}^4 + s_{12}^4 + 2c_{12}^2 s_{12}^2 d_{12})(c_{13}^4 \approx 1) \\ &\approx c_{12}^4 + s_{12}^4 + 2c_{12}^2 s_{12}^2 d_{12} \end{aligned}$$

$$\begin{aligned}
&\approx 1 - 2c_{12}^2 s_{12}^2 + 2c_{12}^2 s_{12}^2 d_{12} \\
&\approx 1 - 2c_{12}^2 s_{12}^2 (1 - d_{12}).
\end{aligned} \tag{1.47}$$

Usando com (1.5) e (1.41) para substituir  $c_{12}$ ,  $s_{12}$  e  $d_{12}$ , obtemos

$$P_{ee}(x) = 1 - \sin^2(2\theta_{12}) \sin^2\left(\frac{\Delta m_{12}^2 x}{4E}\right). \tag{1.48}$$

Que é o mesmo resultado obtido em (1.20). Portanto, a oscilação em duas famílias ( $\nu_e, \nu_\mu$ ) é uma aproximação do caso mais geral da oscilação em três famílias. Na figura 5 podemos comparar a aproximação da probabilidade de sobrevivência do sabor eletrônico em duas famílias ( $\nu_e, \nu_\mu$ ) com o caso geral de três famílias, em função da distância. A figura 6 faz a mesma comparação, mas em função da energia.

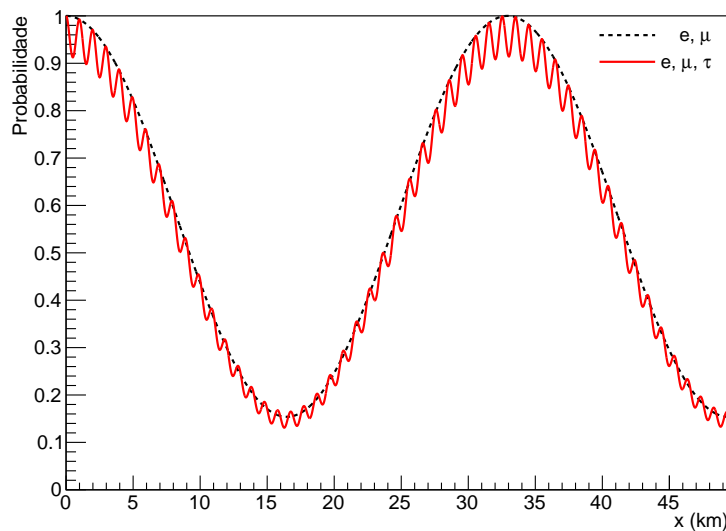


Figura 5 – Aproximação em duas famílias:  $\nu_e, \nu_\mu$ , em função da distância. A linha preta pontilhada indica a probabilidade de sobrevivência do sabor eletrônico em duas famílias:  $\nu_e, \nu_\mu$ , em função da distância  $x$  entre a fonte e o neutrino. A linha vermelha é a mesma probabilidade em três famílias:  $\nu_e, \nu_\mu$  e  $\nu_\tau$ . A energia do neutrino é  $E = 1$  MeV.

Fonte: Do autor.

Também podemos pensar numa aproximação em duas famílias considerando os sabores  $\nu_e$  e  $\nu_\tau$ . Esta é a aproximação de interesse para o experimento de Double Chooz, como veremos depois no capítulo 4. O fato de que a diferença dos quadrados de massas  $\Delta m_{13}^2$  ser 34 vezes maior do que  $\Delta m_{12}^2$  nos sugere que a oscilação, nesse caso, ocorre com comprimento espacial bem menor. O que nos leva a pensar numa aproximação para pequenos valores de  $x$ . Considere novamente a equação de sobrevivência do sabor eletrônico em três famílias

$$P_{ee}(x) = U_{e1}^4 + U_{e2}^4 + U_{e3}^4 + 2U_{e1}^2 U_{e2}^2 d_{12} + 2U_{e1}^2 U_{e3}^2 d_{13} + 2U_{e2}^2 U_{e3}^2 d_{23}, \tag{1.44}$$

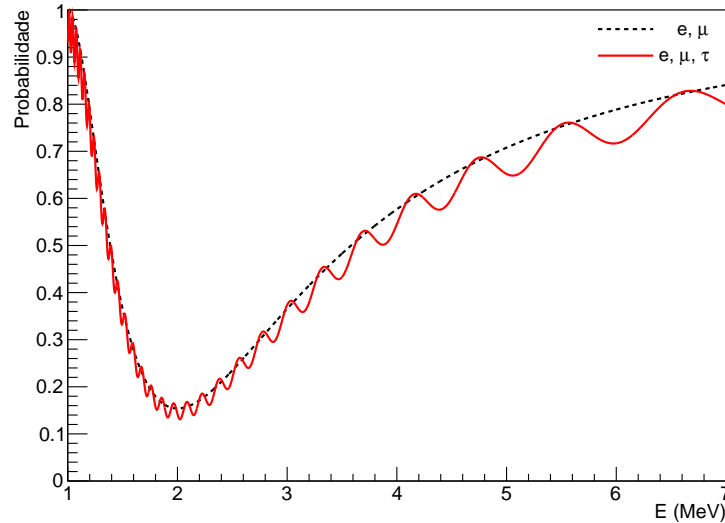


Figura 6 – Aproximação em duas famílias:  $\nu_e, \nu_\mu$ , em função da energia. A linha preta pontilhada indica a probabilidade de sobrevivência do sabor eletrônico em duas famílias:  $\nu_e, \nu_\mu$ , em função da energia  $E$  do neutrino. A linha vermelha é a mesma probabilidade em três famílias:  $\nu_e, \nu_\mu$  e  $\nu_\tau$ . A distância percorrida pelo neutrino entre a fonte e o detector está fixada em  $L = 33$  km.

Fonte: Do autor.

de acordo com os dados do [PDG](8) sabemos que  $\Delta m_{23}^2 = (2,44 \pm 0,06) \times 10^{-3} \approx \Delta m_{13}^2 = (2,52 \pm 0,12) \times 10^{-3}$  (eV<sup>2</sup>). Se  $\Delta m_{23}^2 \approx \Delta m_{13}^2$ , então  $d_{23} \approx d_{13}$ . Para pequenos valores de  $x$  associado com o pequeno valor de  $\Delta m_{12}^2$  em comparação com  $\Delta m_{13}^2$ , podemos fazer  $d_{12} \approx 1$ . Assim, a equação (1.44) se torna em

$$P_{ee}(x) = U_{e1}^4 + U_{e2}^4 + U_{e3}^4 + 2U_{e1}^2 U_{e2}^2 + 2(U_{e1}^2 + U_{e2}^2)U_{e3}^2 d_{13}. \quad (1.49)$$

Utilizando (1.5)

$$\begin{aligned} P_{ee}(x) &= c_{12}^4 c_{13}^4 + s_{12}^4 c_{13}^4 + s_{13}^4 + 2c_{12}^2 c_{13}^2 s_{12}^2 c_{13}^2 + 2(\underbrace{c_{12}^2 c_{13}^2 + s_{12}^2 c_{13}^2}_{c_{13}^2}) s_{13}^2 d_{13} \\ &= (c_{12}^4 + s_{12}^4 + 2s_{12}^2 c_{12}^2) c_{13}^4 + s_{13}^4 + 2s_{13}^2 c_{13}^2 d_{13} \\ &= (s_{12}^2 + c_{12}^2)^2 c_{13}^4 + s_{13}^4 + 2s_{13}^2 c_{13}^2 d_{13} \\ &= s_{13}^4 + c_{13}^4 + 2s_{13}^2 c_{13}^2 d_{13} \\ &= 1 - 2c_{13}^2 s_{13}^2 + 2s_{13}^2 c_{13}^2 d_{13} \\ &= 1 - 2s_{13}^2 c_{13}^2 (1 - d_{13}). \end{aligned} \quad (1.50)$$

Substituindo a definição de  $d_{ij}$  (1.41) e  $s_{ij}, c_{ij}$  na definição de (1.5), obtemos

$$P_{ee}(x) = 1 - \sin^2(2\theta_{13}) \sin^2\left(\frac{\Delta m_{13}^2}{4E} x\right) \quad (1.51)$$

Assim concluímos que a aproximação de famílias, formadas por  $\nu_e$  e  $\nu_\tau$ , para pequenos valores de  $x$ , tem a mesma forma da aproximação de duas famílias, formadas por  $\nu_e$  e  $\nu_\mu$ ,

dada por

$$P_{ee}(x) = 1 - \sin^2(2\theta_{12}) \sin^2\left(\frac{\Delta m_{12}^2}{4E}x\right). \quad (1.48)$$

A figura 7 mostra a aproximação em duas famílias, da probabilidade de sobrevivência do sabor eletrônico ( $\nu_e, \nu_\tau$ ), comparando-a com a mesma probabilidade em três famílias. No caso específico do experimento Double Chooz, o detector, chamado de FD (do inglês

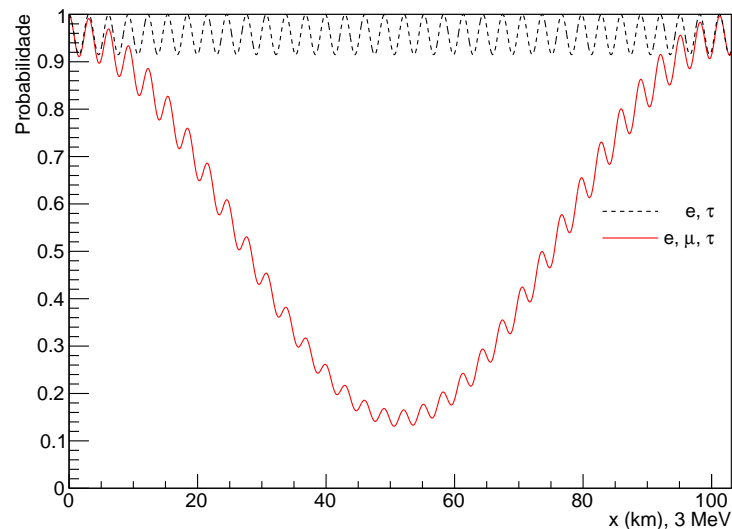


Figura 7 – Aproximação em duas famílias:  $\nu_e, \nu_\tau$ . A linha preta pontilhada indica a probabilidade de sobrevivência do sabor eletrônico em duas famílias ( $\nu_e, \nu_\tau$ ), em função da distância  $x$  entre a fonte e o neutrino. A linha vermelha é a mesma probabilidade em três famílias. A energia do neutrino é  $E = 3$  MeV.

Fonte: Do autor.

*fast detector*), foi posicionado em  $x = 1,05$  km. Como indica a figura 8, ele está localizado próximo ao primeiro mínimo da probabilidade de sobrevivência, afim de deixar mais evidente a influência da oscilação no processo de determinado do ângulo de mistura  $\theta_{12}$ . O quadro geral da oscilação dos neutrinos podem ser visto na figura 9.



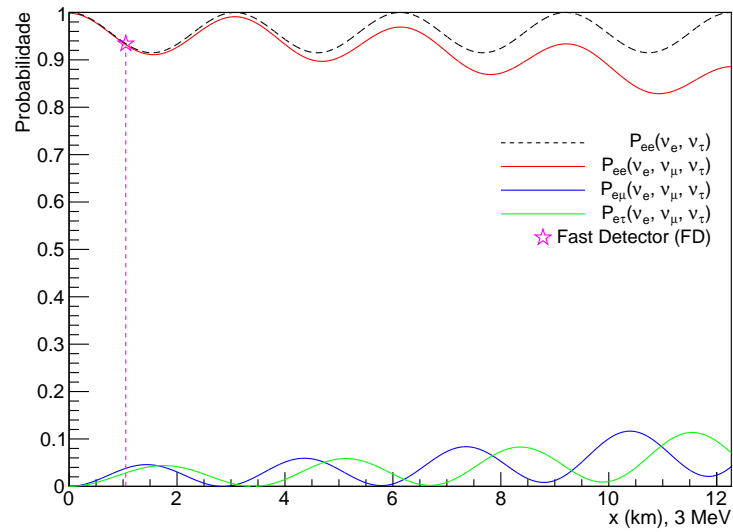


Figura 8 – Indicação das probabilidades de sobrevivência eletrônica nas linhas preta e vermelha, e das as transições para os sabores muônico, linha azul e tauônico, linha verde. A posição do detector, FD, ocorre próxima ao primeiro mínimo de oscilação.

Fonte: Do autor.

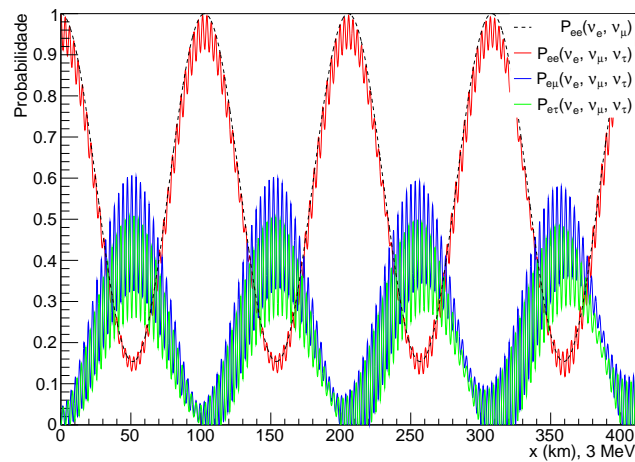


Figura 9 – Visão geral das probabilidades de oscilação dos neutrinos. Linha preta tracejada: oscilação em duas famílias ( $\nu_e, \nu_\mu$ ), linha vermelha: probabilidade de sobrevivência do sabor eletrônico, linha azul probabilidade de transição entre o sabor eletrônico e muônico, linha verde: probabilidade de transição entre o sabor eletrônico e tauônico, todos em três famílias.

Fonte: Do autor.

## 2 ANÁLISE BAYESIANA

Um neutrino não pode ser detectado do mesmo modo como são detectadas outras partículas, como por exemplo, elétrons, fótons, etc, em que os detectores interagem diretamente com essas partículas. No caso dos neutrinos, é necessário buscar por sinais, ou evidências de que ele esteve no detector. O problema é que esses sinais também podem ser gerados por outros fenômenos que ocorrem nas proximidades. Numa situação ambígua como esta, só é possível ter informação relevante se fizermos uso da estatística, mais especificamente da estatística bayesiana, por ser mais indicada para problemas dessa natureza.

O nosso objetivo é distinguir, estatisticamente quais são os sinais que são provenientes de neutrinos e quais provenientes de outras fontes. A estatística necessária para tal análise será apresentada nesse capítulo. Para evitar um delongamento dentro da estatística, não faremos um tratamento auto contido.

### 2.1 Teorema de Bayes

Suponha um espaço amostral  $S$  e dois subconjuntos  $A, B$  desse espaço, com probabilidades  $P(A)$  e  $P(B)$  respectivamente. A probabilidade de ocorrer o evento  $A$ , dado que o evento  $B$  ocorreu é definida como probabilidade condicional  $P(A|B)$ .

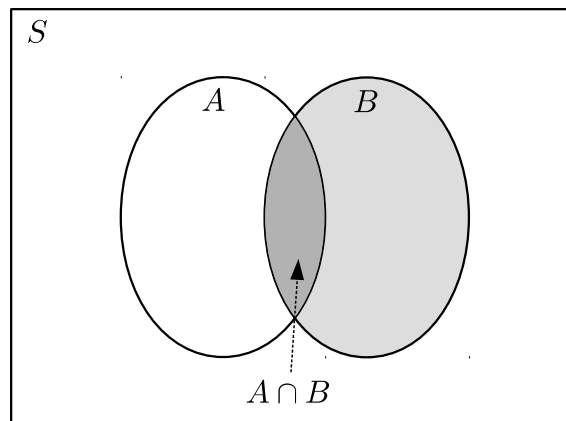


Figura 10 – Probabilidade condicional. A probabilidade  $P(A \cap B)$  compara  $A \cap B$  com  $S$ , enquanto que a probabilidade condicional  $P(A|B)$  compara  $A \cap B$  com  $B$ .

Fonte: Do autor.

Para calcular a probabilidade condicional, podemos pensar que houve um colapso do espaço amostral  $S$  para o subconjunto  $B$ . Então, a interseção  $A \cap B$  deve ser considerada em relação ao subconjunto  $B$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (2.1)$$

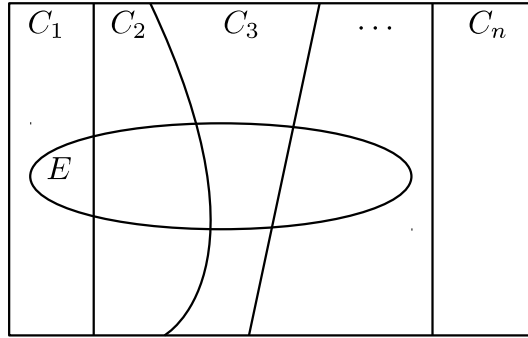


Figura 11 – Partição de um espaço amostral  $S$ . Os subconjuntos  $C_i$  são disjuntos entre si, e sua união resulta em  $S$ .

Fonte: Do autor.

Ao invés de consideramos dois subconjuntos do espaço amostral, vamos supor  $n$  subconjuntos  $C_i, i = 1, 2, \dots, n$  com a seguinte condição

$$\begin{aligned} C_i \cap C_j &= \emptyset, \text{ se } i \neq j, \\ \bigcup_{i=1}^n C_i &= S. \end{aligned} \quad (2.2)$$

Esse conjunto de eventos  $\{C_i\}$  constitui uma *partição* do espaço amostral. A probabilidade de ocorrer o evento  $E$ , dado que  $C_j$  ocorreu, é

$$P(E|C_j) = \frac{P(E \cap C_j)}{P(C_j)}. \quad (2.3)$$

Considere agora a situação inversa. Qual é a probabilidade de  $E$  ter ocorrido a partir de  $C_j$ ?

$$\begin{aligned} P(C_j|E) &= \frac{P(C_j \cap E)}{P(E)} \\ &= \frac{P(E \cap C_j)}{P(E)} \\ &= \frac{P(E|C_j)P(C_j)}{P(E)}, \end{aligned} \quad (2.4)$$

portanto

$$P(C_j|E) = \frac{P(E|C_j)P(C_j)}{P(E)}. \quad (2.4)$$

**Teorema de Bayes** Seja  $C_1, C_2, \dots, C_n$  uma partição de um espaço amostral  $S$ . Suponha que para um evento  $E$  sejam conhecidas as probabilidades  $P(E|C_i)$ . Então, para qualquer  $j$

$$\begin{aligned} P(C_j|E) &= \frac{P(E|C_j)P(C_j)}{P(E)} \\ &= \frac{P(E|C_j)P(C_j)}{\sum_{i=1}^n P(E|C_i)P(C_i)}. \end{aligned} \quad (2.5)$$

O teorema de Bayes é um modo de transformar o problema mais difícil, que é o de calcular a probabilidade  $P(C_j|E)$ , em outro problema mais fácil: calcular as probabilidades  $P(E|C_j)$ , supondo, é claro, que conhecemos também as probabilidades  $P(C_i)$ . Por exemplo, falando de modo geral, um efeito  $E$  pode ter várias causas  $C_j$ . Mas as causas podem ser controladas e os efeitos observados. Com isso podemos calcular todas as probabilidades  $P(E|C_j)$ . Pensando de outra maneira, o teorema de Bayes é um modo sistemático que permite aproveitar o conhecimento prévio que temos sobre um sistema, e calcular outras probabilidades que de fato nos interessam. Convém destacar o seguinte:

- A probabilidade  $P(C_j)$  representa um conhecimento prévio sobre o sistema (espaço amostral), é denominada como *probabilidade a priori*.
- A probabilidade  $P(E|C_j)$  é resultante da aplicação de um modelo teórico que determina o efeito  $E$  a partir da causa  $C_j$ , é chamada de *likelihood* ou *verossimilhança*.
- A probabilidade  $P(C_j|E)$ , que é determinada como resultado da nossa análise, é a *probabilidade posteriori*.

