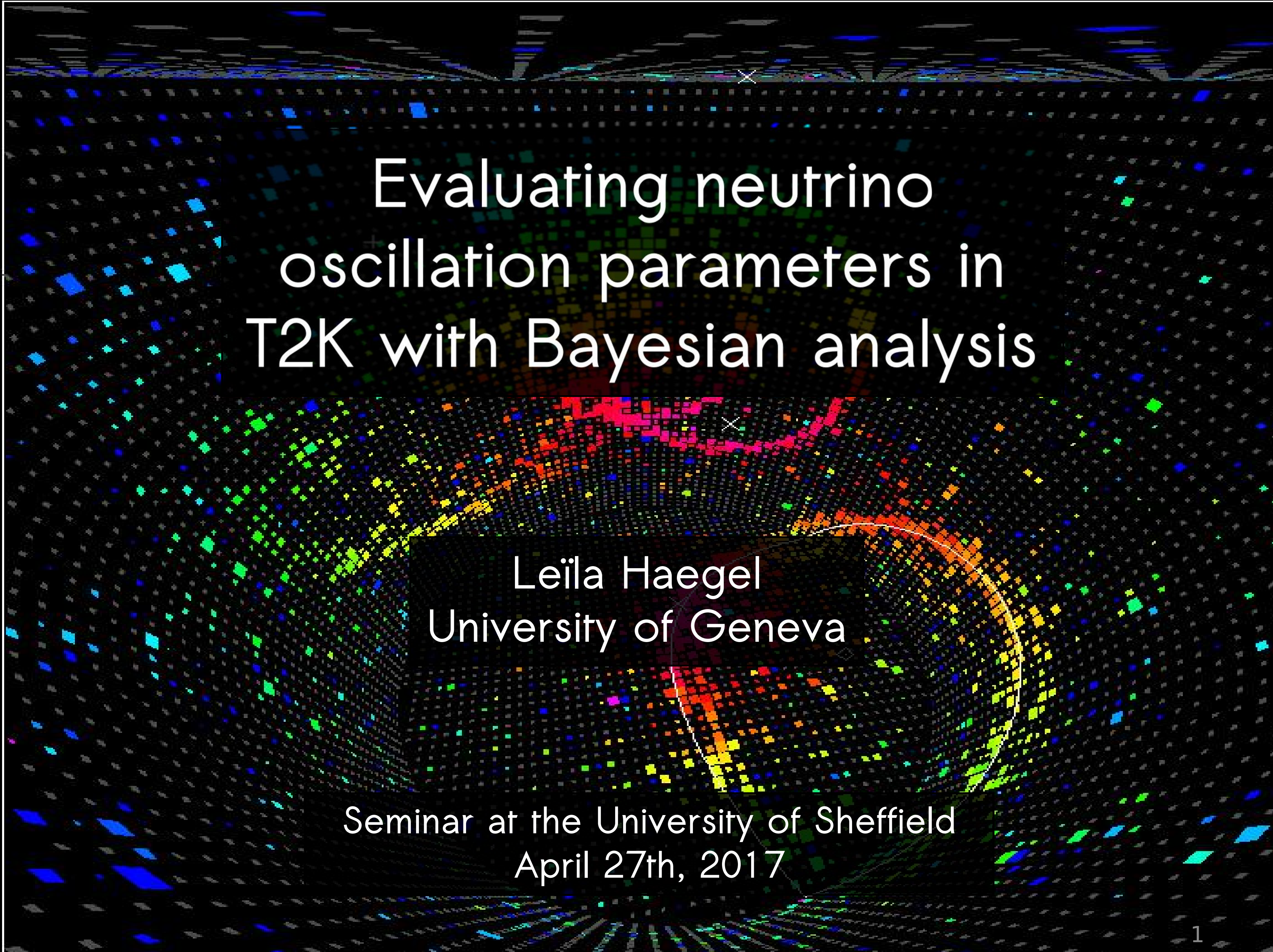


Evaluating neutrino oscillation parameters in T2K with Bayesian analysis

Leïla Haegel
University of Geneva

Seminar at the University of Sheffield
April 27th, 2017

Picture found at: http://www.ps.uci.edu/~tomba/sk/tscan/pictures.html

Neutrinos — What we know

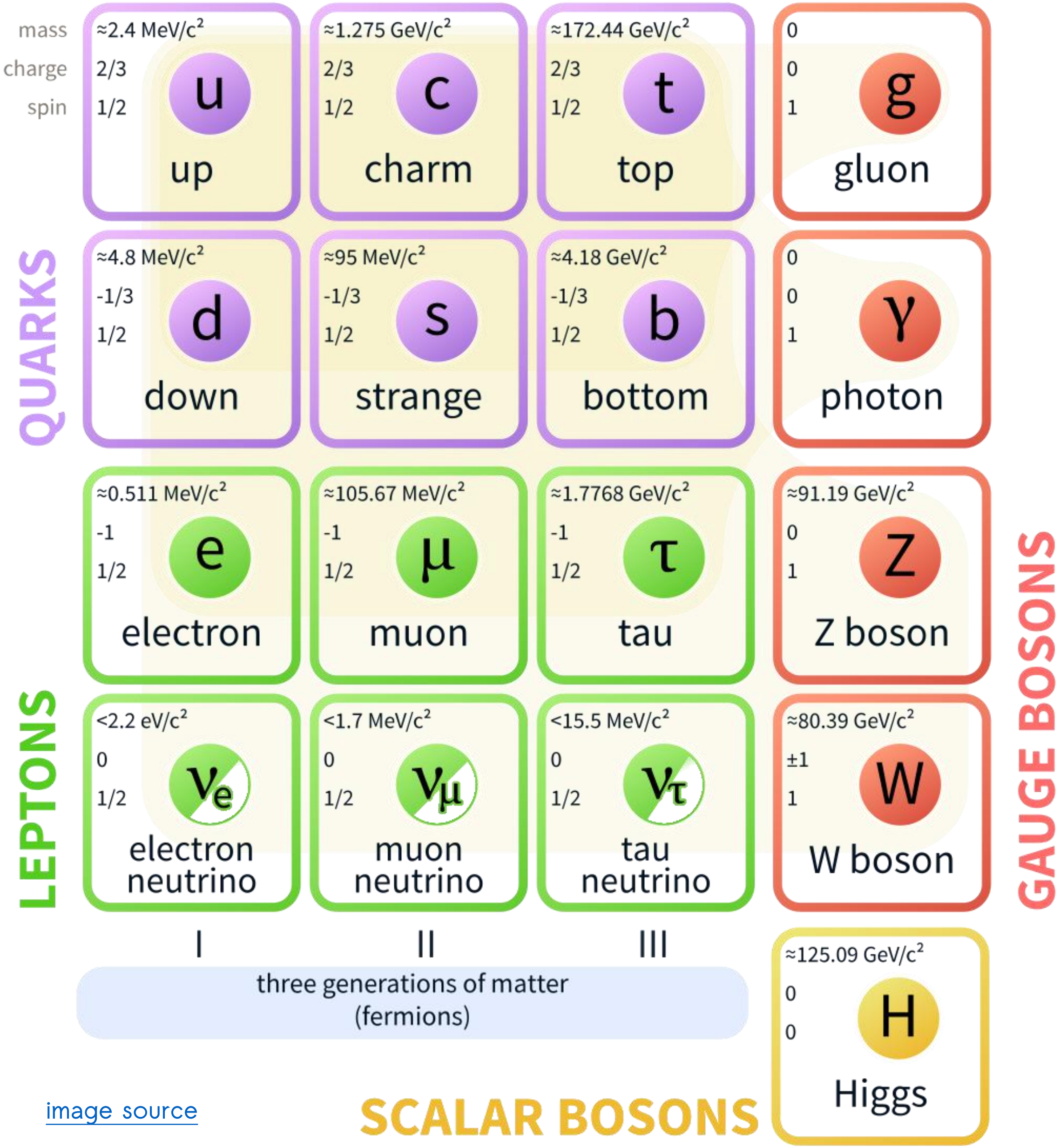


image source

Neutrinos are:

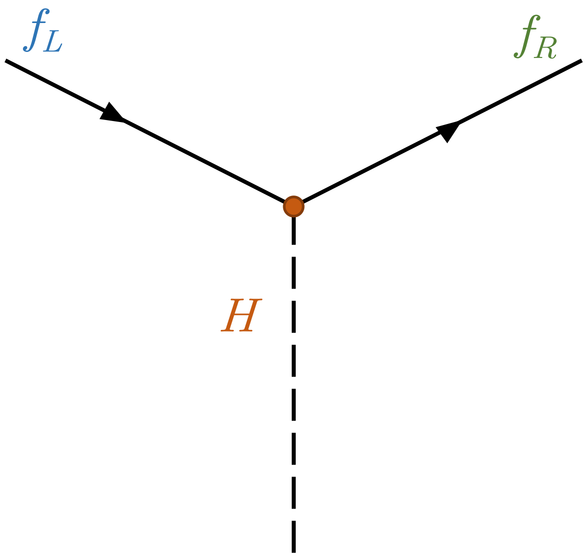
- the only particle known with only left-handed chirality;
- the only particle interacting only through weak interaction;
- the lightest fermions of the Standard Model;
- the only leptons that oscillate. i.e. can be detected in another flavour state than created.

- How do neutrinos aquire mass ?

    Higgs coupling change the chirality of the particles.

    But neutrinos only exist left-handed !

    Are they (non-interacting ?) right-handed neutrinos ?



$$\mathcal{L}_y = \underline{g_y}\ \underline{\bar{\psi}(x)_L}\ \underline{\phi(x)}\ \underline{\psi(x)_R} + h.c.$$

- How do neutrinos acquire mass ?
- Are neutrinos Majorana particles ?

  Majorana particles are they own antiparticles from the action of a charge conjugate operator.

  Adding a right-handed Majorana singlet N_R can explain how neutrinos acquire mass, and why the neutrino mass is small in the "see-saw" model.

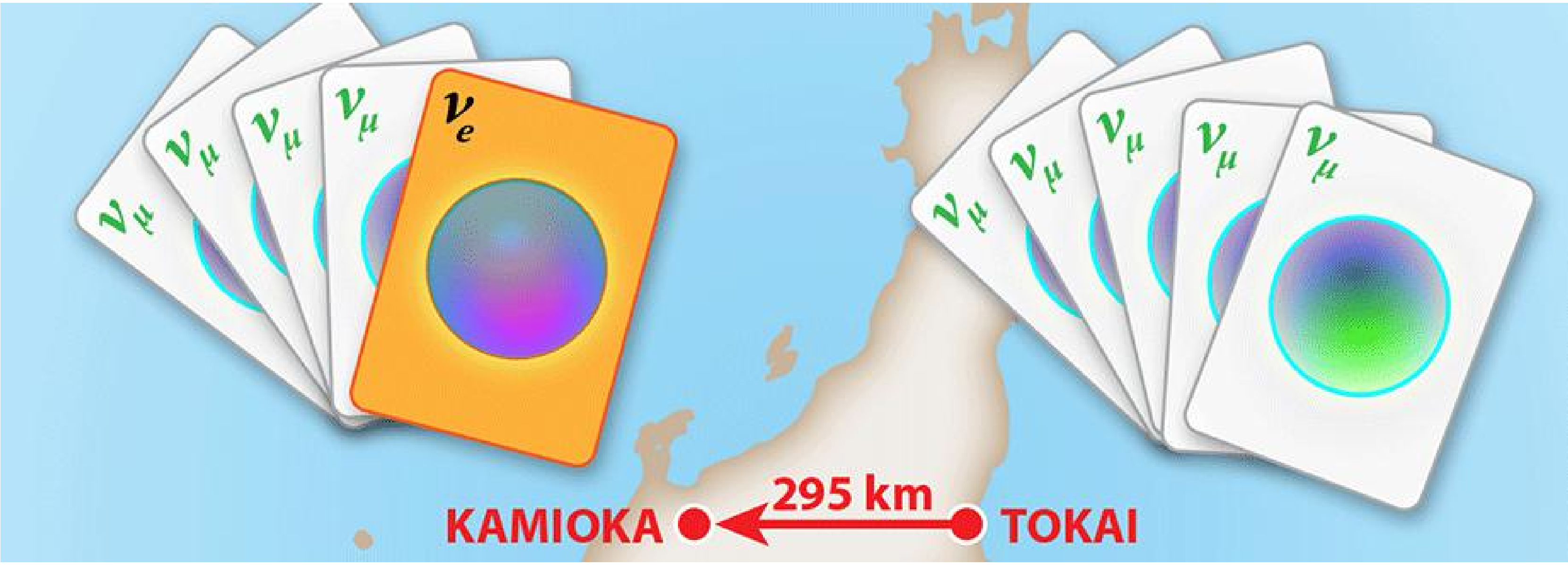
$$\begin{matrix} T_3 = \frac{1}{2} \\ T_3 = -\frac{1}{2} \end{matrix} \quad \underbrace{\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L}_{\text{leptons } l_L}$$

$$\begin{matrix} T_3 = 0 \\ T_3 = 0 \end{matrix} \quad \underbrace{(e^-)_R \;\; (\mu^-)_R f \;\; (\tau^-)_R}_{\text{leptons } l_R}$$

$$\psi_{lepton} = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \nu_L \\ l_L \\ N_R \\ l_R \end{pmatrix}$$

What we don't know

- How do neutrinos acquire mass ?
- Are neutrinos Majorana particles ?
- How do neutrinos oscillate ?
  - What are the values of the oscillation parameters ?
  - Are they free values or do they come from breaking a higher symmetry ?
  - Do neutrinos and antineutrinos oscillate the same way ?



- Mass eigenstates do not coincide with flavour eigenstates.
- Phase shifts occur while they propagate in time.
- The probability to detect a certain flavour eigenstate is encoded by the rotation matrix U_{PMNS}.

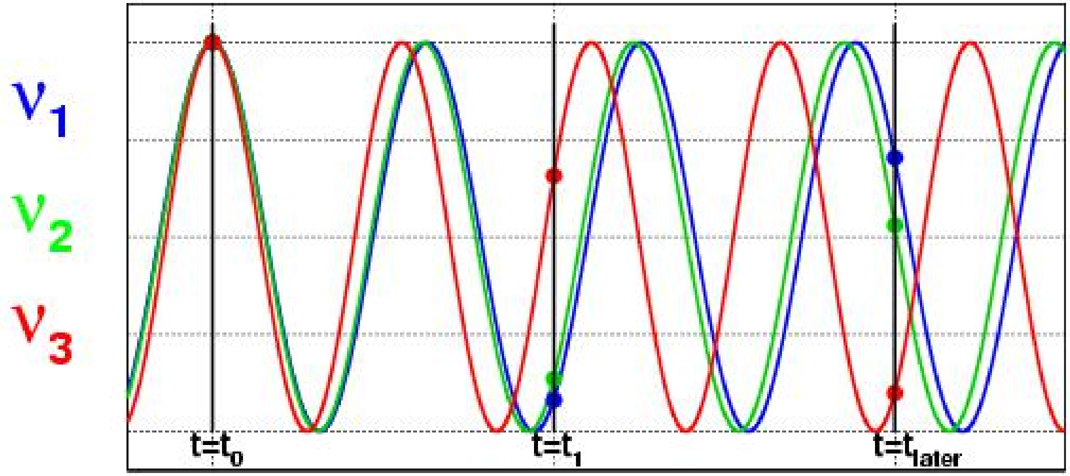


image source

$$\underbrace{\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \\ |\nu_\tau\rangle \end{pmatrix}}_{\nu_\alpha} = \underbrace{\begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix}}_{U_{\mathrm{PMNS}}} \underbrace{\begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix}}_{\nu_i}$$

$$\underbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{solar} \underbrace{\begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 0 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix}}_{accelerator+reactor} \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}}_{atmospheric} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\frac{\alpha_{21}}{2}} & 0 \\ 0 & 0 & e^{i\frac{\alpha_{31}}{2}} \end{pmatrix}}_{Majorana}$$

- Accelerator neutrino experiments probe oscillation through $\overset{(-)}{\nu_\mu}$ disappearance and $\overset{(-)}{\nu_e}$ appearance:

$$P(\overset{(-)}{\nu_\mu} \rightarrow \overset{(-)}{\nu_e}) = s_{23}^2\ s_{13}^2 \left(\frac{\Delta_{13}}{A \pm \Delta_{13}}\right)^2 \sin^2 \frac{|A \pm \Delta_{13}| + L}{2}$$

atmospheric term

$$+ c_{13}^2\ \sin^2 2\theta_{12} \left(\frac{\Delta_{12}}{A}\right)^2 \sin^2 \frac{AL}{2}$$

solar term

CP conserving term

$$+ J\ \cos\delta \left(\frac{\Delta_{12}}{A}\right) \left(\frac{\Delta_{13}}{|A \pm \Delta_{13}|}\right) \cos \frac{\Delta_{13} L}{2} \sin \frac{AL}{2} \sin \frac{|A \pm \Delta_{13}| + L}{2}$$

CP violating term

$$\mp J\ \sin\delta \left(\frac{\Delta_{12}}{A}\right) \left(\frac{\Delta_{13}}{|A \pm \Delta_{13}|}\right) \sin \frac{\Delta_{13} L}{2} \sin \frac{AL}{2} \sin \frac{|A \pm \Delta_{13}| + L}{2}$$

$\pm$ and $\mp$ are for ν or $\bar{\nu}$

$$A = \sqrt{2} G_F N_e$$

$$\Delta_{ij} = \frac{\Delta m_{ij}^2}{2E}$$

$$J = c_{13}^2\ \sin^2 2\theta_{12}\ \sin^2 2\theta_{13}\ \sin^2 2\theta_{23}$$

- We don't know if U_{PMNS} is a unitary matrix.

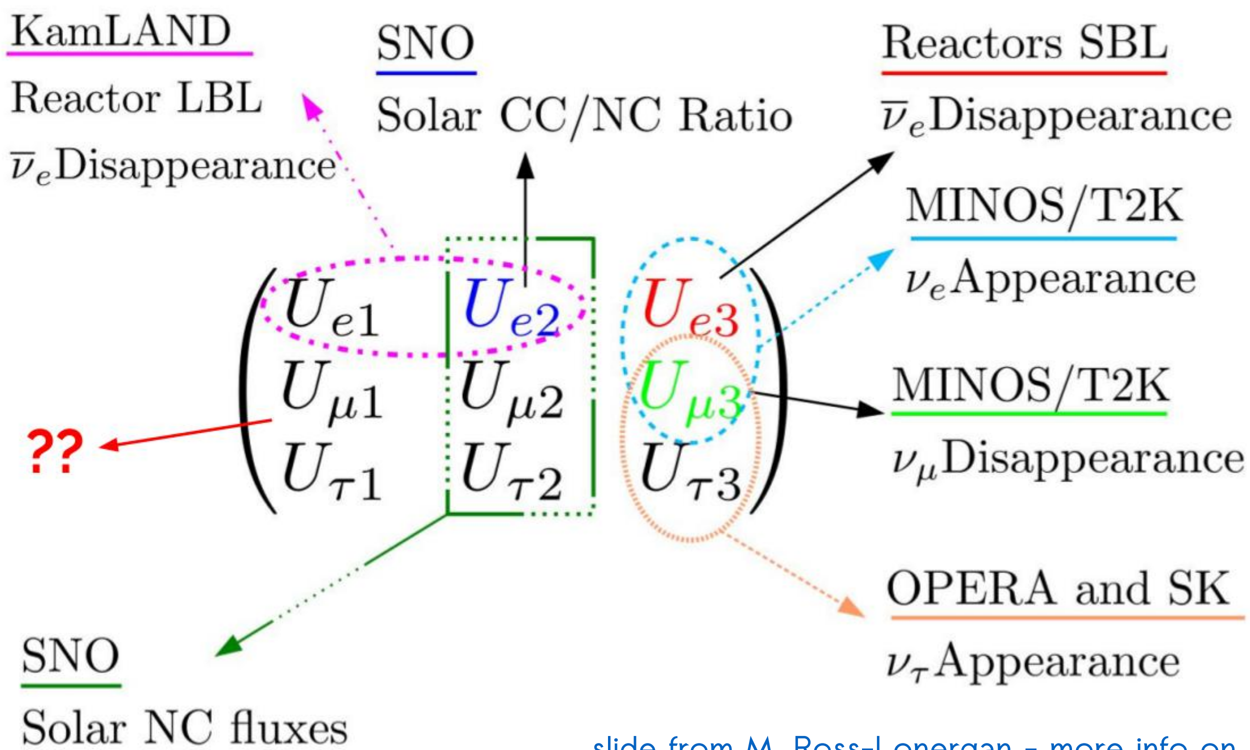
  If not, indication of mixing with another neutrino.