Ao invés de pensarmos em causa  $C_j$  e efeito  $E$ , podemos pensar em hipóteses científicas  $H$  sobre um sistema e  $\{\text{dados}\}$  observados em experimentos com o sistema, tendo como objetivo a verificação se tais hipóteses são ou não verdadeiras. Essa abordagem, feita sobre a luz do teorema de Bayes é conhecida como *análise bayesiana*.

## 2.2 O problema da honestidade de uma moeda

Um ponto importante na análise bayesiana, é a determinação de parâmetros que descrevem distribuições de probabilidades. Vamos começar com um problema bem simples, sem importância prática, mas que pode nos ajudar com os conceitos que precisamos entender.

Suponha uma moeda, tomada como amostra em um cassino. Queremos saber se ela é honesta. Poderíamos pensar numa análise em laboratório para descobrir se a moeda apresenta algum desvio de simetria. Ou podemos fazer um estudo estatístico dos resultados produzidos por vários lançamentos desta moeda.

- *problema fácil*: determinar a probabilidade da moeda em produzir um determinado resultado em  $N$  lançamentos, dado que seu vício ou honestidade sejam conhecidos ou supostos.
- *problema difícil*: determinar a honestidade da moeda, dado que houve um resultado experimental decorrente de  $N$  lançamentos.

Para relacionarmos esses dois problemas, primeiro temos que definir um parâmetro que quantifique a honestidade da moeda, isto é, uma medida numérica que represente o quanto

acreditamos que a moeda seja honesta ou não. Vamos definir um parâmetro  $H$ . Para uma moeda honesta, é natural atribuímos o valor  $H = 0,5$ . No caso extremo, a moeda poderia ter duas coroas ou duas caras, os valores de  $H$  seriam 0 e 1, respectivamente. De um modo geral, o parâmetro  $H$  é uma variável contínua com valores no intervalo  $[0,1]$ . O teorema de Bayes para esse problema pode ser escrito da seguinte forma

$$P(H|\{\text{dados}\}) = \frac{P(\{\text{dados}\}|H)P(H)}{P(\{\text{dados}\})}. \quad (2.6)$$

Na equação (2.6),  $P(H|\{\text{dados}\})$  representa a probabilidade de que um certo valor de  $H$  quantifique a honestidade da moeda, dado que o resultado  $\{\text{dados}\}$  foi observado. Por outro lado,  $P(\{\text{dados}\})$  representa probabilidade total do resultado observado, considerando que ele pode ocorrer para todos os valores possíveis de  $H$ , portanto

$$P(\{\text{dados}\}) = \int_0^1 P(\{\text{dados}\}|H)P(H)dH. \quad (2.7)$$

Uma vez feita a integral, o resultado é uma constante que não depende mais de  $H$ . Por isso podemos retirar  $P(\{\text{dados}\})$  do denominador da equação (2.6) e recompensar depois quando  $P(H|\{\text{dados}\})$  for normalizado. Assim, podemos escrever

$$P(H|\{\text{dados}\}) \propto P(\{\text{dados}\}|H)P(H). \quad (2.8)$$

A equação (2.8) possui o termo  $P(H)$ , que se fosse conhecido, não haveria necessidade de qualquer tipo de inferência. Toda a informação sobre a honestidade da moeda estaria disponível através dele. Para prosseguirmos com a análise, precisamos entrar com alguma informação a respeito de  $P(H)$ . Nesse momento ocorre um fato surpreendente, indicado por simulações: quando o tamanho da amostra é grande o suficiente,  $P(H|\{\text{dados}\})$  é independente da forma funcional de  $P(H)$ .

Vamos considerar primeiro os gráficos com linhas vermelhas da figura 12. No quadro 1, temos a probabilidade uniforme  $P(H)$ . Ela é a estimativa de  $P(H|\{\text{dados}\})$  antes de qualquer resultado experimental. Esta primeira estimativa nos diz que temos a mesma probabilidade de que a moeda seja honesta, ou que tenha duas caras ou duas coras. Na verdade, todas as possibilidades são igualmente prováveis.

No segundo quadro, quando houve o resultado do primeiro lançamento, nossa estimativa para  $P(H|\{\text{dados}\})$  ganha a primeira melhoria: a moeda não pode mais ter duas coroas, ela precisa ter uma cara. Mas, ainda existe a chance da moeda ter duas caras, e por isso o gráfico de  $P(H|\{\text{dados}\})$  não se anula em  $H = 1$ .

No terceiro quadro, a probabilidade da moeda ter duas caras aumenta, por que o resultado de dois lançamentos foram duas caras, ao mesmo tempo que a probabilidade dessa moeda, mostrar uma coroa diminui. Se esse resultado prevalecer nos lançamentos seguintes, a linha vermelha vai se aproximar cada vez mais do eixo  $x$  e da reta  $y = 1$ . Só que essa situação se modifica no quarto quadro, quando aparece a primeira coroa. A moeda

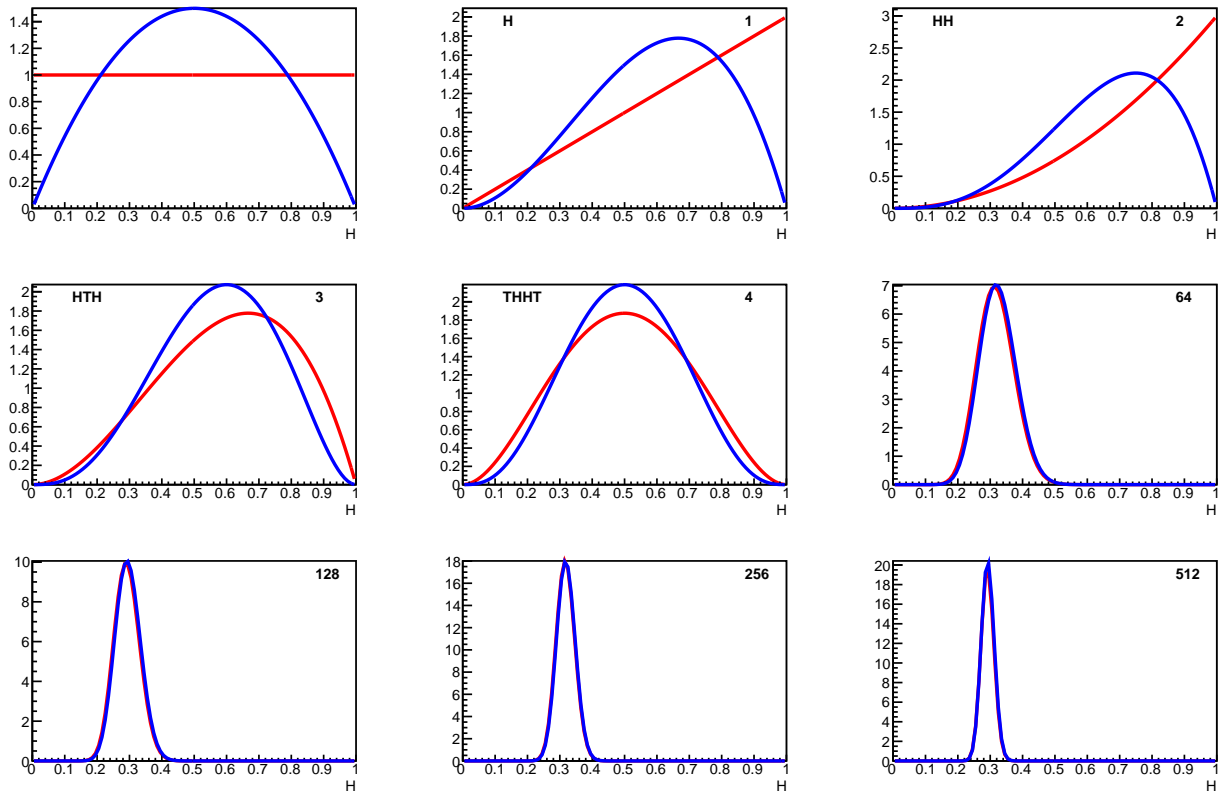


Figura 12 – Simulação do lançamento de uma moeda com vício  $H = 0,3$ . Primeiro quadro: duas possíveis probabilidades a priori  $P(H)$ . Segundo quadro em diante: simulações decorrentes da escolha de  $P(H)$ . O número na parte superior direita é o número de lançamentos.

Fonte: Do autor.

que não pode ter duas coroas, a partir desse momento, não pode também ter duas caras. A medida que os lançamentos prosseguem, e novos dados são obtidos, a curva vermelha procura se acomodar cada vez mais próxima do valor verdadeiro de  $H$ .

A curva azul começa com uma preposição mais realística da situação. De início já se sabia que a moeda tinha cara e coroa e que provavelmente era uma moeda justa. Talvez devido uma boa reputação do Cassino de onde ela veio. Apesar dessa preposição ser bem diferente do que a primeira, as duas convergem para o mesmo resultado  $P(H|\{\text{dados}\})$ , quando o número dados presentes na análise é  $N \gg 1$ . Essa simulação serve para nos trazer a seguinte segurança, quando não tivermos uma boa ideia sobre a probabilidade a priori  $P(H)$ , podemos utilizar uma probabilidade uniforme. Isto trás uma importante simplificação na equação (2.8), que se torna em

$$P(H|\{\text{dados}\}) \propto P(\{\text{dados}\}|H). \quad (2.9)$$

**O princípio de máxima verossimilhança** está baseado na ideia de que o resultado experimental que é observado é o que tem a maior probabilidade de ocorrer no momento que o experimento é realizado. Analisando a equação (2.9) a luz desse princípio

concluimos que a melhor estimativa de  $H$  é a que faz com que o  $P(H|\{\text{dados}\})$  seja máximo.

Vamos generalizar a equação (2.9), supondo que  $X$  seja o nosso parâmetro de interesse, então a probabilidade posteriori  $P \equiv P(X|\{\text{dados}\})$  é

$$P \propto P(\{\text{dados}\}|X). \quad (2.10)$$

Pelo princípio da máxima verossimilhança devemos ter

$$\left. \frac{dP}{dX} \right|_{X_0} = 0. \quad (2.11)$$

Onde  $X_0$  é a melhor estimativa do parâmetro  $X$ . Além do valor de  $X_0$ , precisamos também de uma margem de erro. Como a própria simulação nos mostrou, a função  $P$  tende a se concentrar próxima de  $X_0$ , em forma de pico. Assim, podemos usar a largura do pico como medida da margem de erro. Para avaliarmos a função  $P$  próxima do pico em  $X_0$  vamos expressar  $P$  como uma série de Taylor. Como o logaritmo de uma função tende a ser mais suave que a própria função e além disso, não altera a posição de seus extremos, vamos expandir o logaritmo de  $P$

$$\begin{aligned} L &= \ln(P) \\ &= L(X_0) + \left. \frac{dL}{dX} \right|_{X_0} (X - X_0) + \frac{1}{2} \left. \frac{d^2L}{dX^2} \right|_{X_0} (X - X_0)^2 + \dots \\ &= L(X_0) + \frac{1}{2} \left. \frac{d^2L}{dX^2} \right|_{X_0} (X - X_0)^2 + \dots \end{aligned} \quad (2.12)$$

Escolhendo a aproximação até segunda ordem,

$$L \approx L(X_0) + \frac{1}{2} \left. \frac{d^2L}{dX^2} \right|_{X_0} (X - X_0)^2. \quad (2.13)$$

Tomando a exponencial da equação (2.13)

$$P = A \exp \left[ \frac{1}{2} \left. \frac{d^2L}{dX^2} \right|_{X_0} (X - X_0)^2 \right]. \quad (2.14)$$

A constante  $A$  é para normalização de  $P$ . Sabemos que a derivada segunda de  $L$  é negativa em  $X_0$ , logo equação (2.14) representa uma distribuição de probabilidade normal. Portanto concluimos que, a probabilidade posteriori  $P = P(X|\{\text{dados}\})$  é uma distribuição aproximadamente normal, e podemos estimar a confiança do valor estimado  $X_0$  como sendo

$$\sigma_{X_0} = \left( - \left. \frac{d^2L}{dX^2} \right|_{X_0} \right)^{-1/2}. \quad (2.15)$$

Com esses parâmetros determinados, podemos expressar nossa estimativa como sendo

$$X = X_0 \pm \sigma_{X_0}. \quad (2.16)$$

Vamos então aplicar o resultado ao problema da moeda.

$$\begin{aligned}
P &= P(\{\text{dados}\}|H) \\
&\propto H^R(1-H)^{(N-R)} \\
L &= L_0 + \ln \left[ H^R(1-H)^{(N-R)} \right] \\
\frac{dL}{dH} &= \frac{R}{H} - \frac{N-R}{1-H} \\
\frac{d^2L}{dH^2} &= -\frac{R}{H^2} - \frac{N-R}{(1-H)^2} \\
\left. \frac{dL}{dH} \right|_{H_0} &= \frac{R}{H_0} - \frac{N-R}{1-H_0} = 0 \Rightarrow H_0 = \frac{R}{N} \\
\left. \frac{d^2L}{dH^2} \right|_{H_0} &= -\frac{R(1-H_0)^2 + (N-R)H_0^2}{H_0^2(1-H_0)^2} \\
&= -\frac{N}{H_0(1-H_0)}. \tag{2.17}
\end{aligned}$$

Resumindo, se em  $N$  lançamentos, uma moeda apresenta  $R$  caras, a honestidade da moeda deve ser estimada como

$$H_0 = \frac{R}{N} \pm \sqrt{\frac{H_0(1-H_0)}{N}}. \tag{2.18}$$

Vamos dar um exemplo numérico. Usando o ROOT(12) para fazer uma simulação, uma moeda honesta foi lançada 1 milhão de vezes. O resultado foi  $R = 499\,723$  caras. Com esses valores, obtemos que  $H_0 = 0,49972 \pm 0,00050$ .

### 2.3 Estimativa com dois parâmetros - ruído gaussiano

O caso tratado na seção 2.2 envolve a estimação de um parâmetro, o viés  $H$  da moeda. Vamos agora considerar a estimação de dois parâmetros. Considere um experimento que fornece um conjunto de dados  $\{x_k\}$ ,  $k = 1, 2, \dots, N$ , que são descritos por uma distribuição normal, com média  $\mu$  e desvio padrão  $\sigma$ , os quais pretendemos estimar a partir de dados experimentais.

Pelo teorema de Bayes, podemos escrever

$$P(\mu, \sigma | \{x_k\}) \propto P(\{x_k\} | \mu, \sigma) P(\mu, \sigma). \tag{2.19}$$

Vamos usar a estimativa a priori de  $P(\mu, \sigma) = \text{constante}$ , então

$$P(\mu, \sigma | \{x_k\}) \propto P(\{x_k\} | \mu, \sigma). \tag{2.20}$$

Considerando cada resultado de  $\{x_k\}$  independente dos demais

$$P(\mu, \sigma | \{x_k\}) \propto \prod_{k=1}^N (\sigma\sqrt{2\pi})^{-N} \exp \left[ -\frac{1}{2\sigma^2} (x_k - \mu)^2 \right]$$



$$\propto \sigma^{-N} \exp \left[ -\frac{1}{2\sigma^2} \sum_{k=1}^N (x_k - \mu)^2 \right]. \quad (2.21)$$

Calculando a probabilidade marginal

$$P(\mu|\{x_k\}) = \int_0^{+\infty} \sigma^{-N} \exp \left[ -\frac{1}{2\sigma^2} \sum_{k=1}^N (x_k - \mu)^2 \right] d\sigma. \quad (2.22)$$

Fazendo a mudança de variável  $\sigma = 1/t$  e lembrando que a somatória na exponencial é uma constante, que vamos designá-la por  $a$ . Com isso, temos

$$P(\mu|\{x_k\}) = \int_0^{+\infty} t^{N-2} e^{-\frac{1}{2}at^2} dt. \quad (2.23)$$

Essa integral pode ser resolvida por comparação com a definição da função gama

$$\Gamma(n) = \int_0^{+\infty} x^{n-1} e^{-x} dx. \quad (2.24)$$

Fazendo essa comparação

$$\int_0^{+\infty} t^{N-2} e^{-\frac{1}{2}at^2} dt = 2^{\frac{N-3}{2}} \Gamma\left(\frac{N-1}{2}\right) a^{-\frac{N-1}{2}}. \quad (2.25)$$

Com esse resultado, a probabilidade marginal (2.23) fica como

$$P(\mu|\{x_k\}) \propto \left[ \sum_{k=1}^N (x_k - \mu)^2 \right]^{-\frac{N-1}{2}}. \quad (2.26)$$

A aproximação por série de Taylor, para o cálculo da função  $L$

$$L \approx L_0 + \frac{1}{2!} \frac{d^2 L}{d\mu^2} \Big|_{\mu_0} (\mu - \mu_0)^2. \quad (2.27)$$

Onde  $L$  é definido como

$$L = \ln P(\mu|\{x_k\}) \approx L_0 - \frac{N-1}{2} \ln \left[ \sum_{k=1}^N (x_k - \mu)^2 \right]. \quad (2.28)$$

A derivada primeira de  $L$ , em relação a  $\mu$

$$\frac{dL}{d\mu} = \frac{(N-1) \sum_k (x_k - \mu)}{\sum_k (x_k - \mu)^2} \quad (2.29)$$

A derivada segunda, em relação a  $\mu$

$$\frac{d^2 L}{d\mu^2} = \frac{-N(N-1) \sum_k (x_k - \mu)^2 + 2(N-1) \left[ \sum_k (x_k - \mu) \right]^2}{\left[ \sum_k (x_k - \mu)^2 \right]^2}. \quad (2.30)$$

A estimativa  $\mu_0$  é dada por

$$\frac{dL}{d\mu} \Big|_{\mu_0} = 0 \Rightarrow \sum_k (x_k - \mu) = 0, \quad (2.31)$$

então

$$\mu_0 = \frac{1}{N} \sum_{k=1}^N x_k. \quad (2.32)$$

Substituindo essa estimativa na derivada segunda (2.30), obtemos

$$\frac{d^2 L}{d\mu^2} = -\frac{N(N-1)}{\sum_k (x_k - \mu_0)^2}. \quad (2.33)$$

Usando a equação (2.15)

$$\left( -\frac{d^2 L}{d\mu^2} \Big|_{\mu_0} \right)^{-1} = \frac{1}{N} \frac{1}{N-1} \sum_{k=1}^N (x_k - \mu_0)^2. \quad (2.34)$$

Definindo

$$S^2 = \frac{1}{N-1} \sum_{k=1}^N (x_k - \mu_0)^2, \quad (2.35)$$

temos

$$\left( -\frac{d^2 L}{d\mu^2} \Big|_{\mu_0} \right)^{-\frac{1}{2}} = \frac{S}{\sqrt{N}}. \quad (2.36)$$