  Cosmological fits see 3 weekly interacting neutrinos.

  More neutrinos could be right-handed neutrinos

  ➡ solve the origin of neutrino mass issue if Majorana

  ➡ can be a dark matter candidate



slide from M. Ross-Lonergan - more info on PMNS unitarity here

- We don't know if U_{PMNS} is a unitary matrix;
- We don't know if the U_{PMNS} values come from a higher symmetry

  e.g. symmetry breaking of flavour symmetry predicts the matrix value

  models require accurate measurements of the oscillation parameters to be ruled out

Bimaximal mixing (S_4)

$$U_{\rm BM} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Tri-bimaximal mixing ($A_4/T', S_4$)

$$U_{\rm TBM} = \begin{pmatrix} \sqrt{\frac{2}{3}} & \sqrt{\frac{1}{3}} & 0 \\ -\sqrt{\frac{1}{6}} & \sqrt{\frac{1}{3}} & -\sqrt{\frac{1}{2}} \\ -\sqrt{\frac{1}{6}} & \sqrt{\frac{1}{3}} & \sqrt{\frac{1}{2}} \end{pmatrix}$$

slide from A. Titov - more info on flavour symmetry breaking

- We don't know if U_{PMNS} is a unitary matrix;
- We don't know if the U_{PMNS} values come from a higher symmetry
- CP violation could be an ingredient to explain the disparition of antimatter

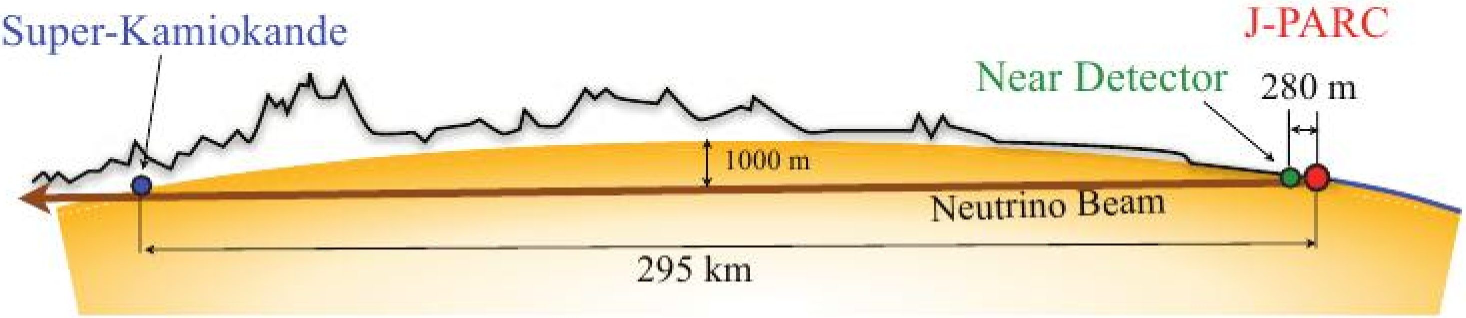
  Sakharov conditions require: ➡ baryon number violation
  ➡ CP violation
  ➡ out of thermal equilibrium interactions

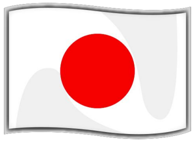
  CP violation in baryon sector (e.g. K decay) is not sufficient.
  Leptogenesis process postulates that a lepton number violation with CP violation in the neutrino sector can be transfered into baryogenesis.

baryon number asymmetry

$$\eta_\ell = \frac{n_\ell - n_{\bar{\ell}}}{n_\gamma} \in [5.8\ ;\ 6.6 \cdot 10^{-10}]$$

nucleosynthesis in the PDG
a paper about leptogenesis



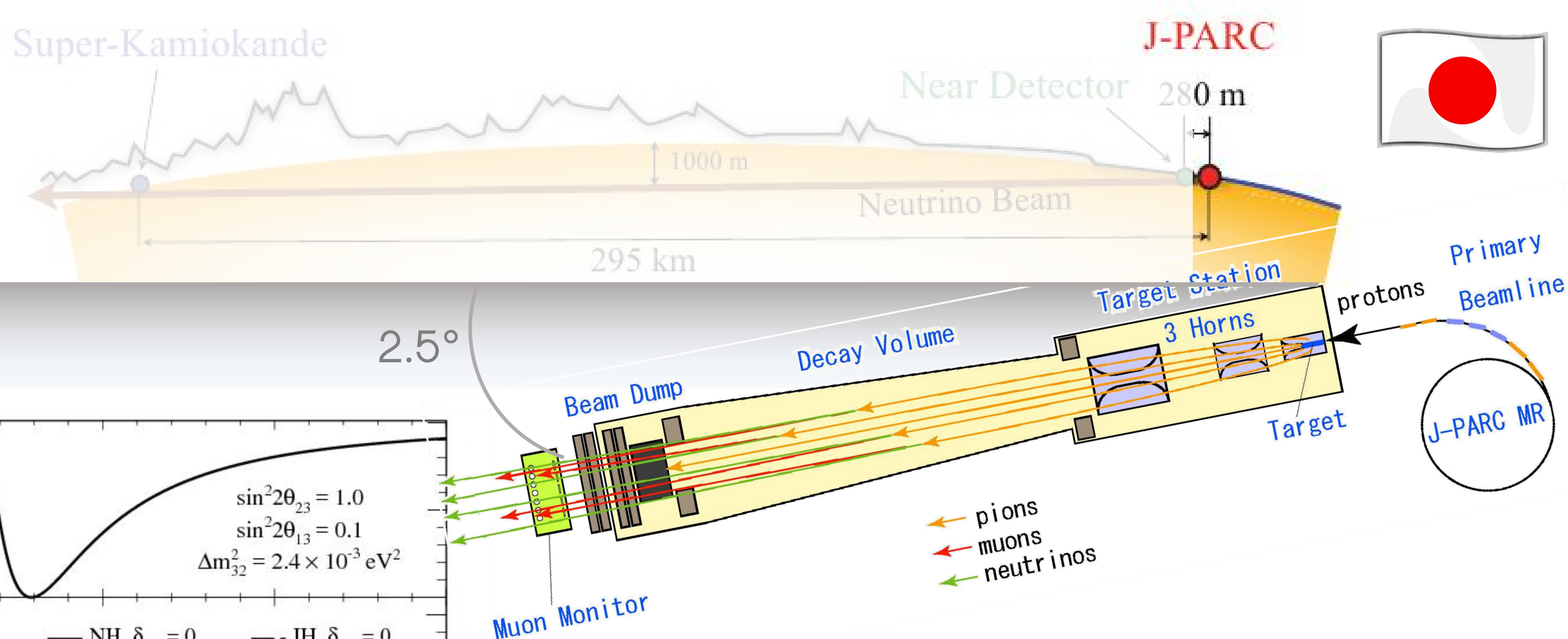


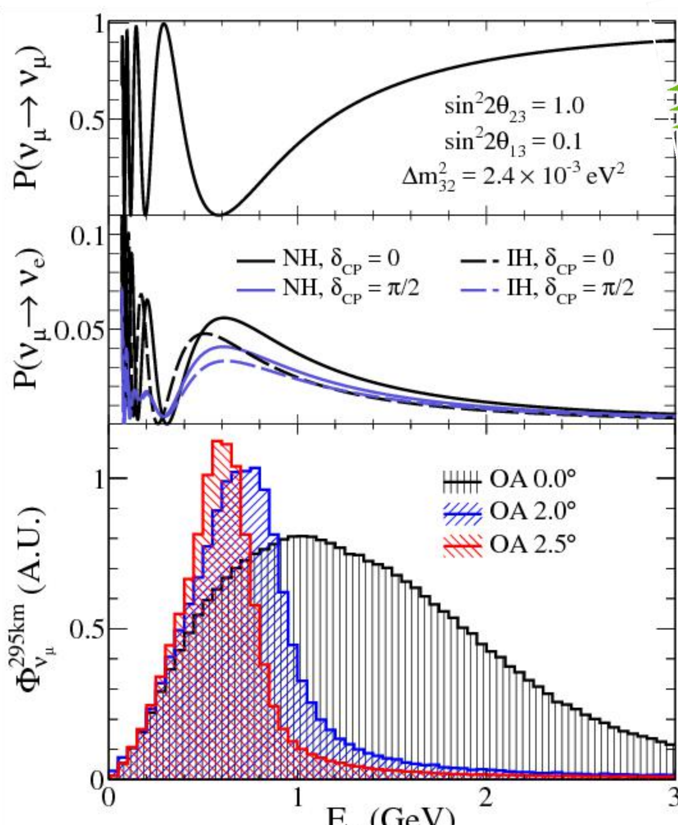
- Long-baseline ($\mathrm{L{=}295\ Km}$) neutrino oscillation experiment.
- Built with aim to measure $\sin^2\theta_{23}$ with great precision.
- First indication of non-0 $\sin^2\theta_{13}$

  Confirmation by reactor experiment opens the door to δ_{CP} measurement.

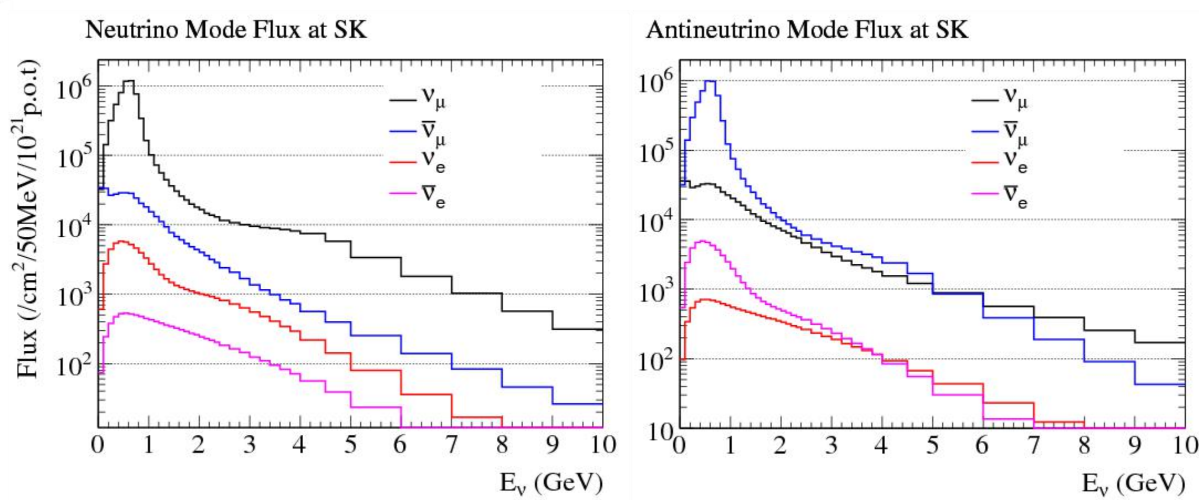
world leading measurement !