Desse modo, podemos expressar a estimativa da média como sendo

$$\mu = \mu_0 \pm \frac{S}{\sqrt{N}}. \quad (2.37)$$

No caso particular do ruído gaussiano, podemos dispensar a aproximação, e usar diretamente a equação (2.26) que permite um desenvolvimento analítico. Vamos definir as seguintes quantidades baseadas nos dados experimentais

$$\bar{x} = \mu_0 = \frac{\sum_{k=1}^N x_k}{N} \quad (2.38)$$

$$V = \sum_{k=1}^N (x_k - \bar{x})^2. \quad (2.39)$$

Em termos de  $V$  e  $\bar{x}$ , a equação (2.26) pode ser escrita como

$$P(\mu|\{x_k\}) \propto [N(\bar{x} - \mu)^2 + V]^{-\frac{N-1}{2}}. \quad (2.40)$$

Esta expressão obedece a distribuição de Student ou distribuição t com  $N - 2$  graus de liberdade. A variável  $t$ , como normalmente é conhecida, pode ser obtida através de

$$t^2 = \frac{n(n+2)(\bar{x} - \mu)^2}{V}, \quad (2.41)$$

onde  $n = N - 2$  é o grau de liberdade da distribuição. Em termos de  $t$  a equação (2.40) é escrita como

$$P(\mu|\{x_k\}) \propto V^{-\frac{n+1}{2}} \left[ \frac{n(n+2)}{V} \frac{(\bar{x} - \mu)^2}{n} + 1 \right]^{-\frac{n+1}{2}}$$

$$\propto \left( \frac{t^2}{n} + 1 \right)^{-\frac{n+1}{2}}, \quad (2.42)$$

que é a forma convencional da distribuição  $t$  com  $n = N - 2$  graus de liberdade. Para estimar o segundo parâmetro  $\sigma$ , devemos substituir as quantidades  $\bar{x}$  (2.38) e  $V$  (2.39) em (2.21) e então calcular a probabilidade marginal  $P(\sigma|\{x_k\})$

$$\begin{aligned} P(\mu, \sigma|\{x_k\}) &\propto \sigma^{-N} \exp\left(-\frac{V}{2\sigma^2}\right) \exp\left[-\frac{N(\bar{x} - \mu)^2}{2\sigma^2}\right] \\ \int_{-\infty}^{+\infty} P(\mu, \sigma|\{x_k\}) d\mu &\propto \sigma^{-N} \exp\left(-\frac{V}{2\sigma^2}\right) \int_{-\infty}^{+\infty} \exp\left[-\frac{N(\bar{x} - \mu)^2}{2\sigma^2}\right] d\mu \\ P(\sigma|\{x_k\}) &\propto \sigma^{-N} \exp\left(-\frac{V}{2\sigma^2}\right) \left(\frac{\pi}{2\sigma^2}\right)^{\frac{1}{2}} \\ &\propto \sigma^{1-N} \exp\left(-\frac{V}{2\sigma^2}\right). \end{aligned} \quad (2.43)$$

Se definirmos a variável  $x = V/\sigma^2$  e fizermos todo o procedimento de normalização, vamos obter

$$P(x|\{x_k\}) = \frac{2x^{3/2}}{V^{1/2}} \chi_{(N-2)}^2(x). \quad (2.44)$$

Os resultados que obtemos com (2.42) e (2.44), nos dão as distribuições exatas para a média e para o desvio do ruído gaussiano. Para fazer a estimação, é bem mais prático continuar com as aproximações. Vamos calcular então o  $L$  e suas derivadas

$$\begin{aligned} L &= \ln P(\sigma|\{x_k\}) \\ &= L_0 + (1 - N) \ln \sigma - \frac{V}{2\sigma^2} \end{aligned} \quad (2.45)$$

$$\frac{dL}{d\sigma} = \frac{1 - N}{\sigma} + \frac{V}{\sigma^3} \quad (2.46)$$

$$\frac{d^2L}{d\sigma^2} = -\frac{1 - N}{\sigma^2} - \frac{3V}{\sigma^4}. \quad (2.47)$$

Agora vamos calcular as estimativas

$$\left. \frac{dL}{d\sigma} \right|_{\sigma_0} = 0, \quad \sigma_0 = \sqrt{\frac{V}{N-1}} \quad (2.48)$$

$$\left. \frac{d^2L}{d\sigma^2} \right|_{\sigma_0} = -\frac{2(N-1)}{\sigma_0^2} \quad (2.49)$$

$$\left( -\left. \frac{d^2L}{d\sigma^2} \right|_{\sigma_0} \right)^{-1/2} = \frac{\sigma_0}{\sqrt{2(N-1)}}. \quad (2.50)$$

Portanto, a estimativa de  $\sigma$  fica como

$$\sigma = \sigma_0 \pm \frac{\sigma_0}{\sqrt{2(N-1)}}. \quad (2.51)$$

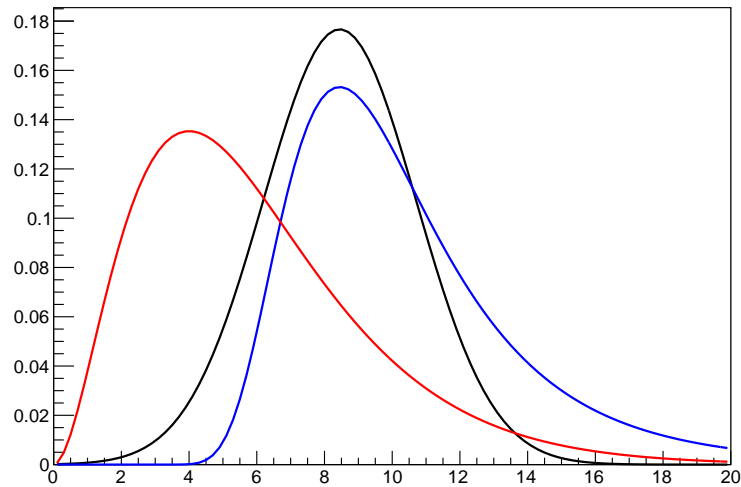


Figura 13 – Simulação de um ruído gaussiano com  $N = 8$  dados e 6 graus de liberdade. Os parâmetros usados na simulação foram  $\mu = 40$ ,  $\sigma = 6$ . A curva vermelha é a distribuição  $\chi^2_6$ . A curva azul é a distribuição exata, e a preta é aproximação usando  $L$ .

Fonte: Do autor.

Vamos dar um exemplo numérico. Uma simulação com o ROOT, ajustada com média  $\mu = 40$  e  $\sigma = 6$  gerou o seguinte ruído gaussiano  $\{37,40,43,29,42,47,49,42\}$  (os valores foram arredondados). Com base nesses valores temos as seguintes quantidades

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = 41,125 \quad (2.52a)$$

$$V = \sum_{k=1}^N (x_k - \bar{x})^2 = 266,875 \quad (2.52b)$$

$$S = \sqrt{\frac{V}{N-1}} = 6,175 \quad (2.52c)$$

$$\mu = \bar{x} \pm \frac{S}{\sqrt{N}} = 41,1 \pm 2,2 \quad (2.52d)$$

$$\sigma = S \pm \frac{S}{\sqrt{2(N-1)}} = 6,2 \pm 1,7. \quad (2.52e)$$

## 2.4 Estimativa com dois parâmetros, caso geral

Vamos continuar com dois parâmetros, porém generalizando um pouco mais. Suponha que queiramos determinar os parâmetros  $X, Y$ . Para isso um experimento é realizado e um conjunto  $\{\text{dados}\}$  é extraído. Com base nesses dados devemos inferir os valores e as confianças desses parâmetros.

É importante chamarmos atenção para o fato de que sendo um parâmetro  $X$  contí-

nuo, com valores reais, a função  $P(X)$  representa uma *função densidade de probabilidade*<sup>1</sup>, normalmente abreviado como p.d.f. (do inglês *probability density function*). Seu valor é tal que a probabilidade de  $X$  assumir valor dentro do intervalo  $X - dX < X < X + dX$  é dada por

$$P(X - dX < X < X + dX) = P(X)dX. \quad (2.53)$$

No caso de dois parâmetros, devemos utilizar o conceito de p.d.f. conjunta  $P(X,Y)$ , onde a probabilidade do parâmetro  $X$  assumir valor dentro do intervalo  $X - dX < X < X + dX$  e o parâmetro  $Y$  assumir valor dentro do intervalo  $Y - dY < Y < Y + dY$ , é dada por

$$P(X - dX < X < X + dX, Y - dY < Y < Y + dY) = P(X,Y)dXdY \quad (2.54)$$

A p.d.f. conjunta  $P = P(X,Y|\{\text{dados}\})$  nos dá toda informação que é possível ter mediante um resultado experimental. E o teorema de Bayes transforma esta p.d.f., no produto de outras duas

$$P(X,Y|\{\text{dados}\}) \propto P(\{\text{dados}\}|X,Y)P(X,Y). \quad (2.55)$$

Quando o número de dados é grande o suficiente, a probabilidade priori  $P(X,Y)$  pode ser estimada por uma constante, e por isso é absorvida pelo sinal de proporção. A likelihood  $P(\{\text{dados}\}|X,Y)$ , é uma p.d.f. mais fácil de se obter do que  $P(X,Y|\{\text{dados}\})$ . Logo, o teorema de Bayes nos proporciona mais um avanço rumo a inferência dos parâmetros, que é expresso na equação

$$P(X,Y|\{\text{dados}\}) \propto P(\{\text{dados}\}|X,Y). \quad (2.56)$$

O princípio da máxima verossimilhança afirma que o resultado experimental que ocorre num experimento é o que tinha a maior probabilidade(14) de ocorrer no momento em que ele é realizado. Por isso, a melhor estimativa é o ponto  $\{X_0, Y_0\}$  que maximiza a probabilidade condicional aos dados observados. Em questão de estimar parâmetros, não precisamos da p.d.f.  $P$  em si, mas apenas do ponto de máximo, e de conhecer o quanto a p.d.f. se concentra em torno desse ponto, o que nos dá a confiança da estimação. Isto nos sugere uma aproximação por série de Taylor da p.d.f. posteriori. Levando-se em conta que o logaritmo suaviza a função melhorando a aproximação, vamos expandir  $\ln P$  em vez de  $P$ . Para isso definimos a função  $L$

$$L = \ln P. \quad (2.57)$$

A localização do máximo  $\{X_0, Y_0\}$  de  $L$ , que é também o de  $P$ , é dada pela resolução simultânea das equações

$$\left. \frac{\partial L}{\partial X} \right|_{X_0, Y_0} = 0, \quad \left. \frac{\partial L}{\partial Y} \right|_{X_0, Y_0} = 0. \quad (2.58)$$

<sup>1</sup> Para mais informação consulte(13).

A série de Taylor para uma função de duas variáveis, é escrita como

$$L = L_0 + \frac{1}{2} \left[ \frac{\partial^2 L}{\partial X^2} \Big|_{X_0, Y_0} (X - X_0)^2 + \frac{\partial^2 L}{\partial Y^2} \Big|_{X_0, Y_0} (Y - Y_0)^2 + 2 \frac{\partial^2 L}{\partial X \partial Y} \Big|_{X_0, Y_0} (X - X_0)(Y - Y_0) \right] + \dots \quad (2.59)$$

Lembrando que o termo de primeira ordem da série é nulo, devido a condição (2.58).

Definindo as constantes

$$A = \frac{\partial^2 L}{\partial X^2} \Big|_{X_0, Y_0}, \quad B = \frac{\partial^2 L}{\partial Y^2} \Big|_{X_0, Y_0}, \quad C = \frac{\partial^2 L}{\partial X \partial Y} \Big|_{X_0, Y_0}, \quad (2.60)$$

temos

$$L \approx L_0 + \frac{1}{2} [A(X - X_0)^2 + B(Y - Y_0)^2 + 2C(X - X_0)(Y - Y_0)]. \quad (2.61)$$

A quantidade  $L$  pode ser rescrita em forma matricial como

$$L \approx L_0 + \frac{1}{2} \begin{pmatrix} X - X_0 & Y - Y_0 \end{pmatrix} \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} X - X_0 \\ Y - Y_0 \end{pmatrix}. \quad (2.62)$$

A parte que interessa de  $L$ , podemos chamá-la de  $Q$

$$Q = \begin{pmatrix} X - X_0 & Y - Y_0 \end{pmatrix} \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} X - X_0 \\ Y - Y_0 \end{pmatrix}. \quad (2.63)$$

As curvas de níveis de  $Q$  são curvas onde  $L$ , e portanto de  $P$ , apresenta valores constantes. Essas curvas são elipses cujos semi eixos possuem a mesma direção dos auto vetores da matriz

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix}. \quad (2.64)$$

Em termos de  $Q$ , a probabilidade é aproximada como

$$P = \exp(L) \approx \exp(L_0 + \frac{1}{2}Q) \propto \exp(\frac{1}{2}Q). \quad (2.65)$$

A probabilidade marginal, com respeito a  $X$ , é dada por

$$\begin{aligned} P(X|\{\text{dados}\}) &= \int_{-\infty}^{\infty} P(X, Y|\{\text{dados}\}) dY \\ &\propto \int_{-\infty}^{\infty} \exp(\frac{1}{2}Q) dy \\ &\propto \exp \left[ \frac{1}{2} \left( A - \frac{C^2}{B} \right) x^2 \right] \int_{-\infty}^{\infty} \exp \left[ \frac{1}{2} B \left( y + \frac{2C}{B} \right)^2 \right] dy, \end{aligned} \quad (2.66)$$

onde  $x = X - X_0, y = Y - Y_0$ . No caso de  $B < 0$ , a integral é uma gaussiana e vale

$$\sqrt{\frac{2\pi}{-B}}, \quad (2.67)$$

resultando em

$$P(X|\{\text{dados}\}) \propto \exp \left[ \frac{1}{2} \left( A - \frac{C^2}{B} \right) (X - X_0)^2 \right]. \quad (2.68)$$

Se  $A$  também for negativo e  $AB > C^2$ , o argumento da exponencial acima será negativo, com isso teremos novamente uma gaussiana, com a qual podemos definir, aproximadamente, a confiança relacionada ao valor estimado  $X_0$ , através de

$$\begin{aligned} -\frac{1}{\sigma_X^2} &= A - \frac{C^2}{B} \\ \sigma_X &= \sqrt{\frac{-B}{AB - C^2}} \end{aligned} \quad (2.69)$$

$$P(X|\{\text{dados}\}) \propto \exp \left[ \frac{(X - X_0)^2}{2\sigma_X^2} \right]. \quad (2.70)$$

Seguindo a mesma lógica, agora para a variável  $Y$ , teremos

$$\sigma_Y = \sqrt{\frac{-A}{AB - C^2}} \quad (2.71)$$

$$P(Y|\{\text{dados}\}) \propto \exp \left[ \frac{(Y - Y_0)^2}{2\sigma_Y^2} \right]. \quad (2.72)$$

Os resultados acima também podem ser obtidos pelas integrais

$$\sigma_X^2 = \langle (X - X_0)^2 \rangle = K \iint (X - X_0)^2 \exp \left( \frac{1}{2} Q \right) dX dY \quad (2.73)$$

$$\sigma_Y^2 = \langle (Y - Y_0)^2 \rangle = K \iint (Y - Y_0)^2 \exp \left( \frac{1}{2} Q \right) dX dY, \quad (2.74)$$

onde podemos incluir

$$\sigma_{XY}^2 = \langle (X - X_0)(Y - Y_0) \rangle = K \iint (X - X_0)(Y - Y_0) \exp \left( \frac{1}{2} Q \right) dX dY, \quad (2.75)$$

sendo  $K$  é a constante de normalização dada por

$$K = \frac{\sqrt{AB - C^2}}{2\pi}. \quad (2.76)$$

Agrupando as variâncias numa matriz de covariâncias

$$\begin{pmatrix} \sigma_{X^2} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{Y^2} \end{pmatrix} = \frac{1}{AB - C^2} \begin{pmatrix} -B & C \\ C & -A \end{pmatrix} = - \begin{pmatrix} A & C \\ C & B \end{pmatrix}^{-1}. \quad (2.77)$$

## 2.5 Estimativa com $M$ parâmetros

Suponha que temos um conjunto de  $M$  parâmetros representados pelo vetor  $\mathbf{X}$ , e também um conjunto de  $N$  dados experimentais indicados pelo vetor  $\mathbf{D}$ . Nosso objetivo é determinar a melhor estimativa  $\mathbf{X}_0$  para os parâmetros  $\mathbf{X}$ , com base nos dados  $\mathbf{D}$ . Começemos pelo teorema de Bayes

$$P(\mathbf{X}|\mathbf{D}) \propto P(\mathbf{D}|\mathbf{X})P(\mathbf{X}), \quad (2.78)$$

usando a estimativa de probabilidade a priori constante

$$P(\mathbf{X}) = \text{constante}, \quad (2.79)$$

obtemos

$$P(\mathbf{X}|\mathbf{D}) \propto P(\mathbf{D}|\mathbf{X}). \quad (2.80)$$

Supondo que os resultados individuais do experimento são independentes

$$P(\mathbf{D}|\mathbf{X}) = \prod_{k=1}^N P(D_k|\mathbf{X}). \quad (2.81)$$

A probabilidade posteriori fica então como

$$P(\mathbf{X}|\mathbf{D}) \propto \prod_{k=1}^N P(D_k|\mathbf{X}). \quad (2.82)$$

Definimos a função  $L$  como

$$L = \ln P(\mathbf{X}|\mathbf{D}) \quad (2.83a)$$

$$\propto \ln P(\mathbf{D}|\mathbf{X}), \quad (2.83b)$$

a melhor estimativa  $\mathbf{X}_0$  é a solução de um conjunto de  $M$  equações simultâneas, indicadas vetorialmente como

$$\nabla L = 0. \quad (2.84)$$

Expandindo  $L$  próxima ao ponto  $\mathbf{X}_0$

$$L(\mathbf{X}) \approx L(\mathbf{X}_0) + \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \frac{\partial^2 L}{\partial X_i \partial X_j} \Big|_{\mathbf{X}_0} (X_i - X_{0i})(X_j - X_{0j}). \quad (2.85)$$

Definindo a matriz  $M \times M$

$$(\nabla \nabla L)_{ij} = \frac{\partial^2 L}{\partial X_i \partial X_j} \Big|_{\mathbf{X}_0}, \quad (2.86)$$

que é a matriz equivalente a (2.64) para o caso de dois parâmetros. A expansão de  $L$  é escrita em forma matricial

$$L(\mathbf{X}) \approx L(\mathbf{X}_0) + \frac{1}{2} (\mathbf{X} - \mathbf{X}_0)^T \nabla \nabla L (\mathbf{X} - \mathbf{X}_0). \quad (2.87)$$

A densidade de probabilidade posteriori conjunta dos parâmetros é dada por

$$P(\mathbf{X}|\mathbf{D}) \propto \exp \left[ \frac{1}{2} (\mathbf{X} - \mathbf{X}_0)^T \nabla \nabla L (\mathbf{X} - \mathbf{X}_0) \right]. \quad (2.88)$$

E finalmente a covariância entre dois parâmetros é dada por

$$\langle (X_i - X_{0i})(X_j - X_{0j}) \rangle = \left[ -(\nabla \nabla L)^{-1} \right]_{ij}. \quad (2.89)$$