The neutrino beam

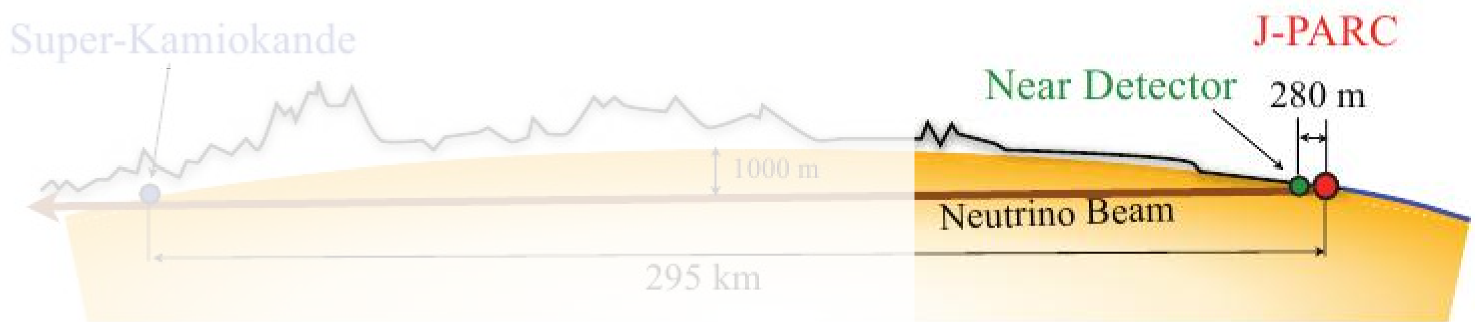


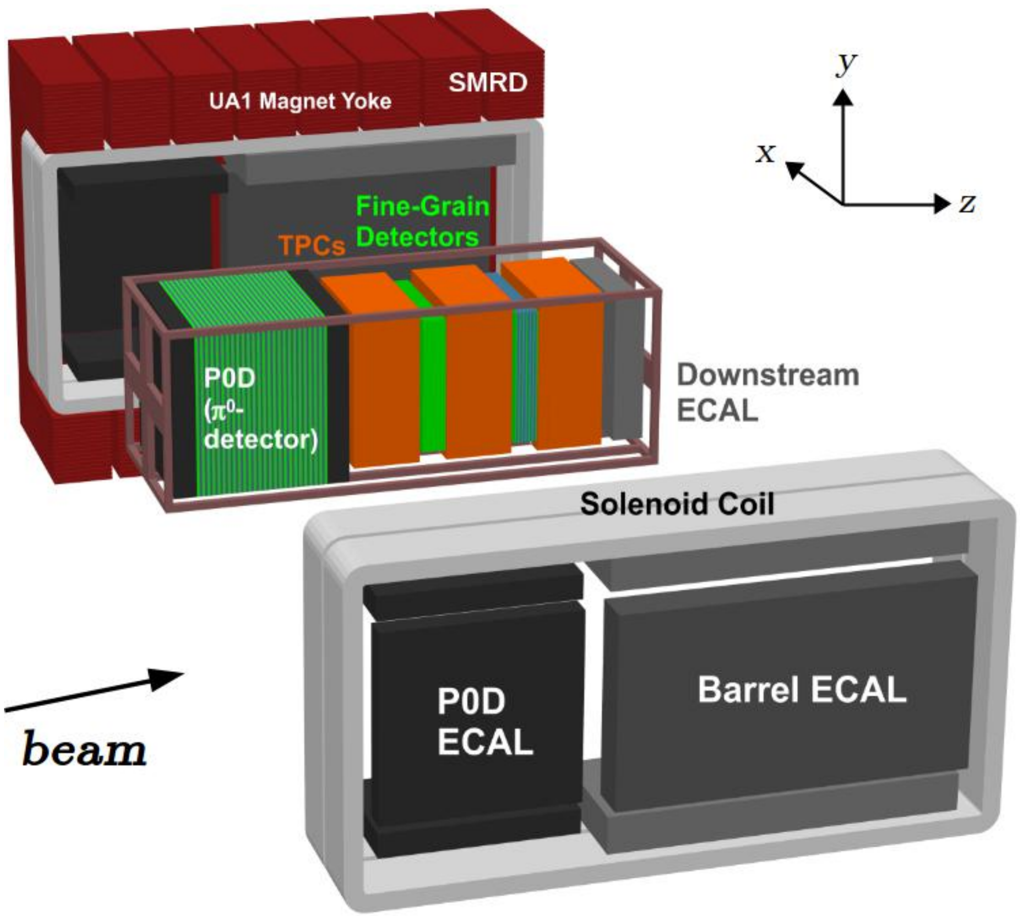


can create ν_μ or $\overline{\nu_\mu}$ enhanced beam



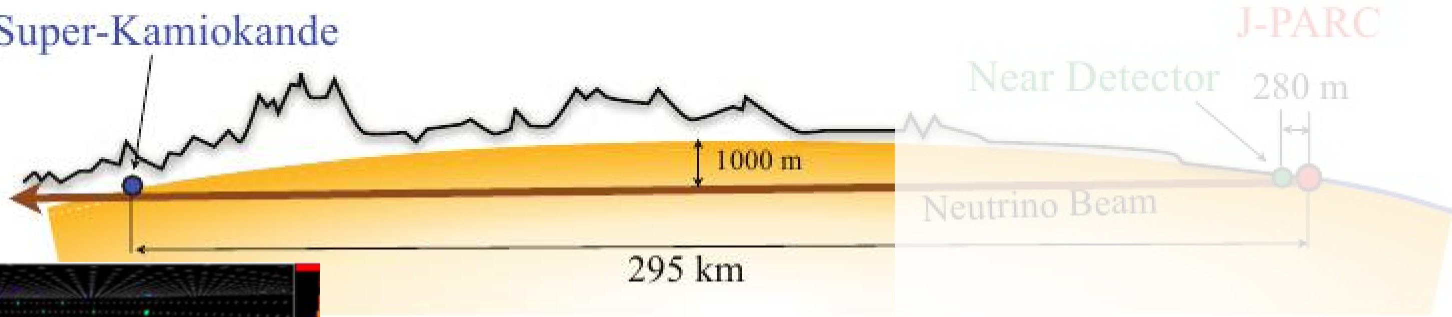
The near detector ND280

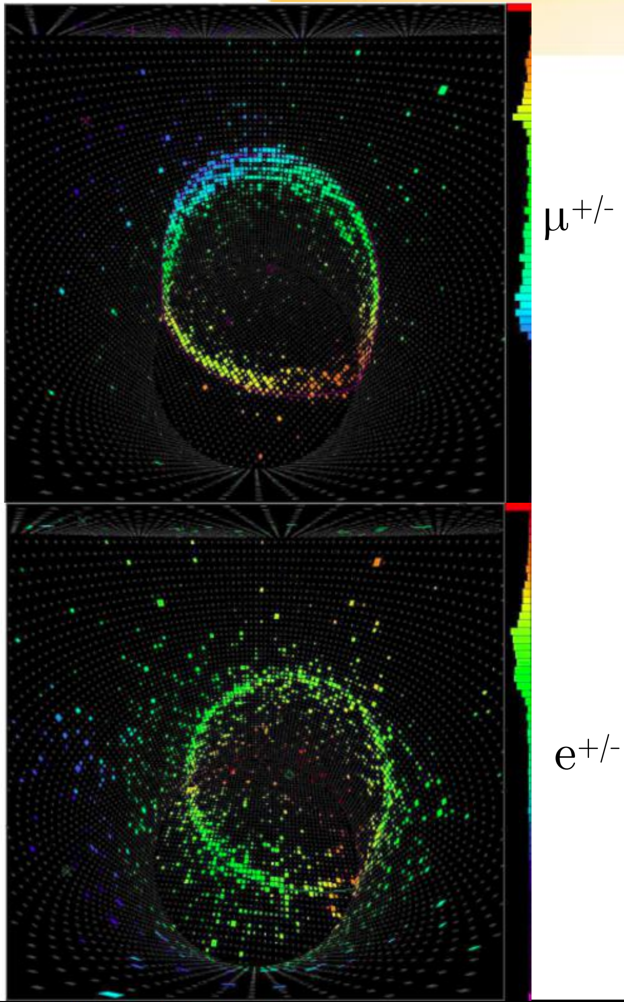




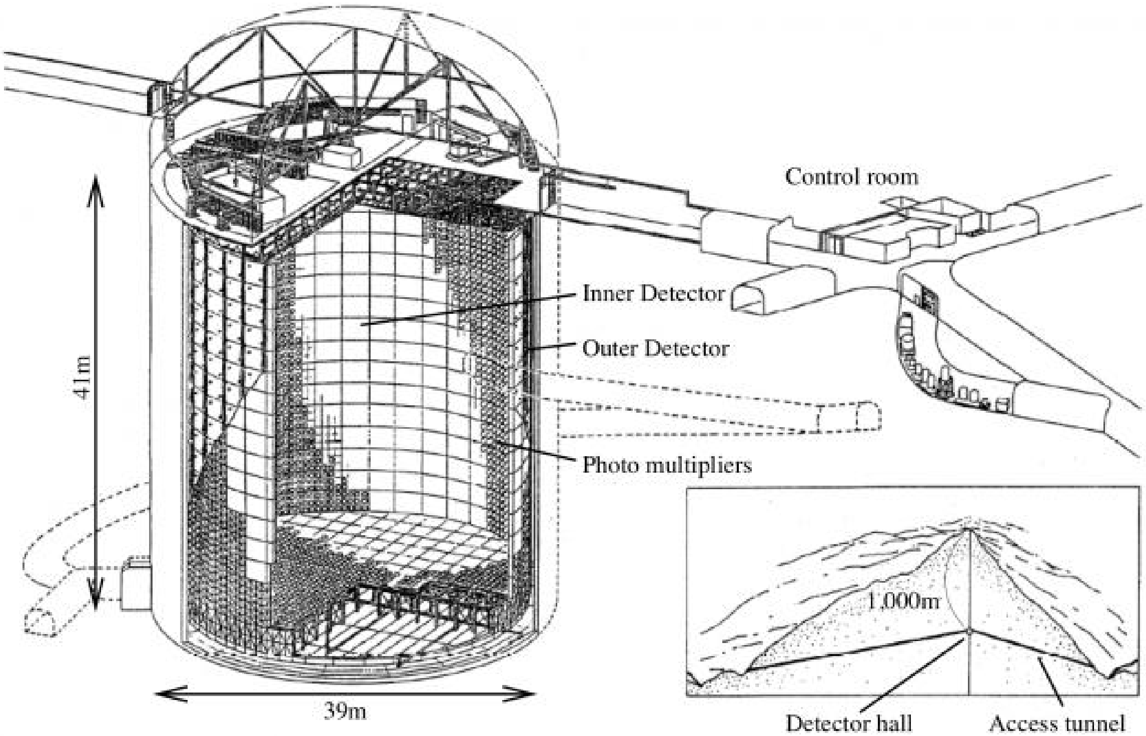
- constrains flux and cross-section model
- target detector is FGD, TPC give PID, magnetic field give charge
- interactions simulated with custom MC generator NEUT
- also allows cross-section measurements

The far detector Super-K



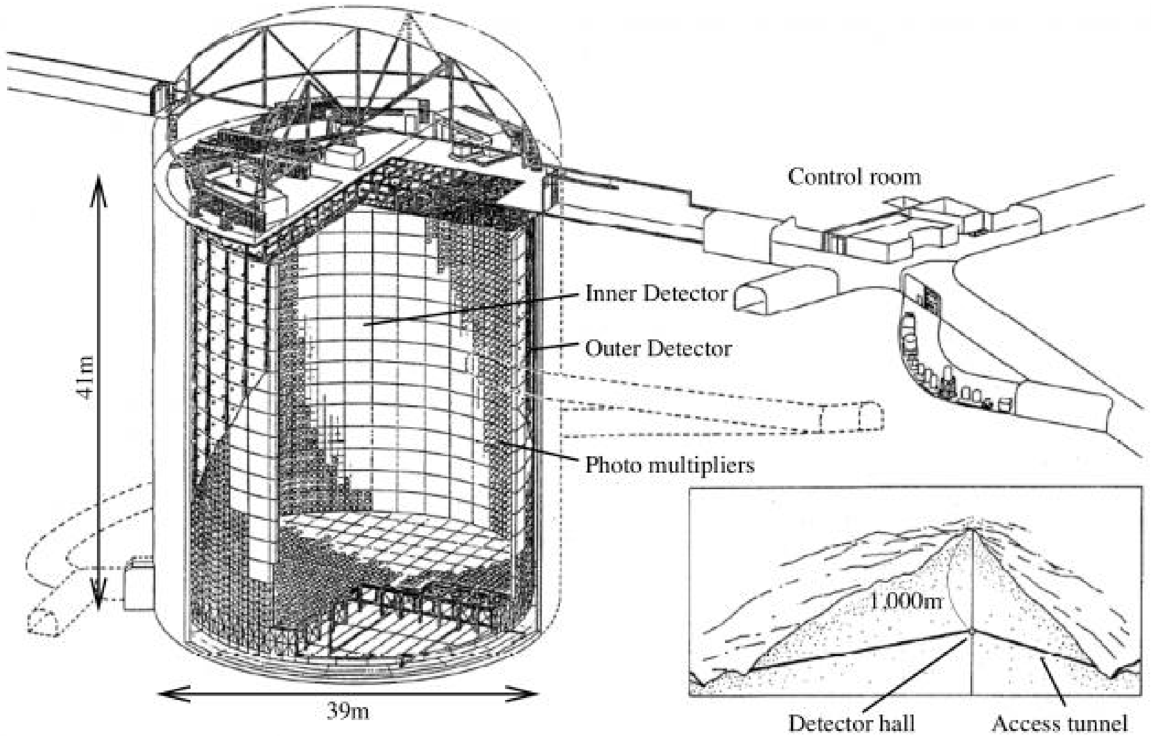


- 50 KTon Cherenkov detector
- granularity of the Cherenkov rings give PID



The dawn of all analysis:

compare data to prediction (MC) and vary prediction (e.g. oscillation parameters) to find which one agree with data best

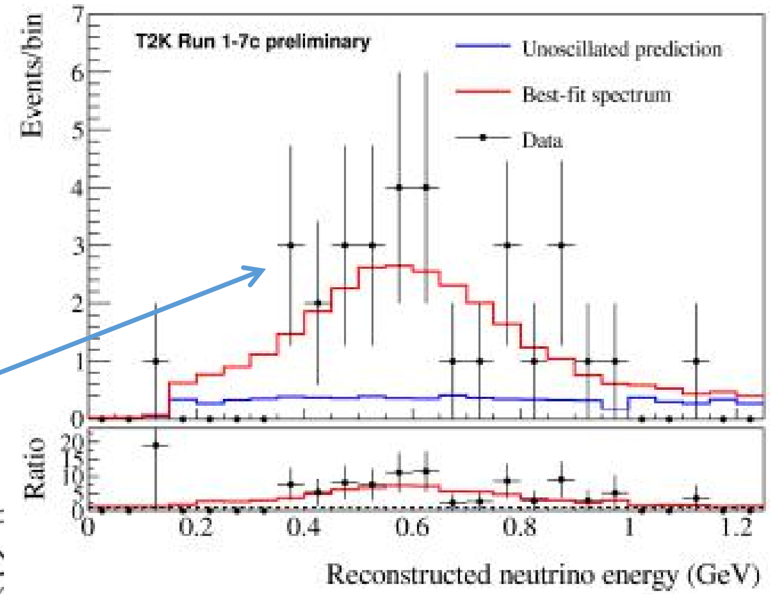


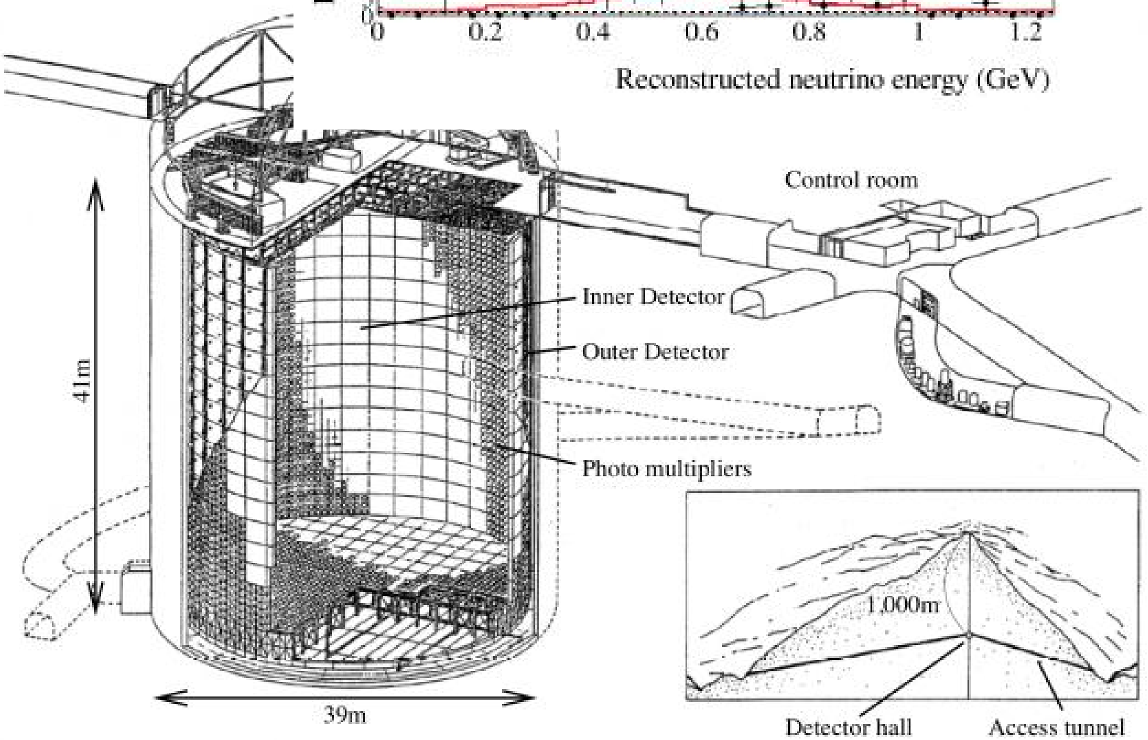
The dawn of all analysis:

compare data to prediction (MC) and vary prediction (e.g. oscillation parameters) to find which one agree with data best

$e^{+/-}$ rings CC0π

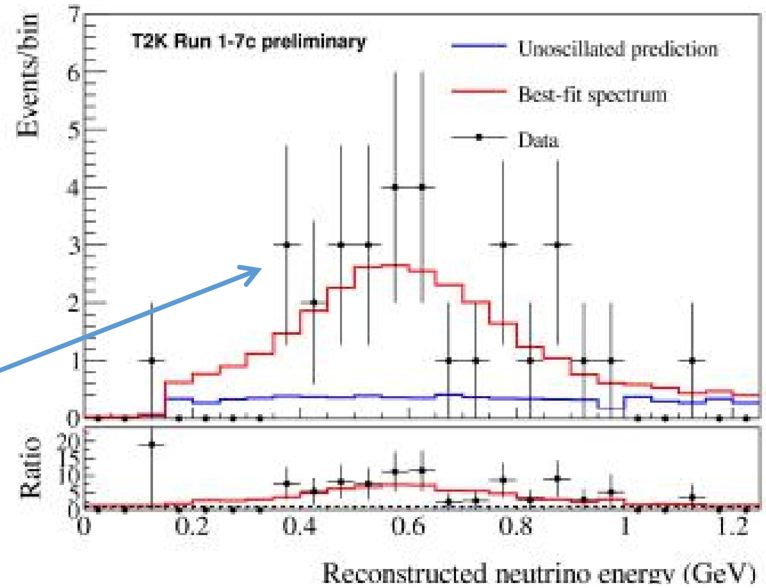
peak ∝ $\sin^2\theta_{13}$
δ_{CP}
$\mathrm{sign}(\Delta m_{23}{}^2)$