## 2.6 Método dos mínimos quadrados

Uma solução analítica do método descrito na seção 2.5, depende de sermos capazes de resolver as equações propostas naquela seção, em particular, a (2.84). Vamos considerar então algumas simplificações visando aplicação analítica do método, mas, que ainda o mantenha útil em situações práticas. Como primeiro passo, vamos supor que o experimento possua um ruído gaussiano, de modo que

$$P(D_k|\mathbf{X}) = \frac{1}{\sigma_k \sqrt{2\pi}} \exp \left[ -\frac{(F_k - D_k)^2}{2\sigma_k^2} \right], \quad (2.90)$$

onde  $F_k$  representa o valor ideal do dado experimental, em relação a um modelo que expresse os dados em função dos parâmetros, representado através de uma função

$$F_k = f_k(\mathbf{X}), \quad (2.91)$$

onde  $k = 1, 2, \dots, N$ . Com isto, a probabilidade

$$\begin{aligned} P(\mathbf{D}|\mathbf{X}) &= \prod_{k=1}^N \frac{1}{\sigma_k \sqrt{2\pi}} \exp \left[ -\frac{(F_k - D_k)^2}{2\sigma_k^2} \right] \\ &= \frac{1}{\sigma_k^N (2\pi)^{N/2}} \exp \left[ -\frac{1}{2} \sum_{k=1}^N \left( \frac{F_k - D_k}{\sigma_k} \right)^2 \right] \\ &\propto \exp \left( -\frac{1}{2} \chi^2 \right), \end{aligned} \quad (2.92)$$

onde

$$\chi^2 = \sum_{k=1}^N \left( \frac{F_k - D_k}{\sigma_k} \right)^2. \quad (2.93)$$

Calculando a função

$$\begin{aligned} L &= \ln P(\mathbf{X}|\mathbf{D}) \\ &\propto \ln P(\mathbf{D}|\mathbf{X}) \\ &= \text{constante} - \frac{1}{2} \chi^2. \end{aligned} \quad (2.94)$$

Assim, o máximo de  $L$  é obtido pelo mínimo das soma de quadrados dados por  $\chi^2$ , e esta é a razão do termo *método dos mínimos quadrados*. O fato de termos assumido um ruído gaussiano faz com expansão em série de Taylor para  $L$  seja exata. Isto se deve ao fato de todas as derivadas de ordem superior a dois, serem nulas na expansão. Nossa próxima simplificação é quanto ao modelo  $F_k$ . Se supormos que ela seja uma relação linear, por exemplo

$$F = T\mathbf{X} + \mathbf{C}, \quad (2.95)$$

onde  $\mathbf{C}$  é um vetor constante qualquer e  $T$  a matriz de coeficientes. O gradiente  $L$  dado por

$$\nabla L = -\frac{1}{2} \nabla \chi^2, \quad (2.96)$$

assume uma forma também linear, como

$$\nabla L = Q\mathbf{X} + \mathbf{P}, \quad (2.97)$$

onde  $\mathbf{P}$  é um vetor constante qualquer e  $Q$  a matriz de coeficientes. Um modo fácil de vermos que  $\nabla L$  é linear, é diferenciá-lo uma vez e verificar que o resultado é uma constante. Primeiro devemos obter  $\nabla L$  a partir de uma relação linear de  $F$

$$\begin{aligned} (\nabla L)_j &= -\frac{1}{2} \frac{\partial \chi^2}{\partial X_j} \\ &= -\frac{1}{2} \sum_{k=1}^N 2 \left( \frac{F_k - D_k}{\sigma_k} \right) \frac{1}{\sigma_k} \frac{\partial F_k}{\partial X_j} \\ &= -\sum_{k=1}^N \frac{F_k - D_k}{\sigma_k^2} T_{kj}, \end{aligned} \quad (2.98)$$

onde usamos

$$\frac{\partial F_k}{\partial X_j} = \frac{\partial}{\partial X_j} (T\mathbf{X} + \mathbf{C}) = T_{kj}. \quad (2.99)$$

Agora derivamos  $(\nabla L)_j$

$$\begin{aligned} \frac{\partial}{\partial X_i} (\nabla L)_j &= (\nabla \nabla L)_{ij} \\ &= -\frac{\partial}{\partial X_i} \sum_{k=1}^N \frac{F_k - D_k}{\sigma_k^2} T_{kj} (\nabla \nabla L)_{ij} \\ &= -\sum_{k=1}^N \frac{T_{ki} T_{kj}}{\sigma_k^2}, \end{aligned} \quad (2.100)$$

que é uma constante.

É importante entendermos que o modelo linear  $F$  dado por (2.95), se refere a relação entre os parâmetros e os dados, e não entre a dependência de  $Y_k$  com  $x_k$ . Por exemplo, dado um conjunto experimental  $\{Y_k, \sigma_k, x_k\}$ , podemos usar o modelo linear da equação (2.95), e uma equação polinomial de grau  $M - 1$  entre  $Y_k$  e  $x_k$ . Nesse caso,  $F$  é dado por

$$\begin{pmatrix} F_1 \\ \vdots \\ F_N \end{pmatrix} = \begin{pmatrix} (x_1)^0 & \cdots & (x_1)^{M-1} \\ \vdots & \ddots & \vdots \\ (x_N)^0 & \cdots & (x_N)^{M-1} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_M \end{pmatrix}. \quad (2.101)$$

Para ajustes polinomiais, o vetor  $\mathbf{C}$  de (2.95) é nulo. O vetor  $\nabla L$  é dado por

$$\nabla L = - \begin{pmatrix} \sum_{k=1}^N \frac{x_k^0}{\sigma_k^2} & \cdots & \sum_{k=1}^N \frac{x_k^{M-1}}{\sigma_k^2} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^N \frac{x_k^{M-1}}{\sigma_k^2} & \cdots & \sum_{k=1}^N \frac{x_k^{2M-2}}{\sigma_k^2} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_M \end{pmatrix} + \begin{pmatrix} \sum_{k=1}^N \frac{x_k^0 Y_k}{\sigma_k^2} \\ \vdots \\ \sum_{k=1}^N \frac{x_k^{M-1} Y_k}{\sigma_k^2} \end{pmatrix}. \quad (2.102)$$

E a matriz de covariância dada por

$$-(\nabla\nabla L)^{-1} = \begin{pmatrix} \sum_{k=1}^N \frac{x_k^0}{\sigma_k^2} & \cdots & \sum_{k=1}^N \frac{x_k^{M-1}}{\sigma_k^2} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^N \frac{x_k^{M-1}}{\sigma_k^2} & \cdots & \sum_{k=1}^N \frac{x_k^{2M-2}}{\sigma_k^2} \end{pmatrix}^{-1}. \quad (2.103)$$

Exemplo: fizemos uma simulação usando ROOT para gerar um conjunto de dados, com ruído gaussiano

$$Y_k = \{-4,649001, -0,332987, 4,781061, 7,92813, 8,327498, \\ 12,736236, 10,653006, 8,880925, 10,702428\} \quad (2.104)$$

baseado numa função do terceiro grau, dada por  $f(x) = 4,4 + 3,6x - 0,75x^2 + 0,05x^3$ . Os valores usados para a simulação foram

$$\sigma_k = \{1; 1; 1; 1,5; 1,5; 1,5; 1; 1; 1\} \quad (2.105)$$

$$x_k = \{-2; -1; 0; 1; 2; 3; 4; 6; 8\} \quad (2.106)$$

$$X_k = \{4,4; 3,6; -0,75; 0,05\}, \quad (2.107)$$

A equação para obter a melhor estimativa é dada por  $\nabla L = 0$ , numericamente temos

$$\begin{pmatrix} -7,3 & -17,6 & -127,2 & -799,0 \\ -17,6 & -127,2 & -799,0 & -5708,5 \\ -127,2 & -799,0 & -5708,5 & -41657,6 \\ -799,0 & -5708,5 & -41657,6 & -313313,8 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = \begin{pmatrix} -42,920707 \\ -219,055476 \\ -1225,460828 \\ -8303,207877 \end{pmatrix}, \quad (2.108)$$

cuja solução é

$$\mathbf{X}_0 = \begin{pmatrix} 4,858916 \\ 3,485780 \\ -0,706219 \\ 0,044497 \end{pmatrix}. \quad (2.109)$$

A matriz de covariância é dada por

$$-(\nabla\nabla L)^{-1} = \begin{pmatrix} 0,406882 & 0,054129 & -0,063036 & 0,006357 \\ 0,054129 & 0,079712 & -0,025465 & 0,001795 \\ -0,063036 & -0,025465 & 0,019677 & -0,001991 \\ 0,006357 & 0,001795 & -0,001991 & 0,000219 \end{pmatrix}, \quad (2.110)$$

onde raiz quadrada dos elementos da diagonal principal ( $i,i$ ) fornece a margem de erro do parâmetro  $i$ . Podemos então expressar nossa melhor estimativa e suas margens de erros, conforme tabela 1

Tabela 1 – Melhor estimativa comparada com o valor real utilizado na simulação.

parâmetro	melhor estimativa	valor real
$X_1$	$4,86 \pm 0,64$	4,4
$X_2$	$3,49 \pm 0,28$	3,6
$X_3$	$-0,71 \pm 0,14$	-0,75
$X_4$	$0,044 \pm 0,015$	0,05

Fonte: Do autor.

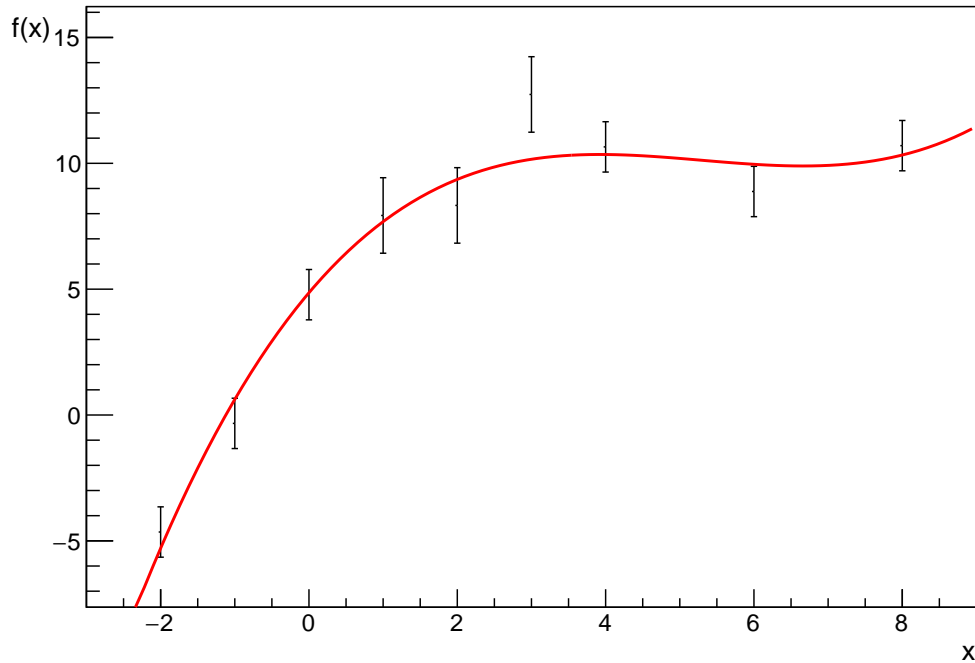


Figura 14 – Simulação com o ROOT gerou um conjunto de nove valores para  $Y_k$ . A função usada para fazer a simulação foi  $f(x) = 4,4 + 3,6x - 0,75x^2 + 0,05x^3$ .

Fonte: Do autor.

Para resolver a equação (2.108) usamos o ROOT que é um software que oferece muitas funcionalidades para tratar com grande volumes de dados. Também podemos usar a biblioteca c++ chamada BAT, que dispõe de classes que fazem a estimação desses parâmetros utilizando análise bayesiana. A tabela 2 mostra o resultado fornecido pelo BAT para o mesmo problema e o compara com o resultado do ROOT.

Tabela 2 – Comparando soluções do BAT e ROOT.

Parâmetro	BAT	ROOT
$X_1$	$4,86 \pm 0,65$	$4,86 \pm 0,64$
$X_2$	$3,49 \pm 0,28$	$3,49 \pm 0,28$
$X_3$	$-0,71 \pm 0,14$	$-0,71 \pm 0,14$
$X_4$	$0,0445 \pm 0,015$	$0,0445 \pm 0,015$

Fonte: Do autor.

Com a biblioteca BAT também podemos gerar histogramas que representam

graficamente a p.d.f. posteriori de cada parâmetro, como mostrado na figura 15 para o parâmetro  $X_1$ .

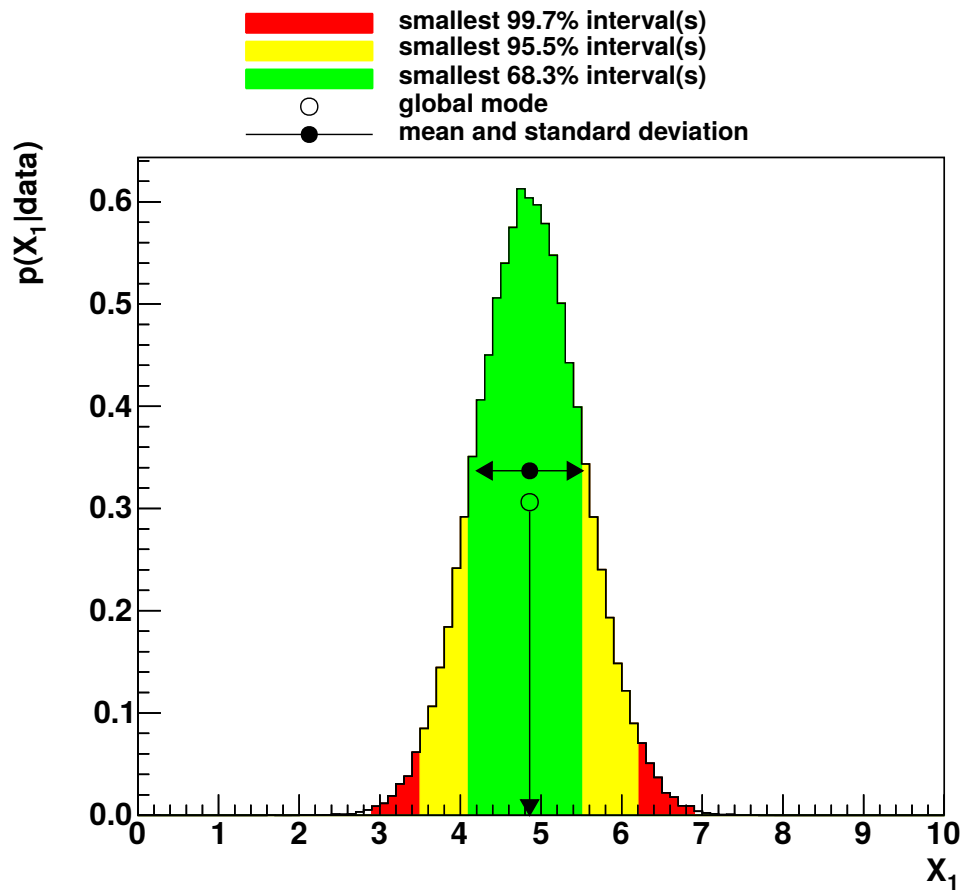


Figura 15 – P.d.f. posteriori do parâmetro  $X_1$  gerada pelo BAT.

Fonte: Do autor.

## 3 O EXPERIMENTO DOUBLE CHOOZ

### 3.1 Neutrinos de reatores

Na década de 1990, o deficit observado tanto nos neutrinos solares quanto nos atmosféricos intrigavam os pesquisadores. A proposta da oscilação de neutrinos era apenas uma entre muitas hipóteses, e a possibilidade de estudar neutrinos oriundos de uma fonte artificial poderia ser uma nova luz sobre o problema: será que neutrinos de reatores também apresentariam um deficit? Resultados anteriores (meados da década de 1980) do experimento Bugey pareciam indicar que sim(15), porém o próprio experimento refutou seus primeiros resultados conforme novas medidas foram sendo tomadas(16). Bugey detectou neutrinos de reatores a diferentes distâncias, sendo a mais longa cerca de 100 m. Ao não observar deficit nestas distâncias, era possível limitar os valores do ângulo de mistura e da escala de massa, mas não determiná-los. Batizado com o nome da cidade e da usina nuclear que pretendia estudar, o experimento Chooz (localizado no norte da França) surge com a proposta de verificar se um deficit nos neutrinos de reatores viria a surgir em uma distância de cerca de 1000 m dos reatores. Seus primeiros resultados surgem em 1999 indicando nenhuma evidência de um deficit(17), mesmo até o final de seu funcionamento(18).

Uma continuação natural do experimento Chooz, Double Chooz (DC) é projetado com o objetivo de averiguar se havia ou não um deficit no fluxo dos neutrinos de reatores(19). Entretanto o contexto científico de Double Chooz é bem diferente de seu antecessor: quando em 2002 o experimento SNO comprova a hipótese da oscilação para os neutrinos solares, fixa-se também a ordem de grandeza da escala de massa  $\Delta m_{12}^2$  e do ângulo de mistura  $\theta_{12}$ , mostrando que este último deveria ser grande (conhecida como *Large Mixing Angle solution*, ou LMA). Com a ordem de grandeza das escalas de massa de neutrinos solares e atmosféricos determinada, ficava claro que se houvesse algum deficit em neutrinos de reatores, este apareceria em distâncias próximas a 1000 m, onde Chooz se localizava. Portanto era necessário aprimorar a sensibilidade e a estatística, com relação ao experimento anterior, uma vez que a amplitude das oscilações deveria ser pequena (já que Chooz não havia observado deficit algum). O “double” de DC ilustra a característica fundamental deste experimento: dois detectores idênticos, um próximo e outro distante. O deficit, se existisse, seria determinado por uma comparação entre os dois, resultando em uma medida independente dos modelos para o combustível nuclear.

Com apenas um detector em funcionamento à 1050 m dos dois reatores da usina de Chooz e comparando o fluxo ali medido com resultados anteriores de Bugey, DC publicou seus primeiros resultados no final de 2011(11), com apenas 101 dias de coleta de dados, indicando a primeira evidência de um deficit no fluxo de neutrinos de reatores: a

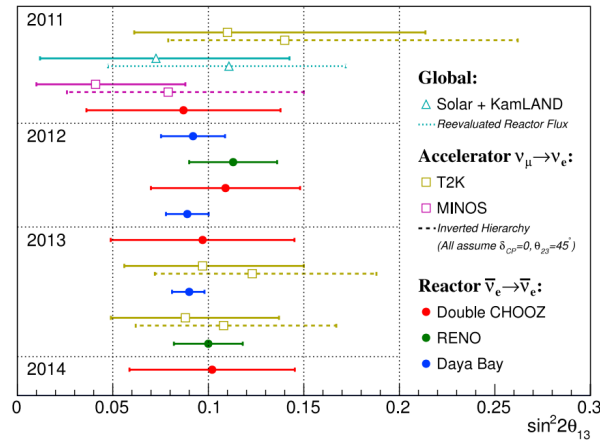


Figura 16 – Valores de  $\sin^2 2\theta_{13}$  obtidos por diversos métodos e experimentos, ao longo do tempo.

Fonte: Soo-Bong Kim (21).

razão entre o número de eventos observados e esperados foi de  $0,944 \pm 0,044$ , com nível de confiança de 90%. Este resultado é compatível um ângulo de mistura equivalente à  $\sin^2 2\theta_{13} = 0,086 \pm 0,053$ , à 90% de nível de confiança. Esta foi a primeira evidência de que o ângulo de mistura  $\theta_{13}$  era diferente de zero. Os últimos resultados publicados de DC antes da ativação do segundo detector(20), considerando 467,90 equivalente-dias de tomada de dados, atualizam o valor do fator de mistura para:

$$\sin^2 2\theta_{13} = 0,090_{-0,029}^{+0,032}. \quad (3.1)$$

O conjunto de dados desta publicação servirá como base para a análise bayesiana apresentada nos próximos capítulos, e o valor da mistura indicada na figura 3.1 servirá como base de comparação com os resultados futuramente obtidos.

No panorama geral dos neutrinos de reatores, outros dois experimentos concorrem com DC na determinação de  $\theta_{13}$ : RENO e Day Bay. Ambos operam com maior massa de detecção que DC (e conseqüentemente, detectam mais eventos) e com detectores a múltiplas distâncias, há mais tempo. Com isto, seus valores para o ângulo de mistura apresentam menor incertezas estatísticas e sistemáticas, quando comparadas à DC. A figura 16 ilustra a evolução das medidas de  $\sin^2 2\theta_{13}$  ao longo do tempo (publicações sucessivas), para cada experimento, incluindo neutrinos de aceleradores.

Em Dezembro de 2014 o detector próximo de DC entrou em operação. Em 2016, a colaboração DC deverá publicar seus primeiros resultados com dois detectores. No que se segue, faremos uma discussão sobre os mecanismos de detecção, a geometria do detector distante e os dados que levaram ao resultado apresentado na equação (3.1). Todas as informações específicas do FD estão disponíveis na referência (11), exceto quando indicado o contrário.

### 3.2 Mecanismo de detecção

Reatores nucleares produzem antineutrinos, especificamente  $\bar{\nu}_e$ , como subproduto das reações termonucleares. Estes neutrinos pode ser detectados através de uma das interações<sup>1</sup> chamadas de decaimento beta inverso (IBD, do inglês *inverse beta decay*). Quando colocada em função apenas dos estados iniciais e finais (ignorando quaisquer estados intermediários), temos:



onde o antineutrino eletrônico interage com um próton em repouso, produzindo um pósitron e um nêutron como estados finais. Esta é a mesma interação que permitiu a descoberta do neutrino por Cowan e Reines em 1956 e o sucesso na detecção de um neutrino depende da identificação destes estados finais e sua correlação, garantindo que foram produzidos pela mesma interação. O pósitron é identificado através da radiação Cherenkov e sua posterior aniquilação, enquanto o nêutron é capturado por núcleos do meio, desencadeando efeitos nucleares secundários que podem ser identificados. A figura 17 ilustra de forma qualitativa as etapas do processo de detecção.

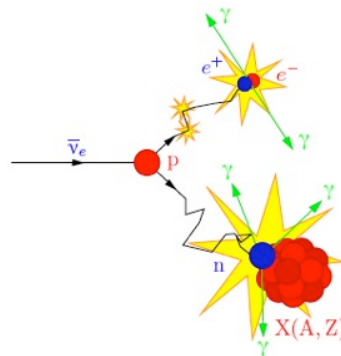


Figura 17 – Ilustração qualitativa da sequência de eventos que têm início com um IBD, da esquerda para a direita: (i) o antineutrino eletrônico interage com um próton em repouso, criando um pósitron e um nêutron; (ii) o pósitron é espalhado pelo material produzindo radiação Cherenkov, até que é aniquilado por um elétron do meio; (iii) o nêutron é espalhado pelo meio e é indetectável, até que é capturado por um núcleo do meio, desencadeando uma cascata nuclear identificável através da emissão de gamas.

Fonte: Do autor.

### 3.3 Estrutura do detector

Os dois detectores de DC são referidos como detector-próximo (*near detector*, ND) e detector-distante (*far detector*, FD). O ND está localizado a uma distância média de 410 m dos núcleos dos reatores, enquanto o FD a 1050 m dos núcleos. Entretanto não é

<sup>1</sup> Outra interação que também recebe este nome é:  $e^- + p^+ \longrightarrow \nu_e + n$ .



apenas a distância reator/detector que importa para DC: como a fonte de neutrinos não é pontual, sendo na verdade dois reatores com característica diferentes, a localização dos dois detectores deve proporcionar uma fração uniforme do fluxo de cada um dos reatores sobre o detector. Em outras palavras, se denotarmos os reatores por  $A$  e  $B$  e os fluxos em cada detector por  $\phi_{ND} = \phi_{ND}^A + \phi_{ND}^B$  e  $\phi_{FD} = \phi_{FD}^A + \phi_{FD}^B$ , então a disposição dos detectores é tal que:

$$\frac{\phi_{ND}^A}{\phi_{FD}^A} = \frac{\phi_{ND}^B}{\phi_{FD}^B}. \quad (3.3)$$

Chamamos esta condição de isofluxo, e sua função é garantir que a incerteza sistemática proveniente das diferenças nas condições de operação de cada reator sejam minimizadas quando os dados de ambos os detectores forem usados para determinar os parâmetros de oscilação.

O detector distante se vale da proteção extra da montanha rochosa sobre sua estrutura, equivalente a 300 metros de água, o que colabora para reduzir a incidência de raios cósmicos(19). Uma vez elencadas as diferenças na localização de cada detector, todas as demais características devem ser idênticas, visando sempre a minimização das incertezas sistemáticas. Desta forma, a descrição que se faz a seguir é válida para ambos os detectores.

Cada detector é composto de quatro vasos cilíndricos concêntricos, sendo os dois mais internos feitos de acrílico transparente, e os dois mais externos, feitos de aço. Destaca-se os seguintes componentes do detector:

*$\nu$ -target* (NT): alvo dos neutrinos. É o volume mais interno que constitui o próprio alvo para as interações com o fluxo de neutrinos incidente. Contém 10 m<sup>3</sup> de cintilador líquido enriquecido com gadolínio (Gd) dissolvido, para aumentar a seção de choque com nêutrons térmicos.

*$\gamma$ -catcher* (GC): capturador de raios gamas. É uma camada de 55 cm de espessura que envolve o NT. Contém líquido cintilador, porém sem gadolínio. Esse meio absorve energia do raio gama e a reemite em forma luz para ser detectado pelas PMTs (*photomultiplier tubes*). É responsável indicação do sinal atrasado (*delayed signal*).

*Buffer* (B) É uma camada de 105 cm de espessura, que envolve o GC. Está preenchida com óleo mineral, não cintilador. É meio destinado para a propagação da luz com eficiência. Sua parede externa é de aço, e possui PMTs fixadas nessa parede, apontadas para o centro do NT. As camadas existentes até o Buffer compõe o detector interno (ou *inner detector*, ID).

*Inner veto* (IV): veto interno. É a camada externa ao Buffer, tem espessura de 50 cm, contém líquido cintilador. Está equipado com 78 PMT, sendo 24 PMTs no topo, 12 estão a meia altura nas paredes laterais e 42 PMTs na base.

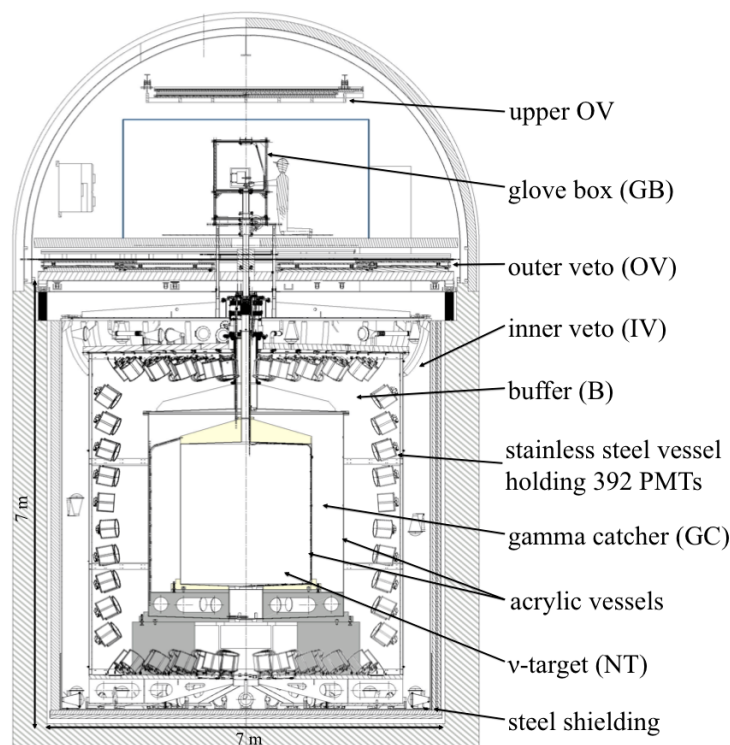


Figura 18 – Visão esquemática do detector FD, de cima para baixo: *upper OV*, primeira camada de identificação de raios cósmicos, composta por placas de cintilador plástico; *glove box (GB)*, estação de operação para manusear as fontes radioativas utilizadas no processo de calibração; *outer veto (OV)*, segunda camada de identificação de raios cósmicos, também composta por placas de cintilador plástico; *inner veto (IV)*, terceira camada de identificação de raios cósmicos e radiação ambiental, composta por cintilador líquido, compassando as camadas mais internas do detector; *buffer (B)*, Camada material, preenchida com óleo mineral não-cintilador, cuja função é garantir espaço para a propagação da luz até as PMTs; *gamma catcher (GC)*, camada de cintilador líquido otimizada para identificação de raios gamma gerados no interior do NT, visando a detecção do sinal atrasado (*delayed signal*);  *$\nu$ -target (NT)*, alvo central, composto de cintilador líquido e dopado com gadolínio, sendo este último responsável por aumentar a seção de choque para nêutrons térmicos, visando a captura do nêutron produzido no evento de decaimento beta inverso.

Fonte: Notas técnicas do experimento Double Chooz(11).

### 3.3.1 Sinal imediato

Quando ocorre no interior do detector, a interação IBD, equação (3.2), inicia uma cadeia de eventos que permite a identificação do neutrino e a medida de sua energia. O primeiro fenômeno visível desta cadeia é a deposição de energia ocasionada pela propagação do pósitron. Este perde energia para o meio até que se encontre no mesmo espaço de fase que um elétron atômico, podendo tanto ser aniquilado quanto capturado. O primeiro caso produz um par de raios gamas característico do processo, enquanto o segundo produz um estado positrônico, cujo tempo de vida retarda a eventual aniquilação do par(22). A

energia depositada antes da aniquilação é chamada de energia do sinal,  $E_{\text{signal}}$ , e permite obter a energia do neutrino com relativa precisão. Para isto, observe que na equação (3.2) que o neutrino deve ter uma energia mínima para que ela ocorra, uma vez que a massa dos estados finais é maior que a inicial e também que o momento do neutrino deve ser conservado, o que implica em pelo menos um dos estados finais ser criado em movimento, relativo ao laboratório. Assumindo que o próton inicial está em repouso, que a massa do nêutron é muito maior que seu momento ( $E_n \cong M_n$ ) e usando o fato de que  $M_n \gg m_e$ , temos

$$E_\nu + M_p = (M_n + m_e) + K_e, \quad (3.4)$$

onde  $K_e$  é a energia cinética do pósitron. Substituindo as massas[PDG] do nêutron  $M_n = 939,565$  MeV, do pósitron  $m_e = 0,511$  MeV e do próton  $M_p = 938,272$  MeV, temos que  $E_\nu \cong K_e + 1,804$  MeV. Ou seja, o limiar energético para que esta reação seja utilizada para detecção de antineutrino é  $E_\nu \geq 1,804$  MeV (repare que  $E_\nu$  é a energia total do neutrino mas que na realidade é constituída apenas da energia cinética, para fins práticos). Do ponto de vista experimental, observa-se a energia total depositada no detector pelo pósitron, incluindo sua aniquilação. Desta forma,  $E_{\text{signal}} = K_e + 2m_e$ . Reescrevendo a equação (3.4) em função de  $E_{\text{signal}}$  ao invés de  $K_e$ , temos:

$$E_\nu \cong E_{\text{signal}} + 0,782 \cong E_{\text{signal}} + 0,8 \text{ MeV}. \quad (3.5)$$

O resultado final é expresso com apenas um algarismo significativo devido à todas as aproximações feitas até aqui. A energia  $E_{\text{signal}}$  é obtida calibrando-se a produção de fótons no interior do detector com fontes radiológicas cujo espectro de energia é conhecido. Uma vez que  $E_{\text{signal}}$  é calibrado, a equação (3.5) é utilizada para obter a energia do neutrino  $E_\nu$  em todos os eventos.

Toda a cadeia de eventos descrita acima termina antes do nêutron ser capturado. Por isso este sinal energético recebe o nome de sinal imediato, do inglês *prompt signal*<sup>2</sup>. A identificação deste sinal é ambígua já que ela não garante que o pósitron observado veio de fato de um IBD e, em última instância, nem que se trata de fato de um pósitron. Por isso a identificação da captura subsequente do nêutron é essencial para a identificação do neutrino.

### 3.3.2 Sinal atrasado

O nêutron produzido pelo IBD não carrega informação sobre a energia do neutrino, porém é fundamental na identificação do IBD em si. Tanto que sua identificação faz parte dos chamados critérios de seleção para IBDs. Em DC, esta identificação é por intermédio de duas formas de capturas de nêutrons: por hidrogênio (H) e por gadolínio (Gd). A primeira pode parecer que seria a mais comum, devido à abundância de hidrogênio nos meios de

<sup>2</sup> No jargão do meio, é comum referir-se à “sinal prompt”, mesmo em português.

detecção, mas a segunda possibilidade corresponde à  $(85,30 \pm 0,08)\%$  das capturas(20). Gadolínio, em particular gadolínio-157, tem uma alta seção de choque de captura de nêutrons (muito maior que a do hidrogênio), e por isso é dissolvido nos meios de detecção com este propósito. Como o processo de reconhecimento de nêutrons capturados por hidrogênio ou por gadolínio é diferente, estes dois mecanismos são tratados como conjuntos de dados diferentes (quase que como experimentos diferentes). Assim, DC atualiza seus resultados organizados de acordo com estes mecanismos: Gd-I(11), Gd-II(23), H-I(24), Gd-III(20) e H-III(25), em ordem de publicação. Este trabalho utiliza apenas os dados do conjunto Gd-III e por isso a captura de nêutrons por hidrogênio não será discutida.

Para que o nêutron seja capturado, ele deve estar no mesmo espaço de fase que o núcleo de Gd, ou seja, momento e posição devem ser semelhantes. O núcleo de Gd está à temperatura ambiente, o que representa uma energia cinética típica da ordem de  $k_B T = 0,025 \text{ eV}$  (para temperaturas entre 17 e 20°C). Portanto o nêutron, gerado com energias  $K_n > 0,025 \text{ eV}$  irá sofrer múltiplos espalhamentos, perdendo energia cinética até atingir este limite, num processo conhecido como termalização, de forma a entrar em equilíbrio térmico com o meio. Para as energia típicas dos nêutrons criados na detecção de antineutrinos de reatores, este processo ocorre com um tempo médio de  $\tau_{\text{therm}} \cong 5 \mu\text{s}$ . Quando o equilíbrio é atingido, o nêutron está à mercê da captura, o que depende da seção de choque do processo e da concentração de agente no meio, no caso Gd-157. Gadolínio-157 tem a maior seção de choque de captura de nêutrons dentre os elementos estáveis: 259 kb<sup>3</sup>. A concentração de Gd no alvo central de DC foi escolhida de forma que o tempo médio de captura seja  $\tau_{\text{nGd}} = 30 \mu\text{s}$ , sendo que o processo de captura é governado por uma distribuição exponencial e a probabilidade de um nêutron térmico ser capturado, em um intervalo de tempo  $\Delta t$ , é dada por:

$$P_{\text{capture}}(0 \leq t \leq \Delta t) = \int_0^{\Delta t} \frac{e^{-t/\tau_{\text{nGd}}}}{\tau_{\text{nGd}}} dt, \quad (3.6)$$

ou seja, 99% dos nêutrons térmicos são capturados em  $0 \leq \Delta t \leq 138 \mu\text{s}$ . Uma vez capturado pelo núcleo de Gd, o nêutron provoca uma cascata nuclear onde cerca de 8MeV são emitidos na forma de gamas. Estes gamas são detectados e compõe o chamado sinal atrasado (do inglês, *delayed signal*). Quando se leva em consideração também o tempo de termalização, temos que o sinal imediato precisa ser seguido de um segundo sinal atrasado em uma janela de tempo tal que:

$$0,5 \leq \Delta t \leq 150 \mu\text{s}. \quad (3.7)$$

Um evento de detecção é chamado de candidato-IBD se for composto dois sinais dentro da janela de tempo definida por (3.7) e também se as energia depositadas nos dois

<sup>3</sup> Unidade de área utilizada para medir seções de choque: 1 b (bar) =  $10^{-32} \text{ cm}^2 = 100 \text{ fm}^2$ .

sinais satisfizer as seguintes condições:

$$0,5 \leq E_{\text{prompt}} \leq 20 \text{ MeV} \quad \text{e} \quad (3.8)$$

$$4 \leq E_{\text{delayed}} \leq 10 \text{ MeV}, \quad (3.9)$$

e também uma condição geométrica, de forma que a origem de ambos os sinais no interior do detector não estejam mais distantes do que  $\Delta R < 100 \text{ cm}$ . Estes quatro critérios formam a chamada regra de seleção para candidatos-IBD. Mesmo assim, sempre haverá interações de neutrinos que acabam sendo excluídas por estes critérios, bem como candidatos-IBD que não são provenientes de IBD reais. Ambos os casos são levados em consideração e discutidos mais adiante.

### 3.3.3 Eficiência na identificação de IBD

Double Chooz opera com um sistema de aquisição de dados sem tempo-morto, cujo *trigger* atinge 100% de eficiência para energias acima de 500 keV. A eficiência de detecção do sinal imediato foi determinada como sendo muito próxima de 100%, com incerteza desprezível. Já o sinal atrasado sofre diversas influências e por isso a eficiência de sua detecção é determinada experimentalmente e comparada com simulações. Isto é feito com uso de uma fonte radiológica de emissão de nêutrons (califórnio-252). Seu espectro é conhecido e também o mecanismo de captura, assim toma-se a razão entre o número de eventos identificados pelo número de eventos esperados. Esta razão é medida em diferentes alturas do detector, já que a fonte é introduzida através de uma abertura no topo e pode ser posicionada ao longo do eixo vertical do alvo de neutrinos. O resultado é uma medida da eficiência na identificação de nêutrons ao longo do eixo z, como indicada na figura 19, tanto para as medidas quanto para a simulação. Com exceção dos cerca de 35 cm finais, nas tampas do alvo, a eficiência é medida em  $98,58 \pm 0,19\%$ . Perto das tampas, cai para não menos do que 94%. Simulações da fonte de califórnio e do processo de detecção indicam que a eficiência deveria ser  $98,62 \pm 0,02\%$ , ligeiramente maior do que o valor experimental. Este procedimento de comparação permite recalibrar as simulações para que elas resultem exatamente na eficiência medida, o que é feito através de um coeficiente de correção  $C_{\text{Eff,IBD}} = 0,9996 \pm 0,0021$ , que nada mais é do que a razão entre a eficiência experimental e a obtida na primeira versão da simulação. Esta correção é importante pois se propaga para a simulação final dos nêutrons de IBD, que são essenciais para medir qualquer deficit nos neutrinos de reatores.