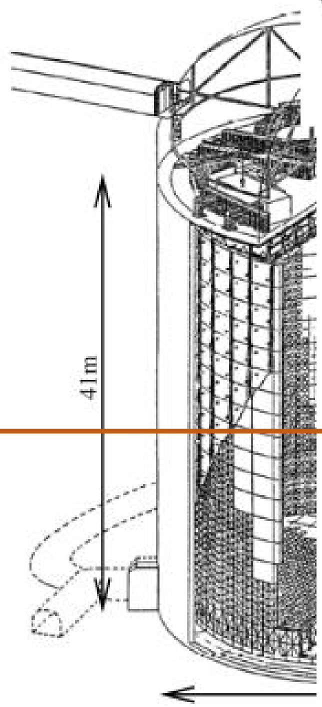
The dawn of all analysis:

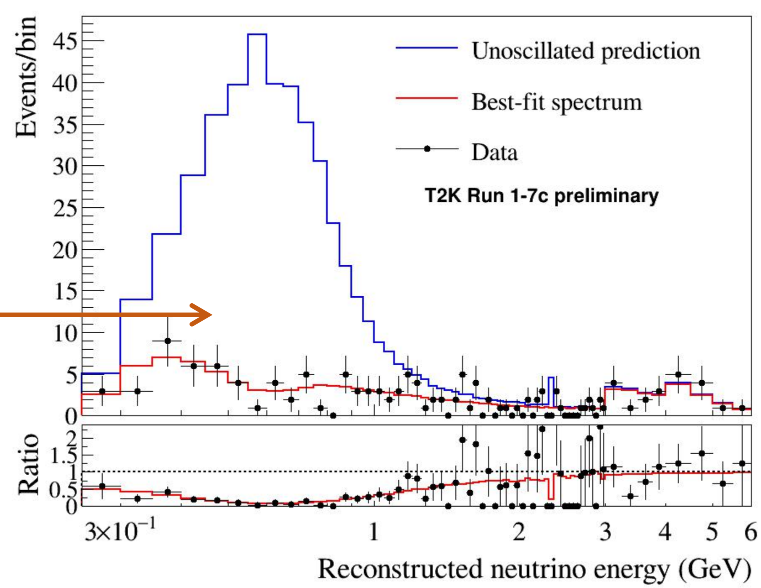
compare data to prediction (MC) and vary prediction (e.g. oscillation parameters) to find which one agree with data best



$e^{+/-}$ rings CC0π

peak $\propto$ $\sin^2\theta_{13}$
δ_{CP}
$\mathrm{sign}(\Delta m_{23}^2)$





$\mu^{+/-}$ rings CC0π

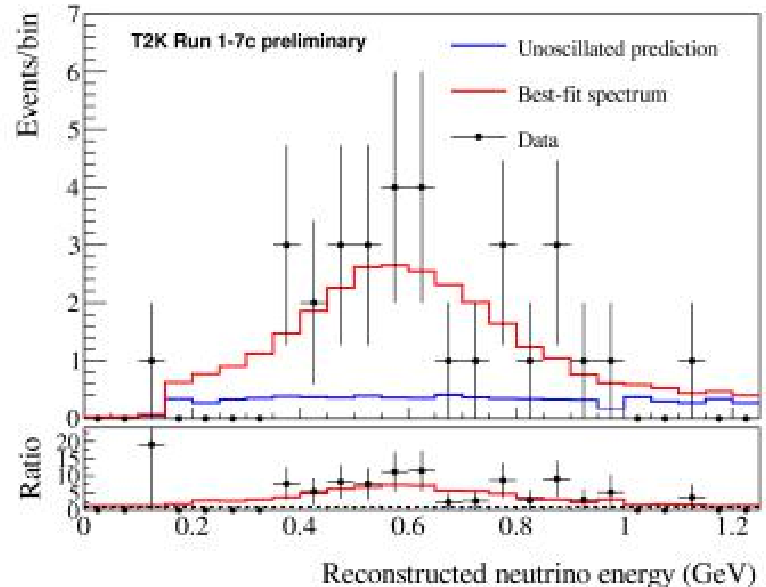
location of dip $\propto$ Δm_{23}^2

depth of dip $\propto$ $\sin^2\theta_{23}$

The oscillation analysis — Systematics sources

The dawn of all analysis:

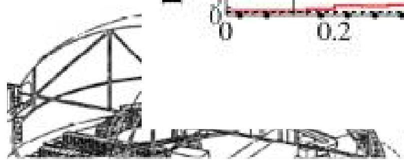
compare data to prediction (MC) and vary prediction (e.g. oscillation parameters) to find which one agree with data best



The sunset of all analysis?

the uncertainties on the prediction

$N_e(E) = \varphi_\mu(E)\ \sigma_e(E)\ \varepsilon_e(E)\ P(\nu_\mu \rightarrow \nu_e)$



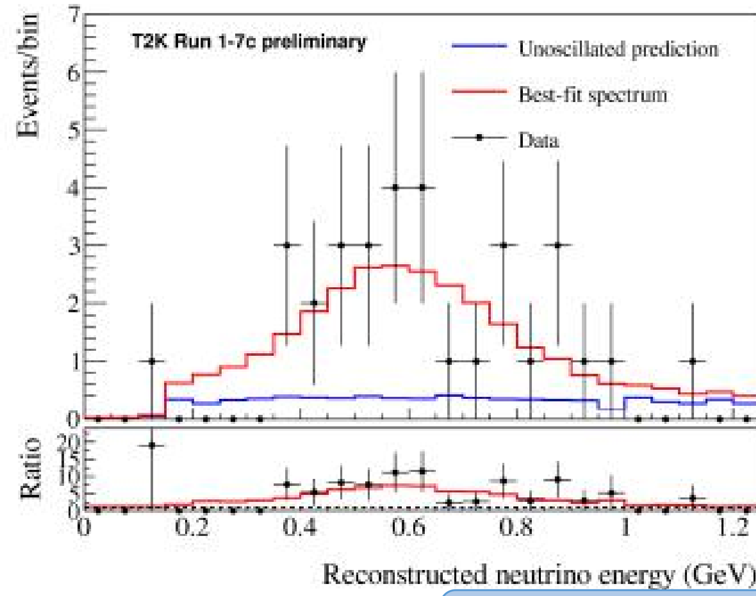
$e^{+/-}$ rings CC0π

Systematic error source	$\Delta N_{SK}/N_{SK}$ before ND fit
Flux	
Cross section	
Flux and cross section	
Final state/secondary interactions at SK	
SK detector	
Total	11.9%

59m Detector hall Access tunnel

The dawn of all analysis:

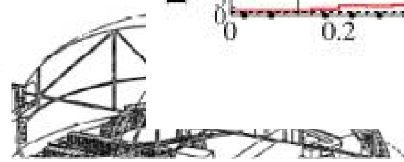
compare data to prediction (MC) and vary prediction (e.g. oscillation parameters) to find which one agree with data best



e$^{+/-}$ rings CC0π

The sunset of all analysis?

the uncertainties on the prediction



$$N_e(E)=\varphi_\mu(E)\ \sigma_e(E)\ \varepsilon_e(E)\ P(\nu_\mu \to \nu_e)$$

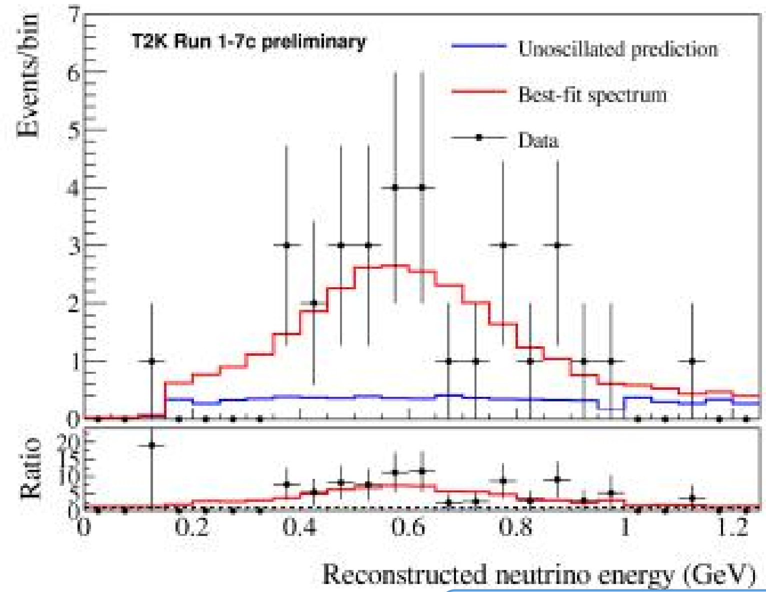
Systematic error source	$\Delta N_{SK}/N_{SK}$ before ND fit
Flux	
Cross section	
Flux and cross section	
Final state/secondary interactions at SK	2.5%
SK detector	2.5%
Total	11.9%

estimated from atmospheric ν in Super-K ← (Final state/secondary interactions at SK; SK detector)

39m Detector hall Access tunnel

The dawn of all analysis:

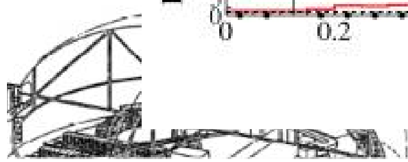
compare data to prediction (MC) and vary prediction (e.g. oscillation parameters) to find which one agree with data best



e$^{+/-}$ rings CC0π

The sunset of all analysis?

the uncertainties on the prediction



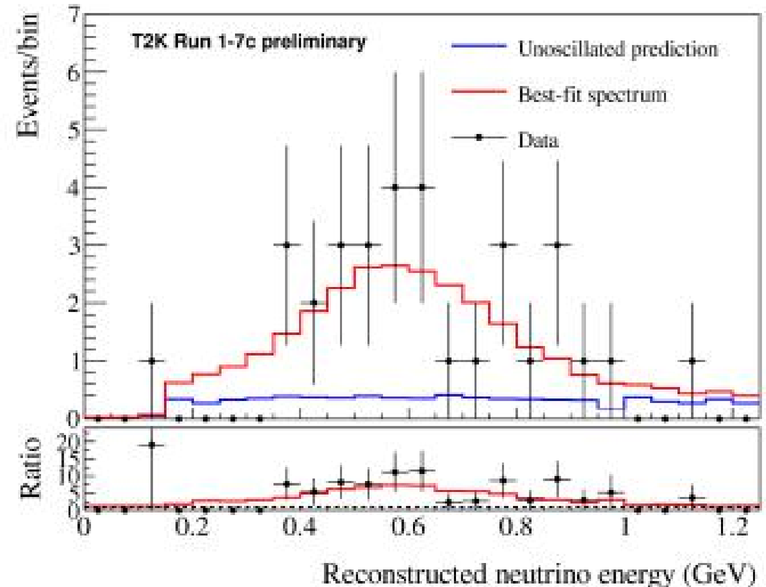
$$N_e(E) = \varphi_\mu(E)\ \sigma_e(E)\ \varepsilon_e(E)\ P(\nu_\mu \rightarrow \nu_e)$$

flux model constrained by NA61/SHINE + beam monitor measurements

Systematic error source	$\Delta N_{SK}/N_{SK}$ before ND fit
Flux	8.8%
Cross section	
Flux and cross section	
Final state/secondary interactions at SK	2.5%
SK detector	2.5%
Total	11.9%

The dawn of all analysis:

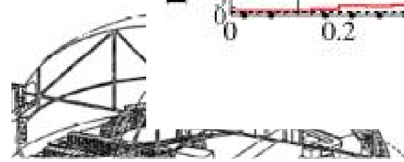
compare data to prediction (MC) and vary prediction (e.g. oscillation parameters) to find which one agree with data best



$e^{+/-}$ rings CC0π

The sunset of all analysis?

the uncertainties on the prediction



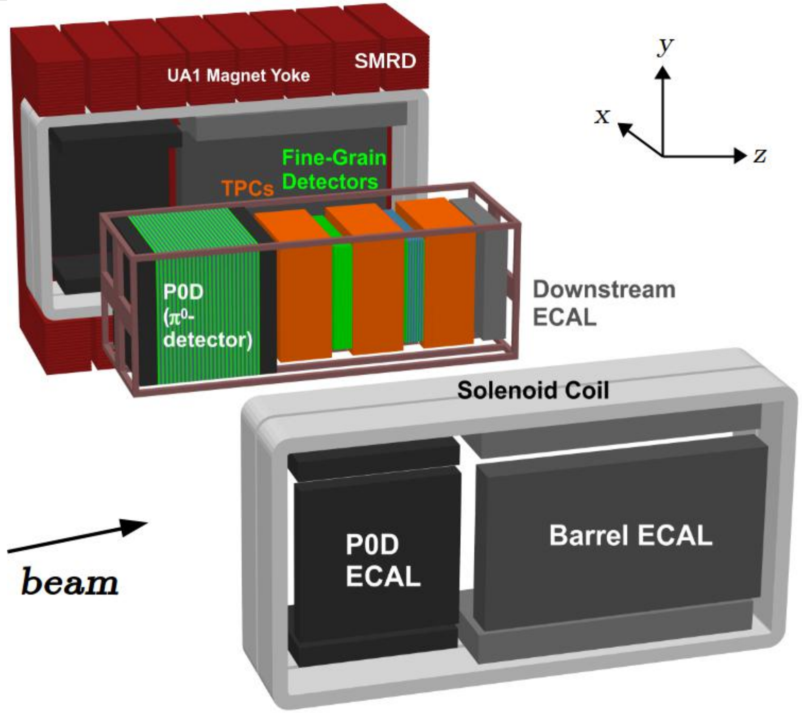
$$N_e(E) = \varphi_\mu(E)\ \sigma_e(E)\ \varepsilon_e(E)\ P(\nu_\mu \rightarrow \nu_e)$$

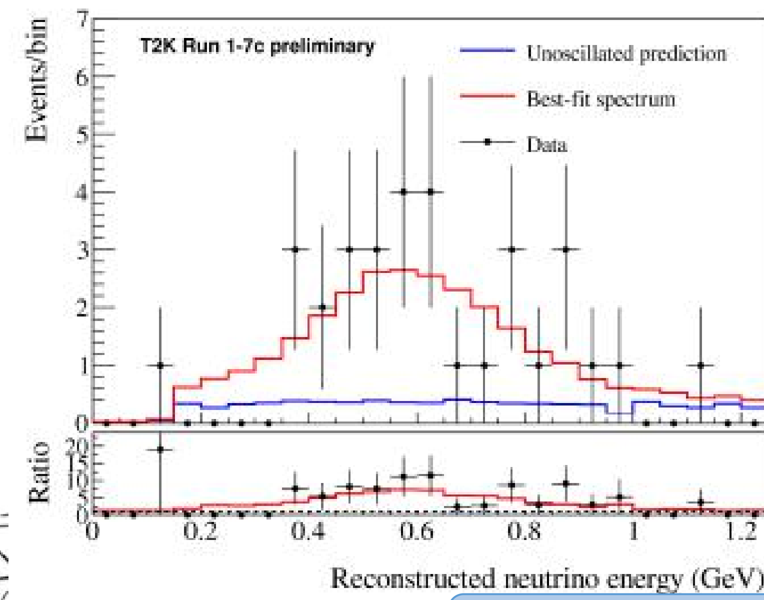
Systematic error source	$\Delta N_{SK}/N_{SK}$ before ND fit
Flux	8.8%
Cross section	7.1%
Flux and cross section	
Final state/secondary interactions at SK	2.5%
SK detector	2.5%
Total	11.9%

cross-section model constrained by measurements from other experiments

59m Detector hall Access tunnel

The oscillation analysis — Systematics sources





$e^{+/-}$ rings CC0π

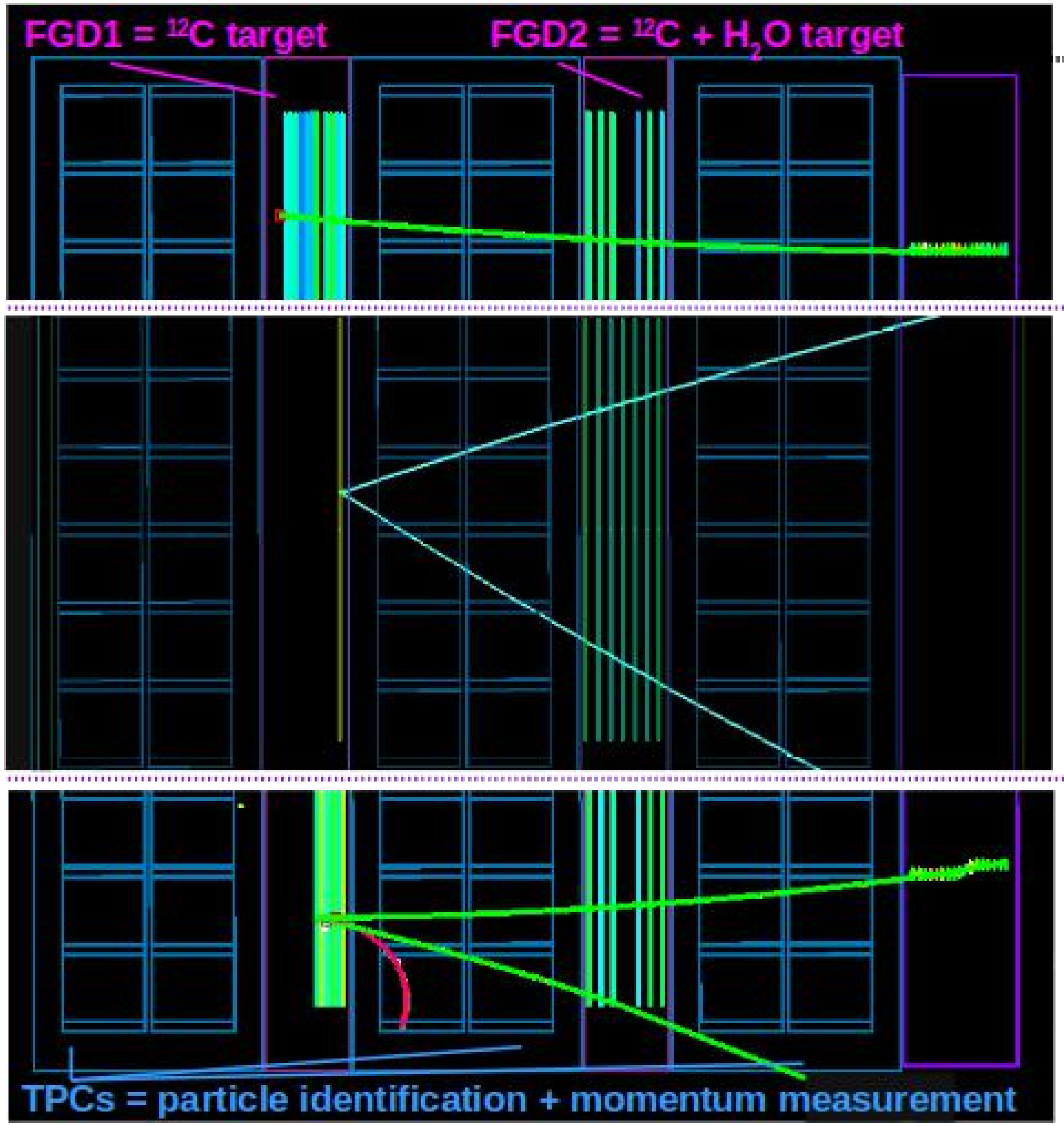
$N_e(E) = \varphi_\mu(E)\ \sigma_e(E)\ \varepsilon_e(E)\ P(\nu_\mu \rightarrow \nu_e)$