A mesma fonte de califórnio é utilizada para calibrar as medidas de energia do detector. Em todos os processos de detecção discutidos até aqui, o verdadeiro observável é o número de fotoelétrons (p.e., do inglês *photoelectrons*) produzidos nas PMTs ao redor do alvo. O número de fotoelétrons deveria ser proporcional à energia depositada no material alvo, porém é preciso determinar esta constante de proporcionalidade. Ao relacionar um certo número de fotoelétrons observados com determinada energia depositada, temos a

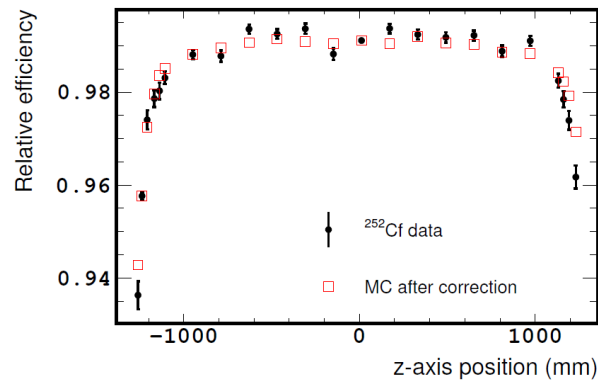


Figura 19 – Eficiência na reconstrução de nêutrons, provenientes do espectro do califórnio-252, cuja fonte é posicionada no interior do NT, ao longo do seu eixo z. Os pontos (preto) indicam os dados, enquanto os quadrados (vermelho) mostram o resultado das simulações deste processo.

Fonte: Notas técnicas do Experimento Double Chooz(20).

calibração do detector. A captura de nêutrons pelo hidrogênio da água produz fótons únicos de gamas com um pico de emissão em 2,223 MeV, e a distribuição da energia ao redor deste pico é bem conhecida. Ao inserirmos a fonte de califórnio-252 no centro do alvo, cujo espectro também é conhecido, obtém-se a relação 186,2 p.e./MeV. A simulação do mesmo procedimento resulta em um valor teórico de 186,6 p.e./MeV, com uma diferença quase imperceptível. A figura 20 mostra a comparação entre os dados de calibração e a simulação. Deste ponto em diante, os depósitos de energia observados no interior do detector através da carga elétrica depositada pelos fotoelétrons nas PMTs recebe o nome de energia visível,  $E_{\text{vis}}$ , e passa a ser nosso novo observável.

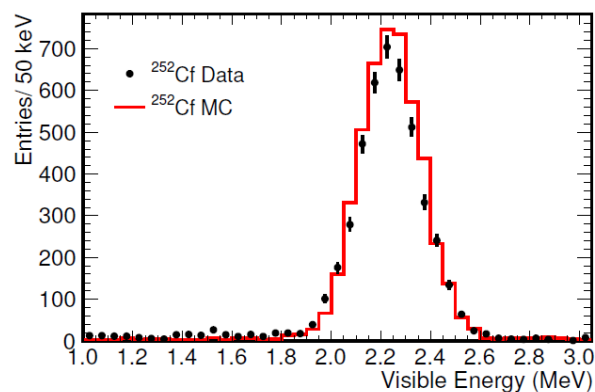


Figura 20 – Pico de captura do nêutron em hidrogênio: os pontos indicam medidas feitas com uma fonte de nêutrons, califórnio-252, inserida no centro do detector. O histograma em vermelho indica o resultado das simulações.

Fonte: Notas técnicas do experimento Double Chooz(20).

Uma vez que a escala de energia é calibrada em função do número de fotoelétrons

observados, é necessário confirmar se esta é de fato linear. Dois efeitos não lineares podem contribuir: a não-linearidade de carga (QNL) e não-linearidade da luz (LNL), principalmente o último. A contribuição de cada um é modelada e aplicada à escala, obtendo-se assim uma escala final que é linear. Com esta finalidade, outras fontes radioativas como cobalto-60, céσιο-137 e germânio-68 são utilizadas para verificar se a relação entre energia visível e a energia dos p.e. permanece linear em pontos diferentes da escala. Estes elementos são escolhidos pela sua emissão de fótons únicos na faixa de gamas, com energias na faixa  $0,5 \leq E_{\text{vis}} \leq 1,5$  MeV. Além disso, a calibração já feita com califórnio-252 através do pico do hidrogênio é comparada com fótons individuais provenientes da captura de nêutrons do califórnio pelo gadolínio. Apesar da energia total emitida neste processo ser cerca de 8 MeV, ela é distribuída em fótons individuais com energia média de 2,2 MeV, e portanto podem ser comparadas com a captura na água. Os resultados do estudo de linearidade são indicados na figura 21, onde vemos que de fato os efeitos da LNL fazem com que fótons de gamas com menor energia apresentem uma energia visível ainda menor. Por exemplo, o espectro de energia visível das medidas com germânio-68 é cerca de 1,5% menor do que a energia esperada para este decaimento. Aplicando o modelo da LNL é possível corrigir esta escala, trazendo-a de volta à linearidade. Porém, é mais simples e mais eficiente tratar ambos os efeitos, LNL e QNL, de uma única vez e considerando apenas um ajuste polinomial entre a energia visível esperada  $E_{\text{vis}}$  e a energia visível obtida via simulação  $E_{\text{vis}}^{MC}$ , obtendo-se assim uma calibração final. Assumindo-se um polinômio de segundo grau, temos:

$$E_{\text{vis}} = a + b \left( E_{\text{vis}}^{MC} \right) + c \left( E_{\text{vis}}^{MC} \right)^2, \quad (3.10)$$

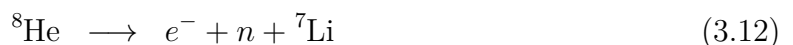
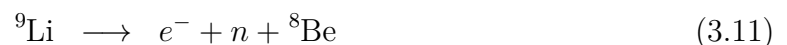
onde os coeficientes  $a$ ,  $b$  e  $c$  bem como suas incertezas e correlações são exibidos no Capítulo 4, nas equações 4.28 e 4.29, respectivamente.

### 3.3.4 Identificações espúrias

Entre os candidatos a IBD existe um subconjunto chamado de fundo (ou *background*), composto de eventos espúrios, ou seja, que não podem ser vetados pois apresentam elevada semelhança com eventos IBD reais. As fontes de eventos espúrios foram classificadas em três tipos: cosmogênicos, correlacionados e acidentais.

#### 3.3.4.1 Cosmogênicos

A passagem de raios cósmicos pela matéria dá origem aos chamados elementos cosmogênicos, isótopos de vida longa como  ${}^9\text{Li}$  e  ${}^8\text{He}$ . Lítio-9 é o isótopo radioativo do lítio com a segunda maior meia-vida, 178,3 ms, enquanto hélio-8 tem meia-vida igual a 119,0 ms. Os decaimentos de interesse são:





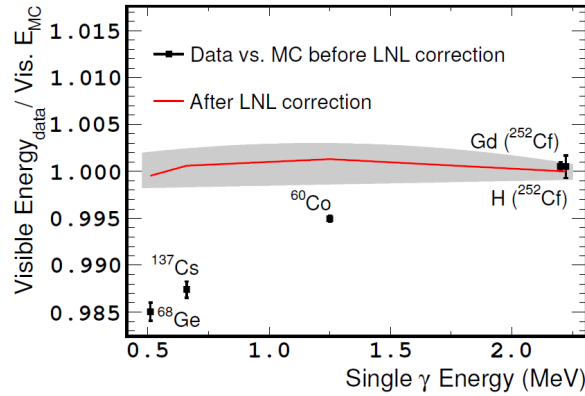


Figura 21 – Dados do ensaio de linearidade: O gráfico indica a razão entre a energia visível medida e a simulada,  $E_{vis}/E_{vis}^{MC}$ , sendo que os pontos se referem às medidas não corrigidas, enquanto a linha vermelha indica a nova escala já corrigida para efeitos de LNL e QNL. A faixa cinza indica a incerteza sistemática remanescente.

Fonte: Notas técnicas do experimento Double Chooz(20).

onde ambos são emissores  $\beta - n$ , ou seja, são capazes de emular tanto o sinal imediato quanto o sinal atrasado. Parte dos sinais gerados por estes elementos podem ser vetados com a aplicação de cortes específicos no sinal, porém não pode eliminá-lo por completo. Simulações indicam que mesmo após os cortes espera-se uma taxa de eventos com origem cosmogênica de  $\alpha_{Li+He} = 0,97^{+0,41}_{-0,16}$  eventos/dia.

### 3.3.4.2 Correlacionados

A passagem de múons pelas rochas ao redor do detector pode ocasionar a espalção de nêutrons rápidos (FN, *fast neutros*). Estes nêutrons podem mimetizar um sinal do tipo IBD se o nêutron interagir com um próton, espalhando-o e gerando um sinal na faixa de energia esperada para o sinal imediato. O nêutron que causou o espalhamento, ou um outro aleatório, é subsequentemente termalizado e capturado pelo gadolínio. Outra possibilidade é um múon chegar ao NT através da tubulação utilizada para descer as amostras radioativas de calibração. Se este múon, por coincidência, decair no interior do detector (SM, *Stopping Muon*), os acionamentos sequenciais causados pelo múons e pelo elétron originário de seu decaimento também podem imitar um IBD. Esta fonte é referida como sinais correlacionados (ou apenas correlacionados), e podem ser vetados quando o múon que gera o evento deposita energia no veto interno, os chamados *IV-tagged*. Uma pequena parte não deixa sinal algum no IV e por isso não pode ser vetado. A estimativa da taxa de eventos correlacionados é feita através destes eventos *IV-tagged*, que em geral seriam vetados da base de dados. Quando isolados, é possível ver que seu espectro é constante ao longo das energias esperadas para antineutrinos de reatores. Este fato abre a possibilidade de checar a medida da taxa observando-se o volume total de candidatos



a IBD com energias acima do espectro de antineutrinos ( $E_{vis} \geq 20$  MeV). A figura 22 mostra a comparação de três conjuntos de dados: candidatos a IBD, candidatos IV-tagged e candidatos com energias na faixa  $20 \leq E_{vis} \leq 30$  MeV. Este estudo indica que a taxa esperada para este fundo é de  $\alpha_{FN+SM} = 0,604 \pm 0,051$  eventos/dia.

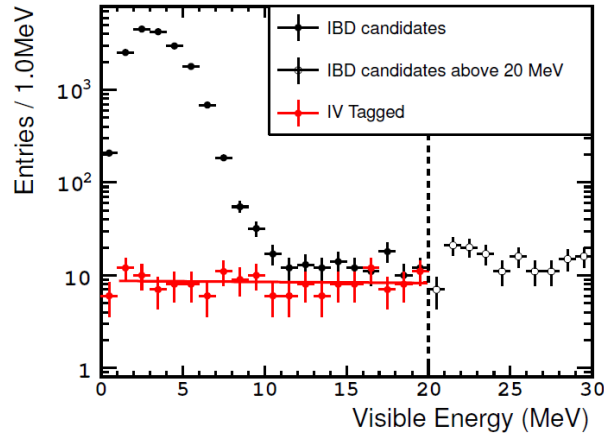


Figura 22 – Espectros de energia depositada no sinal imediato, provenientes de três conjuntos de dados distintos: candidatos a IBD (pontos pretos preenchidos), candidatos tipo IV-tagged que seriam vetados (pontos vermelhos) e candidatos com energias entre 20 e 30 MeV, onde não deveriam haver sinais reais.

Fonte: Notas técnicas do experimento Double Chooz(20).

### 3.3.4.3 Acidentais

A mera coincidência estatística de eventos únicos que aparecem na mesma janela de tempo é uma fonte de fundo. A taxa e o espectro deste sinal é medido através do método *off-time*, onde adquire-se eventos aleatórios em uma janela de tempo que tem início mais de 1 s após o sinal atrasado de um candidato IBD, durando até o sinal imediato do evento seguinte. Esta coleção de sinais aleatórios permite calcular observar a frequência das coincidências, bem como suas energias aparentes, como indicado na figura 23, resultando em uma taxa de  $\alpha_{Acc} = 0,0701 \pm 0,0026$  eventos/dia.

## 3.4 Resultados

Após a aquisição de 467,90 dias-equivalente de dados, os resultados obtidos são exibidos na figura 24. Nesta figura também podem ser vistas as contribuições de fundo, bem como o ajuste da previsão de oscilação, resultando no fator de mistura  $\sin^2 2\theta_{13} = 0,090$ , equivalente a  $\theta_{13} = 8,7^\circ$ . Este resultado é obtido através de uma análise frequentista dos dados, onde o ângulo de mistura é obtido através da minimização de um  $\chi^2$ . No Cap. 4 desenvolveremos a análise bayesiana destes dados, obtendo assim um valor para o ângulo de mistura  $\theta_{13}$  = por intermédio desta estatística, que será posteriormente comparado com o obtido pela colaboração Double Chooz.

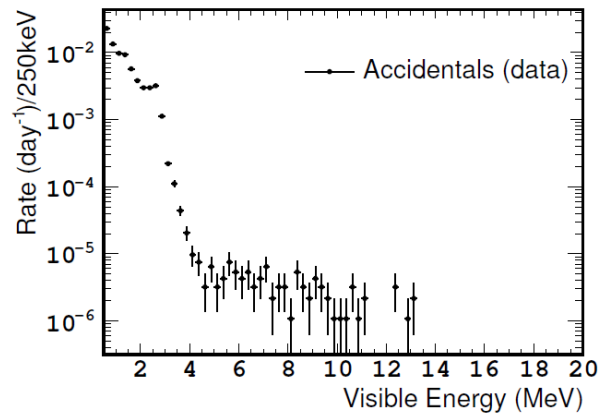


Figura 23 – Espectro de energia para sinais imediatos, obtidos durante a janela de off-time do detector.

Fonte: Notas técnicas do experimento Double Chooz(20).

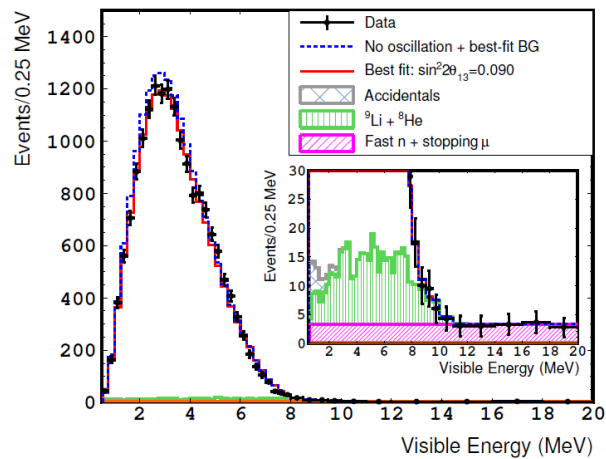


Figura 24 – O espectro de energia visível medido a partir dos sinais imediatos dos candidatos a IBD (pontos em preto), superposto aos espectros esperados para os fundos cosmogênico (verde), correlacionado (rosa) e acidental (azul). A linha azul tracejada mostra o espectro esperado se não houvesse oscilação, enquanto a linha contínua vermelha representa o resultado do ajuste do modelo de oscilação.

Fonte: Notas técnicas do experimento Double Chooz(20).

## 4 DETERMINANDO $\theta_{13}$

### 4.1 Os parâmetros

O experimento Double Chooz tem por objetivo determinar o ângulo de mistura  $\theta_{13}$ . De fato, o parâmetro que determinamos é o valor de  $\sin^2 2\theta_{13}$ . O que vamos fazer aqui é a aplicação da teoria desenvolvida no capítulo 2, utilizando análise bayesiana para determinar um conjunto de parâmetros com base em dados experimentais. Além do parâmetro principal, o desenvolvimento de todo processo experimental nos leva a incluir uma série de outros parâmetros que serão necessários para atender todas as particularidades do experimento. No caso do Experimento Double Chooz, são necessários nove parâmetros, que listamos a seguir

- (a)  $\sin^2 2\theta_{13}$  O valor desse parâmetro envolve o ângulo de mistura  $\theta_{13}$  relacionado com o fenômeno de oscilação dos neutrinos.
- (b)  $\Delta m_{13}^2$  Diferença dos quadrados de massas dos neutrinos físicos associados com o ângulo de mistura  $\theta_{13}$ . Esse parâmetro já possui valor conhecido, mesmo assim ele entra como parâmetro a ser determinado, como será explicado mais tarde.
- (c)  $\mathbf{a}'$ ,  $\mathbf{b}'$ ,  $\mathbf{c}'$  São três parâmetros necessários para obtenção da energia, estimada por um algoritmo de Monte Carlo, que determina a energia visível do sinal de prompt com base nas cargas existentes no conjunto de PMT's do detector. Os valores desses parâmetros também são conhecidos, como no caso de  $\Delta m_{13}^2$ .
- (d)  $\alpha_{Li+He}$ ,  $\alpha_{FN+MS}$ ,  $\alpha_{acc}$ ,  $\alpha_{res}$  Parâmetros para ajustes de backgrounds.

### 4.2 Mínimos quadrados

O método dos mínimos quadrados pode ser aplicado quando o erro de uma medida experimental, varia de modo gaussiano, em torno de um valor central, o qual é dado por um modelo teórico representando os dados. Este método consiste em minimizar a seguinte quantidade

$$\chi^2 = \sum_{i=1}^N \left( \frac{F_i - D_i}{\sigma_i} \right)^2, \quad (2.93)$$

onde  $F_i$  representa o modelo teórico e  $D_i$  os dados observados. Podemos reescrever essa equação como

$$\chi^2 = \sum_{i=1}^N \sum_{j=1}^N \left( \frac{F_i - D_i}{\sigma_i} \right) \delta_{ij} \left( \frac{F_j - D_j}{\sigma_j} \right) \quad (4.1)$$

$$= \sum_{i=1}^N \sum_{j=1}^N (F_i - D_i) \frac{\delta_{ij}}{\sigma_i \sigma_j} (F_j - D_j). \quad (4.2)$$

Desenvolvendo o elemento de matriz, encontramos

$$\frac{\delta_{ij}}{\sigma_i \sigma_j} = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}^{-1}. \quad (4.3)$$

O resultado acima, mostra que a matriz em (4.2) representa o inverso de uma matriz de covariância,

$$M_{ij} = \delta_{ij} \sigma_i \sigma_j. \quad (4.4)$$

O fato da matriz ser diagonal, indica ausência de correlação entre os dados. De um modo mais geral, podemos escrever

$$\chi^2 = \sum_{i=1}^N \sum_{j=1}^N (F_i - D_i) M_{ij}^{-1} (F_j - D_j), \quad (4.5)$$

onde a matriz de covariâncias  $M_{ij}$  pode ser não diagonal, incorporando informações de correlação entre os dados observados.

### 4.3 Definição do $\chi^2$ do DCIII

No caso do experimento de Double Chooz, a função  $\chi^2$  é definida como

$$\begin{aligned} \chi^2 = & \sum_i^B \sum_j^B (N_i^{pred} - N_i^{obs}) M_{ij}^{-1} (N_j^{pred} - N_j^{obs}) \\ & + \frac{(\Delta m^2 - \Delta m_{ee}^2)^2}{(\sigma_{ee}^{\{up,down\}})^2} + \frac{(\alpha_{residual} - 1)^2}{(\sigma_{residual})^2} \\ & + \frac{(\alpha_{Li+He} - 1)^2}{(\sigma_{Li+He}^{\{up,dow\}})^2} + \frac{(\alpha_{acc} - 1)^2}{(\sigma_{acc(syst)})^2} + \frac{(\alpha_{FN+SN} - 1)^2}{(\sigma_{FN+SN})^2} \\ & + [(a' - a'_{CV}), (b' - b'_{CV}), (c' - c'_{CV})] \\ & \begin{bmatrix} (\sigma_{a'})^2 & \rho_{a'b'} (\sigma_{a'} \sigma_{b'}) & \rho_{a'c'} (\sigma_{a'} \sigma_{c'}) \\ \rho_{b'a'} (\sigma_{b'} \sigma_{a'}) & (\sigma_{b'})^2 & \rho_{b'c'} (\sigma_{b'} \sigma_{c'}) \\ \rho_{c'a'} (\sigma_{c'} \sigma_{a'}) & \rho_{c'b'} (\sigma_{c'} \sigma_{b'}) & (\sigma_{c'})^2 \end{bmatrix}^{-1} \begin{bmatrix} (a' - a'_{CV}) \\ (b' - b'_{CV}) \\ (c' - c'_{CV}) \end{bmatrix} \\ & + 2 \left( N_{off}^{obs} \cdot \ln \left[ \frac{N_{off}^{obs}}{N_{off}^{pred}} \right] + N_{off}^{pred} - N_{off}^{obs} \right). \end{aligned} \quad (4.6)$$

A primeira linha de (4.6) é a equação (4.5), só que agora com a matriz  $M_{ij}$  bem ampliada e não mais diagonal. A função  $\chi^2$  depende de todos os nove parâmetros indicados antes. Nossa tarefa é encontrar quais são os valores desses parâmetros que minimizam o  $\chi^2$ ,

obtendo assim a melhor estimativa para eles, o que inclui o valor de  $\sin^2 2\theta_{13}$ , que é o objetivo principal do capítulo.

As notas técnicas publicadas pela Colaboração do Experimento de Double Chooz(26), fornecem os valores necessários para calcular a quantidade  $\chi^2$ .

## 4.4 Termos do $\chi^2$

### 4.4.1 Intervalos do histograma

Os índices  $i, j$  das somatórias em (4.6) se referem aos intervalos do histograma de energia do neutrino, escolhidos para representar o espectro do sinal do prompt. Esses intervalos devem ser construídos conforme o modelo

Tabela 3 – Intervalos do histograma para o sinal de prompt

Intervalo	Largura	Número de intervalos
0,5 - 8 MeV	250 keV	30
8 - 10 MeV	500 KeV	4
10 - 12 MeV	1 MeV	2
12 - 20 MeV	2 MeV	4
0,5 - 20 MeV		40

Fonte: NT Double Chooz(11).

### 4.4.2 Neutrinos previstos

É o número de neutrinos previstos, considerando a soma da oscilação com os backgrounds. Esse valor é uma função de todos os parâmetros, o que podemos resumir como

$$N_i^{pred} = \sum_{R=1,2} N_i^{\bar{\nu}_e, R} \left( \sin^2 2\theta_{12}, \Delta m^2, a', b', c' \right) + \alpha_{Li+He} \cdot N_i^{Li+He} + \alpha_{FN+SM} \cdot N_i^{FN+SM} + \alpha_{acc} \cdot N_i^{acc}. \quad (4.7)$$

### 4.4.3 Matriz de covariância

É uma generalização do caso mais simples indicado pela equação (4.4). Ela é composta pelas seguintes partes

$$M_{ij} = M_{ij}^{stat} + M_{ij}^{reactor} + M_{ij}^{efficiency} + M_{ij}^{Li+He} + M_{ij}^{acc(stat)} \quad (4.8)$$

- $M_{ij}^{stat}$  é uma matriz diagonal com os erros de Poisson provenientes do vetor predição  $N_i^{pred}$ , então:  $M_{ii}^{stat} = N_i^{pred}$ ,  $M_{ij}^{stat} = 0$  quando  $i \neq j$ .
- $M_{ij}^{reactor}$  codifica a incerteza da predição do reator. Foi feita usando a ferramenta DCRxtr de DOGS. É fornecida no arquivos de dados.

- $M_{ij}^{efficiency}$  codifica as incertezas devido a ineficiência na detecção do nêutron, ineficiência FV e IV, likelihood Li e as incertezas no número de prótons. Ela tem a forma  $M_{ij}^{efficiency} = \epsilon_{eff}^2 N_i^{pred} N_j^{pred}$ , sendo  $\epsilon_{eff} = 0,63\%$  o valor fornecido pela nota técnica.
- $M_{ij}^{Li+He}$  é a forma do erro no espectro do  $Li + He$ . Essa matriz também é fornecida nos arquivos de dados.
- $M_{ij}^{acc(stat)}$  é uma matriz diagonal que codifica a taxa da componente estatística do background acidental, sendo

$$N_i^{acc} = \frac{N_i^{obs}}{N_{total}^{obs}} \times acc\_rate\_perday \times run\_time \quad (4.9)$$

$$acc\_rate\_perday = 0,0701 \pm 0,0003 (stat) \pm 0,0054 (syst) \quad (4.10)$$

$$\sigma_{\{stat\ acc\}} = 0,0003/0,0701 \quad (4.11)$$

$$M_{ii}^{acc} = N_i^{acc} \times N_{total}^{acc} \times \sigma_{\{stat\ acc\}}^2 \quad (4.12)$$

#### 4.4.4 Pull terms

Para explicar os *pull terms* vamos tomar como exemplo o parâmetro  $\Delta m^2$ . Ele já é conhecido de outros experimentos com neutrinos e seu valor é  $\Delta m^2 = \Delta m_{31}^2 = (2,53 \pm 0,12) \times 10^{-3} \text{ eV}^2$ . Então, em princípio, não deveríamos ter um parâmetro e sim, uma constante. Mas, como existe uma incerteza, devemos permitir que o algoritmo que minimiza  $\chi^2$ , encontre o melhor valor dentro da liberdade que a incerteza proporciona. Fazemos isto informando  $\Delta m^2$  como parâmetro e não como uma constante. Mas, precisamos evitar que o algoritmo busque um valor para esse parâmetro, muito diferente do valor da constante que ele representa. O modo de fazer isto é introduzir um termo no  $\chi^2$  que puxe o valor do parâmetro para próximo do valor da constante. O termo acrescentado para o caso de  $\Delta m^2$  é

$$\left( \frac{\Delta m^2 - \Delta m_{ee}^2}{\sigma_{ee}} \right)^2, \quad (4.13)$$

com esse termo, se o algoritmo tentar um valor muito fora do valor central, o  $\chi^2$  aumenta, fazendo que tal valor seja estimado novamente.

Os *pull terms* também podem ser introduzidos como produtos. Por exemplo, se o  $\chi^2$  depende de uma constante  $c \pm \sigma_c$ , então podemos definir um parâmetro  $\alpha$ , e usar  $\alpha c$  em lugar de  $c$ . Assim estamos permitindo uma liberdade de ajuste na constante  $c$ . Nesse caso, o *pull term* acrescentado em  $\chi^2$  deve ser da forma

$$\begin{aligned} \left( \frac{\alpha c - c}{\sigma_c} \right)^2 &= \left( \frac{\alpha - 1}{\frac{\sigma_c}{c}} \right)^2 \\ &= \left( \frac{\alpha - 1}{\sigma_\alpha} \right)^2, \end{aligned} \quad (4.14)$$

onde  $\sigma_\alpha = \sigma/c$ , como exemplo, veja equação (4.11). Com as equações (4.13) e (4.14) explicamos *pull terms* presentes na segunda e terceira linha de (4.6).

Vamos considerar agora os *pull terms* que começam na quarta linha. Eles se referem aos parâmetros de calibração da energia visível,  $a'$ ,  $b'$ ,  $c'$ , com os quais é feito um ajuste polinomial, expresso por

$$E_{vis} = a' + b' \cdot E_{vis}^{MC} + c' \cdot (E_{vis}^{MC})^2. \quad (4.15)$$

Esses parâmetros já foram determinados em outra etapa do experimento, sendo seus valores indicados  $a'_{CV}$ ,  $b'_{CV}$ ,  $c'_{CV}$ . Em princípio, a inclusão desses *pull terms* deveriam ser apenas como

$$\left(\frac{a' - a'_{CV}}{\sigma_{a'}}\right)^2 + \left(\frac{b' - b'_{CV}}{\sigma_{b'}}\right)^2 + \left(\frac{c' - c'_{CV}}{\sigma_{c'}}\right)^2. \quad (4.16)$$

Porém, os parâmetros  $a'$ ,  $b'$ ,  $c'$  são quantidades correlacionadas, isto implica que a variação que ocorrer com um deles, acarretará variações nos demais. Para considerar tal correlação, devemos utilizar *pull terms* na forma da equação (4.5). Vamos representar o conjunto  $a'$ ,  $b'$ ,  $c'$  por  $a_i$ . Os *pull terms* serão

$$\sum_{i=1}^3 \sum_{j=1}^3 (a_i - a_{0i}) M_{ij}^{-1} (a_j - a_{0j}), \quad (4.17)$$

onde  $M_{ij} = \sigma_{a_i a_j}^2$ . O coeficiente de correção é definido como

$$\rho_{ij} = \frac{\sigma_{a_i a_j}^2}{\sigma_{a_i} \sigma_{a_j}}, \quad (4.18)$$

assim

$$M_{ij} = \sigma_{a_i a_j}^2 = \rho_{ij} \sigma_{a_i} \sigma_{a_j} \quad (4.19)$$

$$\begin{aligned} & \sum_{i=1}^3 \sum_{j=1}^3 (a_i - a_{0i}) M_{ij}^{-1} (a_j - a_{0j}) = \\ & \begin{bmatrix} a' - a'_{CV} & b' - b'_{CV} & c' - c'_{CV} \end{bmatrix} \begin{bmatrix} \sigma_{a'}^2 & \rho_{a'b'} \sigma_{a'} \sigma_{b'} & \rho_{a'c'} \sigma_{a'} \sigma_{c'} \\ \rho_{b'a'} \sigma_{b'} \sigma_{a'} & \sigma_{b'}^2 & \rho_{b'c'} \sigma_{b'} \sigma_{c'} \\ \rho_{c'a'} \sigma_{c'} \sigma_{a'} & \rho_{c'b'} \sigma_{c'} \sigma_{b'} & \sigma_{c'}^2 \end{bmatrix}^{-1} \begin{bmatrix} a' - a'_{CV} \\ b' - b'_{CV} \\ c' - c'_{CV} \end{bmatrix} \end{aligned} \quad (4.20)$$

onde usamos o fato de que  $\rho_{ii} = 1$ . Com isto, ficam explicados os *pull terms* da quarta linha  $\chi^2$ .

O último termo de  $\chi^2$  não é um *pull term*, mas sim uma contribuição para o logaritmo da likelihood referente a distribuição de Poisson que quantifica a ocorrência ocasional de neutrinos no detector, quando os reatores estão desligados. Seja  $N_{off}^{pred} = N^{\{p\}}$  o número de  $\nu_e$  previstos para o período total em que os dois reatores ficam desligados. A probabilidade de que seja observado uma quantidade  $N_{off}^{obs} = N$ , nesse período, é determinada pela distribuição de Poisson, sendo

$$prob(N|\alpha_{res}) = \frac{(N^{\{p\}})^N e^{-N^{\{p\}}}}{N!}. \quad (4.21)$$

Essa probabilidade deve ser incluída na função  $L$

$$\begin{aligned} L &= -\frac{1}{2}\chi^2 + \ln \left[ \frac{(N^{\{p\}})^N e^{N^{\{p\}}}}{N!} \right] \\ &= -\frac{1}{2}\chi^2 + (N \ln N^{\{p\}} - N^{\{p\}} - \ln N!). \end{aligned} \quad (4.22)$$

Usando a aproximação de Stirling  $N! \approx N \ln N - N$

$$\begin{aligned} L &= -\frac{1}{2}\chi^2 + (N \ln N^{\{p\}} - N^{\{p\}} - N \ln N + N) \\ &= -\frac{1}{2} \left[ \chi^2 + 2(-N \ln N^{\{p\}} + N^{\{p\}} + N \ln N - N) \right] \\ &= -\frac{1}{2} \left[ \chi^2 + 2 \left( N \ln \frac{N}{N^{\{p\}}} + N^{\{p\}} - N \right) \right]. \end{aligned} \quad (4.23)$$

Isolando somente o termos de Poisson, obtemos o último termo da equação (4.6)

$$2 \left( N_{\text{off}}^{\text{obs}} \ln \frac{N_{\text{off}}^{\text{obs}}}{N_{\text{off}}^{\text{pred}}} + N_{\text{off}}^{\text{pred}} - N_{\text{off}}^{\text{obs}} \right) \quad (4.24)$$

Significado dos *pull terms* utilizados

- $\Delta m^2$  Diferença de massa do neutrino ao quadrado, medida em  $\text{eV}^2$ .
- $\alpha_{\text{residual}}$  Taxa prevista de neutrinos residuais no reator desligado.
- $\alpha_{\text{Li+He}}$  Taxa de  ${}^9\text{Li}$  e  ${}^8\text{He}$ . O decaimento dos isótopos de  ${}^9\text{Li}$  e  ${}^8\text{He}$  geram sinais de prompts e delayed que não podem ser distinguidos dos mesmos sinais gerados por neutrinos.
- $\alpha_{\text{acc}}$  Taxa acidental, representa a incidência de neutrinos, mesmo quando os reatores estão desligados.
- $\alpha_{\text{FN+SM}}$  Taxa de nêutrons rápidos + múons parando.
- $a'$ ,  $b'$ , e  $c'$  parâmetros de calibração da energia.

#### 4.4.5 Termos de reator desligado

O termo final da equação (4.6) é a contribuição da likelihood de Poisson ao termo  $\chi^2$  proveniente do período em que o reator encontra-se desligado.

$$\begin{aligned} N_{\text{off}}^{\text{pred}} &= (\alpha_{\text{Li+He}} R_{\text{Li+He}} + \alpha_{\text{FN+SM}} R_{\text{FN+SM}} + \alpha_{\text{acc}} R_{\text{acc}}) \cdot T_{\text{off}} \\ &\quad + \alpha_{\text{residual}} \cdot P_{\text{surv}}(\theta_{13}, \langle \alpha \rangle) \cdot N_{\text{residual}}^{\text{pred}} \end{aligned} \quad (4.25)$$

$$\langle \alpha \rangle = \left\langle \sin^2 \frac{\Delta m_{ee}^2}{4E} \right\rangle \simeq 0,75 \quad (4.26)$$



## 4.5 Resumo das entradas e melhor ajuste

### 4.5.1 Escala de energia

$$E_{vis}^{MC} \rightarrow a' + b' \cdot E_{vis}^{MC} + c' \cdot (E_{vis}^{MC})^2 \quad (4.27)$$

Os parâmetros  $a'$ ,  $b'$ , e  $c'$  na equação (4.27) possuem os seguintes valores

$$\begin{aligned} a'_{CV} \pm \sigma_{a'} &= -0,0271 \pm 0,0062 \\ b'_{CV} \pm \sigma_{b'} &= 1,0120 \pm 0,0077 \\ c'_{CV} \pm \sigma_{c'} &= -0,0001 \pm 0,0006 \end{aligned} \quad (4.28)$$

$$\begin{pmatrix} \rho_{a'a'} & \rho_{a'b'} & \rho_{a'c'} \\ \rho_{b'a'} & \rho_{b'b'} & \rho_{b'c'} \\ \rho_{c'a'} & \rho_{c'b'} & \rho_{c'c'} \end{pmatrix} = \begin{pmatrix} 1 & -0,303868 & 0,00707935 \\ -0,303868 & 1 & -0,285526 \\ 0,00707935 & -0,285526 & 1 \end{pmatrix} \quad (4.29)$$

### 4.5.2 Backgrounds

- Taxa acidental:  $R_{acc} = 0,0701 \pm 0,0054$  (syst)  $\pm 0,0003$  (stat) dia<sup>-1</sup>
- Taxa cosmogênico:  $R_{Li+He} = 0,97^{+0,41}_{-0,16}$  dia<sup>-1</sup>
- Taxa de nêutrons rápidos + múons parando:  $R_{FN+SM} = 0,60 \pm 0,05$  dia<sup>-1</sup>

### 4.5.3 Reator desligado

- Número de  $\bar{\nu}_e$  residuais no reator desligado  $N_{residual}^{prd} = 1,57 \pm 0,47$
- Fator de oscilação médio  $\langle \alpha \rangle = 0,75$
- Tempo de desligamento do reator = 7 238 dias.

### 4.5.4 Resumo das entradas

Com a tabela (4) resumimos os valores de entradas para os oito parâmetros extras que devemos usar para encontrar a melhor estimativa de todos os parâmetros, os extras e o principal  $\sin^2 2\theta_{13}$ .

### 4.5.5 Resultados

Com os valores iniciais dos parâmetros, a quantidade  $\chi^2$  está concluída e o seu valor pode ser obtido para um conjunto particular de dados observados  $\mathbf{D}$  e parâmetros  $\mathbf{X}$

$$\chi^2 = \chi^2(\mathbf{X}; \mathbf{D}). \quad (4.30)$$

Tabela 4 – Valores iniciais para os parâmetros.

Parâmetro	Valor inicial
$\sin^2 2\theta_{13}$	—
$\Delta m^2$ ( $10^{-3}$ eV <sup>2</sup> )	$2,44^{+0,09}_{-0,10}$
$a'$	$-0,0271 \pm 0,006$
$b'$	$1,0120 \pm 0,0077$
$c'$	$-0,0001 \pm 0,0006$
Li+He ( $d^{-1}$ )	$0,97^{+0,41}_{-0,16}$
FN+SM ( $d^{-1}$ )	$0,604 \pm 0,0026$
Acidental ( $d^{-1}$ )	$0,0701 \pm 0,0026$
Residual	$1,57 \pm 0,47$

Fonte: NT Double Chooz(11).

Então, calculamos a likelihood

$$P(\{\mathbf{D}|\mathbf{X}\}) \propto \exp \left[ -\frac{1}{2} \chi^2(\mathbf{X}; \mathbf{D}) \right], \quad (4.31)$$

e finalmente o logaritmo natural da likelihood

$$L = const - \frac{1}{2} \chi^2(\mathbf{X}; \mathbf{D}). \quad (4.32)$$

Como vimos no capítulo 2, seção 2.5, a melhor estimativa para o conjunto de parâmetros  $\mathbf{X}$  é dada pela solução da equação

$$\nabla L = 0. \quad (4.33)$$

Devido a dimensão do problema proposto em (4.6), temos que fazer uma abordagem computacional afim obter os nove parâmetros do Experimento de Double Chooz. Esta abordagem será feita utilizando o kit de ferramentas para análise bayesiana, conhecido como BAT(27). Em sua home page, encontramos tutoriais que nos auxiliam na escrita do código C++ necessário para utilizar tal ferramenta. Fizemos o código C++ e o executamos algumas vezes. O tempo médio da execução foi de 4h40. Os resultados obtidos para os parâmetros estão indicados na tabela 5.

A seguir apresentamos a matriz de covariância, tabela 6 e a matriz correlação, tabela 7, ambas determinadas pela análise da ferramenta BAT.

A figura 25 dá a representação gráfica dos dados experimentais, os backgrounds, o melhor ajuste e o espectro sem oscilação. Na figura 26 temos a p.d.f do parâmetro  $\sin^2 2\theta_{13}$  que é uma probabilidade marginal obtida da p.d.f posteriori conjunta que representa todos os parâmetros calculados. O cálculo da p.d.f conjunta e das p.d.f marginais, são feitos também pela ferramenta BAT.

Tabela 5 – Valores iniciais e resultado final.

Parâmetro	Valor inicial	Melhor ajuste
$\sin^2 2\theta_{13}$	—	$0,098 \pm 0,030$
$\Delta m^2 (10^{-3} \text{ eV}^2)$	$2,44^{+0,09}_{-0,10}$	$2,513 \pm 0,095$
$a'$	$-0,0271 \pm 0,006$	$(-2,44 \pm 0,59) \times 10^{-2}$
$b'$	$1,0120 \pm 0,0077$	$(1,0069 \pm 0,0056)$
$c'$	$-0,0001 \pm 0,0006$	$(-6,7 \pm 5,8) \times 10^{-3}$
Li+He ( $\text{d}^{-1}$ )	$0,97^{+0,41}_{-0,16}$	$1,02 \pm 0,16$
FN+SM ( $\text{d}^{-1}$ )	$0,604 \pm 0,0026$	$0,594 \pm 0,038$
Acidental ( $\text{d}^{-1}$ )	$0,0701 \pm 0,0026$	$0,0711 \pm 0,0053$
Residual	$1,57 \pm 0,47$	$1,60 \pm 0,47$

Fonte: Do autor.

Tabela 6 – Matriz de covariância entre os parâmetros.

$\sigma^2$	$\sin^2 2\theta_{13}$	$\Delta m_{13}^2$	$a'$	$b'$	$c'$	$\alpha_{Li+He}$	$\alpha_{FN+SN}$	$\alpha_{Acc}$	$\alpha_{Res}$
$\sin^2 2\theta_{13}$	0,000882								
$\Delta m_{13}^2$	-0,000000	0,000000							
$a'$	0,000004	0,000000	0,000035						
$b'$	-0,000070	0,000000	-0,000013	0,000032					
$c'$	0,000000	-0,000000	0,000000	-0,000001	0,000000				
$\alpha_{Li+He}$	0,000803	-0,000000	0,000093	-0,000231	-0,000012	0,028362			
$\alpha_{FN+SN}$	-0,000034	-0,000000	0,000000	0,000006	-0,000000	-0,001491	0,003979		
$\alpha_{Acc}$	0,000003	-0,000000	0,000008	0,000006	-0,000000	-0,000123	-0,000042	0,005801	
$\alpha_{Res}$	0,000184	-0,000000	0,000001	-0,000018	0,000000	0,000174	0,000014	-0,000118	0,089065

Fonte: Do autor.

(4.34)

Tabela 7 – Matriz de correlação entre os parâmetros.

$\rho$	$\sin^2 2\theta_{13}$	$\Delta m_{13}^2$	$a'$	$b'$	$c'$	$\alpha_{Li+He}$	$\alpha_{FN+SN}$	$\alpha_{Acc}$	$\alpha_{Res}$
$\sin^2 2\theta_{13}$	1								
$\Delta m_{13}^2$	-0,128514	1							
$a'$	0,021638	0,027478	1						
$b'$	-0,419094	0,033212	-0,401370	1					
$c'$	0,019869	-0,021518	0,074693	-0,391470	1				
$\alpha_{Li+He}$	0,160632	-0,012000	0,093307	-0,243349	-0,118474	1			
$\alpha_{FN+SN}$	-0,018281	-0,006980	0,001226	0,015861	-0,002309	-0,140400	1		
$\alpha_{Acc}$	0,001507	-0,002198	0,017224	0,014155	-0,005312	-0,009602	-0,008836	1	
$\alpha_{Res}$	0,020798	-0,005643	0,000692	-0,010788	0,000777	0,003458	0,000744	-0,005172	1

Fonte: Do autor.

(4.35)

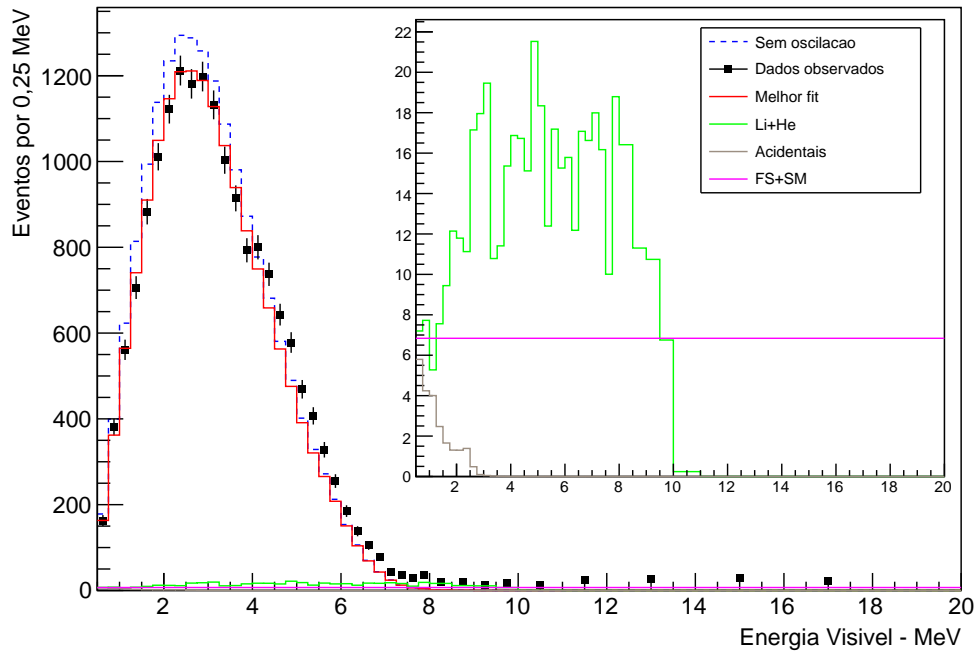


Figura 25 – A figura representa graficamente os dados de Double Chooz, fase III e o melhor ajuste feito com auxílio da biblioteca C++ BAT.

Fonte: Do autor.

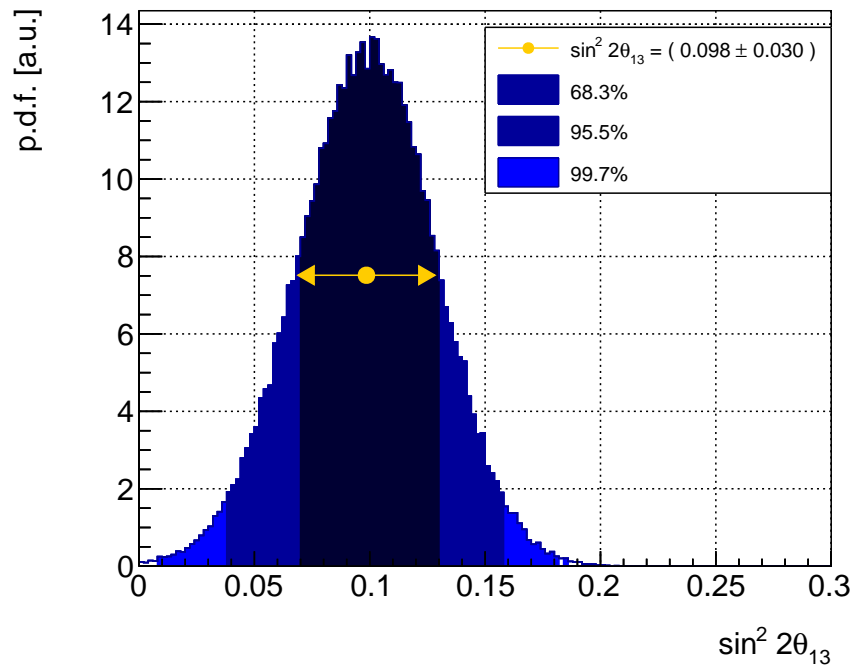


Figura 26 – P.d.f posteriori marginalizada do parâmetro  $\sin^2 2\theta_{13}$ , indicando a melhor estimativa como  $0,098 \pm 0,030$ .

Fonte: Do autor.

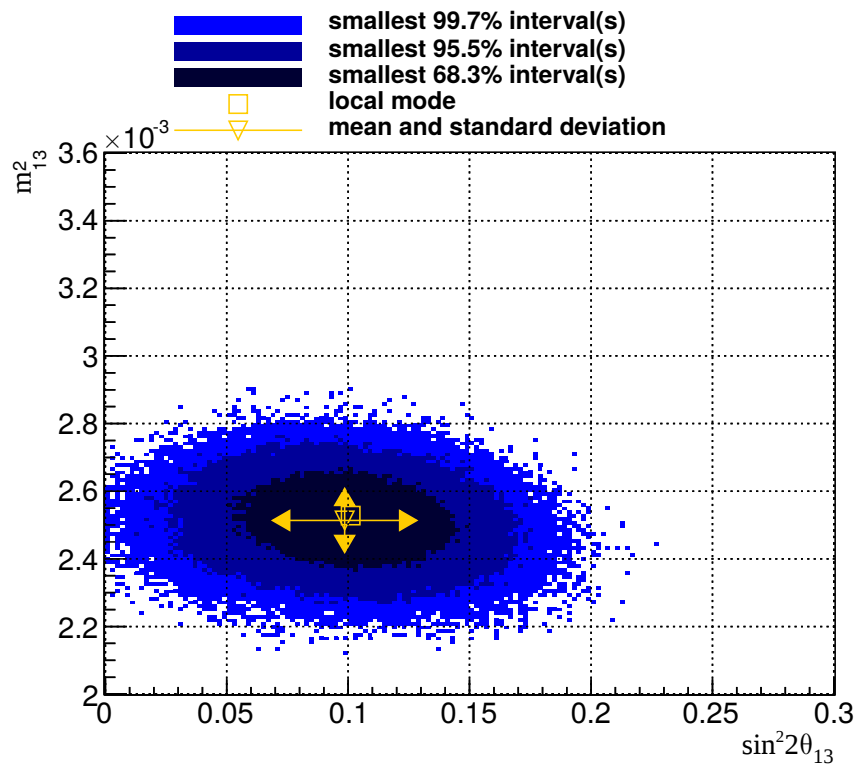


Figura 27 – Correlação entre os parâmetros  $\sin^2 2\theta_{13}$  e  $\Delta m^2_{13}$ . O gráfico mostra uma leve correlação negativa entre esses parâmetros.

Fonte: Do autor.

## 5 CONCLUSÃO

Não é sem razão que a mídia apelidou os neutrinos como partículas fantasmas, dado a grande dificuldade de sua detecção. Bem mais difícil do que detectar o neutrino, é detectá-lo junto com seu espectro de energia, onde podemos saber quantos neutrinos estão presentes em cada intervalo de energia possível. Essa grande façanha de enxergar o invisível, foi possível mediante a grande imaginação e competência dos físicos experimentais e teóricos, provenientes de todas partes do mundo, inclusive do Brasil, na Colaboração Double Chooz.

Nosso trabalho começa no capítulo 1 onde tratamos com a compreensão elementar do processo de oscilação dos neutrinos. Vimos que podemos chegar a frequência de oscilação, abordando o problema com mecânica quântica básica. A oscilação ocorre por que dois observáveis básicos na descrição dos neutrinos, simplesmente não comutam. Como vimos, são os observáveis *sabor* e *massa*. A oscilação de sabor do neutrino está diretamente ligada ao número de neutrinos observados em cada intervalo de energia, que é um dos dados determinados pelo experimento DC.

O método que o experimento DC utiliza para determinar os eventos de neutrinos e os outros eventos que não podem ser distinguidos de neutrinos, está baseado em sincronismos desses eventos. Você detecta um sincronismo e a partir dele faz série de varreduras em uma grande quantidade de informações. Todo o conjunto de informação é então analisado com estatística bayesiana, por isso tivemos o capítulo 2, apresentando de forma sucinta tudo que era necessário para tratar com o problema estatístico de identificar o neutrino. Nesse capítulo também encontramos exemplos simples, resolvidos com Kit de ferramentas BAT.

No capítulo 3 fizemos a descrição do modo operante do experimento em si. No capítulo 4 é o momento em que todas as informações são reunidas com o objetivo de determinar o ângulo de mistura  $\theta_{13}$ . Nossa análise se baseou numa série de informações disponíveis nas notas técnicas do experimento(26), e também, é claro, nos arquivos contendo os dados reais medidos pela Colaboração. O resultado principal de todo nosso trabalho pode ser resumido com

$$\sin^2 2\theta_{13} = 0,098 \pm 0,030 \quad (5.1)$$

Sendo o valor do ângulo de mistura em graus dado por

$$\theta_{13} = (9,1 \pm 1,5)^\circ \quad (5.2)$$

O resultado encontrado apresenta uma margem de erro em torno de 31% para o parâmetro  $\sin^2 2\theta_{13}$  o que é uma margem grande comparada com a margem de erro de outros experimentos, como temos indicado na figura 16, página 47. Entretanto, quando fazemos análise bayesiana temos uma vantagem que a análise frequentista não possui. Ela fornece uma base matemática que permite incorporar na análise, a compreensão subjetiva que

pesquisadores possuem sobre as distribuições de probabilidades relacionadas com os parâmetros em estudo. Temos experiência que tais compreensões frequentemente estavam presentes em grandes descobertas na Física.

## REFERÊNCIAS

- 1 PAULI, Wolfgang. *Liebe radioaktive damen und herren*. 1930. 1 p. Disponível em: <<https://www.bibnum.education.fr/physique/physique-nucleaire/chers-mesdames-et-messieurs-radioactifs>>. Acesso em: 10 abr 2016.
- 2 VALDIVIESSO, Gustavo A.; GUZZO, Marcelo M. Compreendendo a oscilação dos neutrinos. *Revista Brasileira de Ensino de Física*, scielo, v. 27, p. 495–506, 12 2005. ISSN 1806-1117. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1806-11172005000400002&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-11172005000400002&nrm=iso)>. Acesso em: 10 abr 2016.
- 3 CLEVELAND, B. T.; T., Daily; DAVIS, R. Measurement of the solar electron neutrino flux with the homestake chlorine detector. *The Astrophysical Journal*, v. 496, p. 505–526, Mar 1998.
- 4 ABDURASHITOV, J. N. Solar neutrino flux measurements by the soviet-american gallium experiment (sage) for half the 22-year solar cycle. *Journal of Experimental and Theoretical Physics*, v. 95(2), p. 181–193, Aug 2002.
- 5 ALTMANN, M. Complete results for five years of gno solar neutrino observations. *Phys. Lett. B*, v. 616, p. 174–190, Apr 2005.
- 6 VALDIVIESSO, Gustavo A. *Introdução à fenomenologia da oscilação de neutrinos, no vácuo e na matéria*. 114 p. Dissertação (Mestrado em Física) — Universidade Estadual de Campinas - UNICAMP, 2004.
- 7 AHARMIM, B. et al. Determination of the  $\nu_e$  and total  ${}^8\text{B}$  solar neutrino fluxes using the sudbury neutrino observatory phase i data set. *Phys. Rev. C*, American Physical Society, v. 75, p. 045502, Apr 2007. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevC.75.045502>>. Acesso em: 10 abr 2016.
- 8 OLIVE, K. A. et al. Review of particle physics. *Chin. Phys.*, C38, p. 090001, 2014.
- 9 ADAMSON, P. et al. *Precision measurement of the speed of propagation of neutrinos using the MINOS detectors*. 2015.
- 10 WIKIPÉDIA. *Família dos léptons*. 2016. Disponível em: <[https://en.wikipedia.org/wiki/Elementary\\_particle](https://en.wikipedia.org/wiki/Elementary_particle)>. Acesso em: 10 abr 2016.
- 11 ABE, Y. et al. Indication of reactor  $\bar{\nu}_e$  disappearance in the double chooz experiment. *Phys. Rev. Lett.*, American Physical Society, v. 108, p. 131801, Mar 2012. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevLett.108.131801>>. Acesso em: 10 abr 2016.
- 12 ROOT – An object oriented data analysis framework. 2015. Disponível em: <<https://root.cern.ch>>. Acesso em: 10 abr 2016.
- 13 COWAN, Glen. *Statistical data analysis*. [S.l.]: Oxford University Press, 1998.
- 14 VUOLO, José Henrique. *Fundamentos da teoria de erros*. 2. ed. São Paulo: Edgard Blücher LTDA, 1996.



- 15 CAVAINAC, J.F. et al. Indication for neutrino oscillation from a high statistics experiment at the bugey reactor. *Physics Letters B*, v. 148, n. 4–5, p. 387 – 394, 1984. ISSN 0370-2693. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0370269384901096>>. Acesso em: 10 abr 2016.
- 16 ACHKAR, B. et al. Search for neutrino oscillations at 15-meters, 40-meters, and 95-meters from a nuclear power reactor at bugey. *Nucl. Phys.*, B434, p. 503–534, 1995.
- 17 APOLLONIO, M et al. Initial results from the {CHOOZ} long baseline reactor neutrino oscillation experiment. *Physics Letters B*, v. 420, n. 3–4, p. 397 – 404, 1998. ISSN 0370-2693. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0370269397014767>>. Acesso em: 10 abr 2016.
- 18 APOLLONIO, M. et al. Limits on neutrino oscillations from the {CHOOZ} experiment. *Physics Letters B*, v. 466, n. 2-4, p. 415–430, 1999. ISSN 0370-2693. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0370269399010722>>. Acesso em: 10 abr 2016.
- 19 ARDELLIER, F. et al. Letter of intent for double chooz: a search for the neutrino mixing angle  $\theta_{13}$ . *arXiv:hep-ex/0405032*, 2004.
- 20 ABE, Y. et al. Improved measurements of the neutrino mixing angle  $\theta_{13}$  with the double chooz detector. *Journal of High Energy Physics*, v. 2014, n. 10, p. 1–44, 2014. ISSN 1029-8479. Disponível em: <[http://dx.doi.org/10.1007/JHEP10\(2014\)086](http://dx.doi.org/10.1007/JHEP10(2014)086)>. Acesso em: 10 abr 2016.
- 21 KIM, Soo-Bong. *Status and prospects of reactor neutrino experiments*. 2015. Disponível em: <<http://arxiv.org/pdf/1504.08268.pdf>>. Acesso em: 10 abr 2016.
- 22 ABE, Y. et al. Ortho-positronium observation in the double chooz experiment. *Journal of High Energy Physics*, 2014.
- 23 ABE, Y. et al. Reactor  $\bar{\nu}_e$  disappearance in the double chooz experiment. *Phys. Rev. D*, American Physical Society, v. 86, p. 052008, Sep 2012. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevD.86.052008>>. Acesso em: 10 abr 2016.
- 24 ABE, Y. et al. First measurement of from delayed neutron capture on hydrogen in the double chooz experiment. *Physics Letters B*, v. 723, n. 1–3, p. 66 – 70, 2013. ISSN 0370-2693. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0370269313003390>>. Acesso em: 10 abr 2016.
- 25 ABE, Y. et al. Measurement of  $\theta_{13}$  in double chooz using neutron captures on hydrogen with novel background rejection techniques. *Journal of High Energy Physics*, v. 2016, n. 1, p. 1–29, 2016. ISSN 1029-8479. Disponível em: <[http://dx.doi.org/10.1007/JHEP01\(2016\)163](http://dx.doi.org/10.1007/JHEP01(2016)163)>. Acesso em: 10 abr 2016.
- 26 THE DOUBLE CHOOZ ANALYSIS GROUP. *Internal note of DC-III electron antineutrino disappearance analyses using energy spectrum*. [S.l.], 2014.
- 27 BAT – The Bayesian analysis toolkit. 2015. Disponível em: <<http://mpp.mpg.de/bat>>. Acesso em: 10 abr 2016.