Systematic error source	$\Delta N_{SK}/N_{SK}$ before ND fit	$\Delta N_{SK}/N_{SK}$ after ND fit
Flux	8.8%	3.2%
Cross section	7.1%	4.7%
Flux and cross section	11.4%	2.7%
Final state/secondary interactions at SK	2.5%	
SK detector	2.5%	
Total	11.9%	5.2%

flux and cross-section models can be constrained with ND280 data

The samples

7 samples of charged-current (CC) interactions:

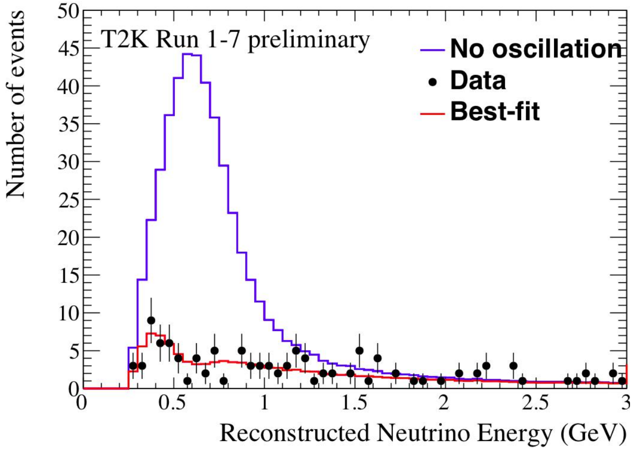


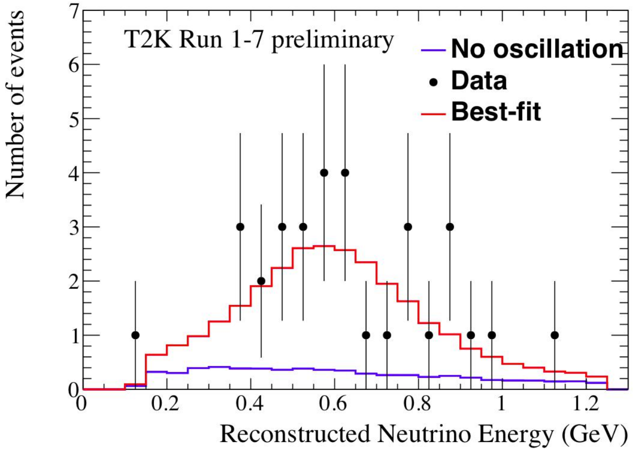
ν_μ in ν mode	$\overline{\nu}_\mu$ in $\overline{\nu}$ mode	ν_μ in $\overline{\nu}$ mode
CC-0π: only 1 μ^- detected	CC-0π: only 1 μ^+ detected	CC-0π: only 1 μ^- detected
CC-1π: 1 μ^- + 1 π^+ detected	CC-other: 1 μ^+ +something other detected	CC-other: 1 μ^- +something other detected
CC-other: 1 μ^- + something other than 1 π^+ detected		

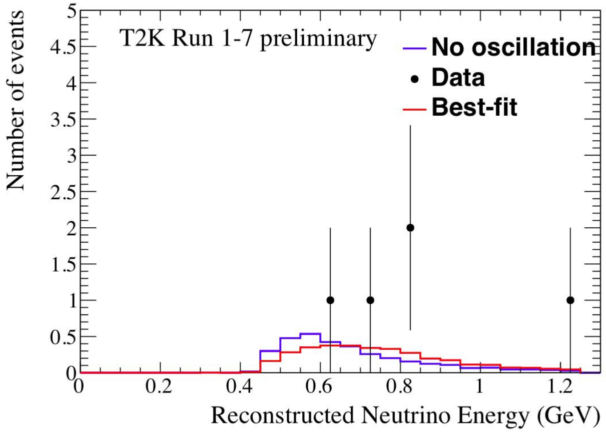
5 samples of charged-current (CC) interactions:

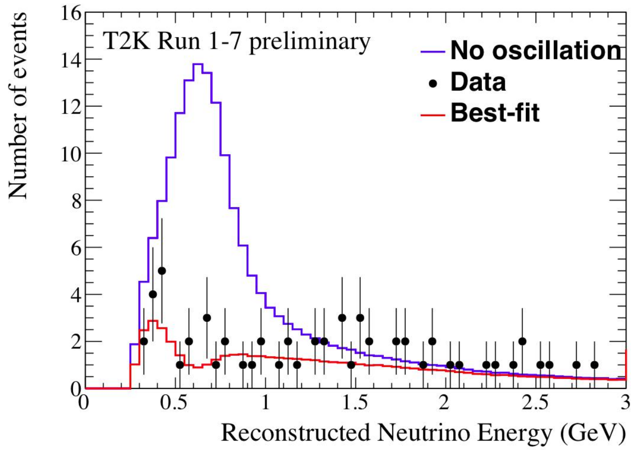
1st row is selection in ν mode, 2nd in $\overline{\nu}$ mode

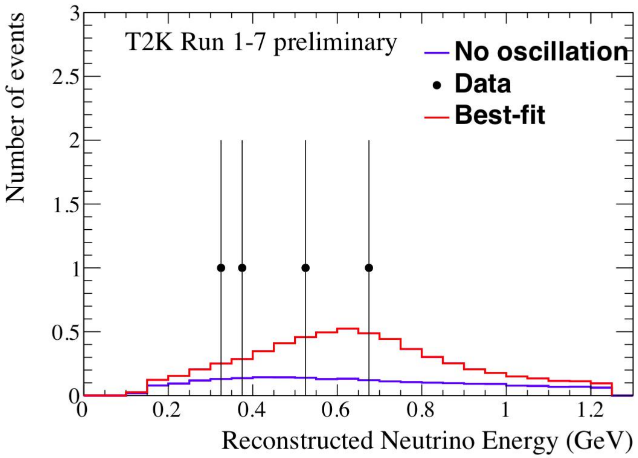
$\mu^{+/-}$ rings CC0π	$e^{+/-}$ rings CC0π	$e^{+/-}$ rings CC1π









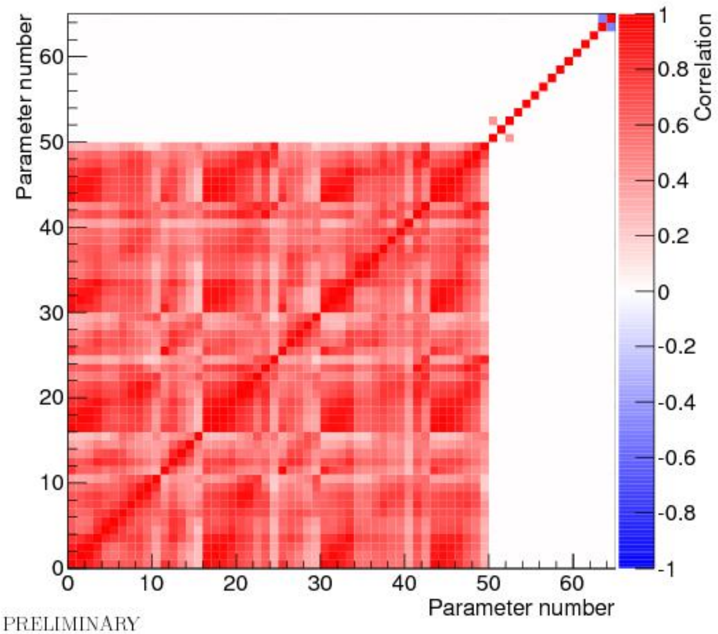


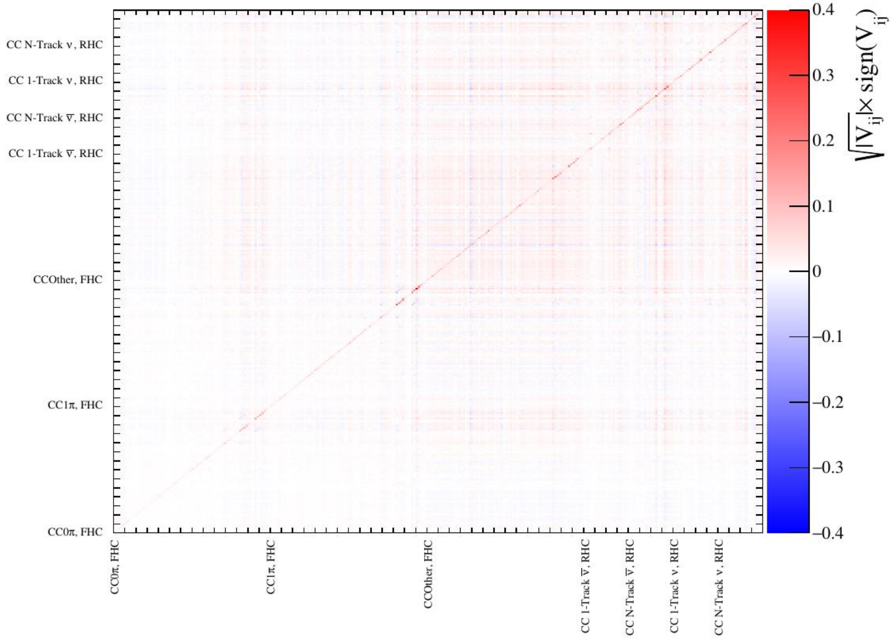
CC-0π 1R$_\mu$ rings	CC-0π 1R$_e$ rings	CC-1π^+ 1R$_e$ rings
event in SK fiducial volume	event in SK fiducial volume	event in SK fiducial volume
exactly 1 Cherenkov ring	exactly 1 Cherenkov ring	exactly 1 Cherenkov ring
PID is μ-like	PID is e-like	PID is e-like
$p_\mu^{rec} > 200$ MeV	visible energy > 100 MeV	visible energy > 100 MeV
≥ 1 electron decay	0 electron decay	$= 1$ electron decay
	reconstructed $E_\nu < 1.25$ GeV	reconstructed $E_\nu < 1.25$ GeV
	fitQun π^0 rejection	fitQun π^0 rejection

The issue ? Multidimensionality

- 7 samples in ND280 + 5 samples in Super-K
- 100 parameters for the flux model
- 26 parameters for the cross-section model
- 580 parameters for the ND280 detector systematics model
- 45 parameters for the Super-K detector systematics model

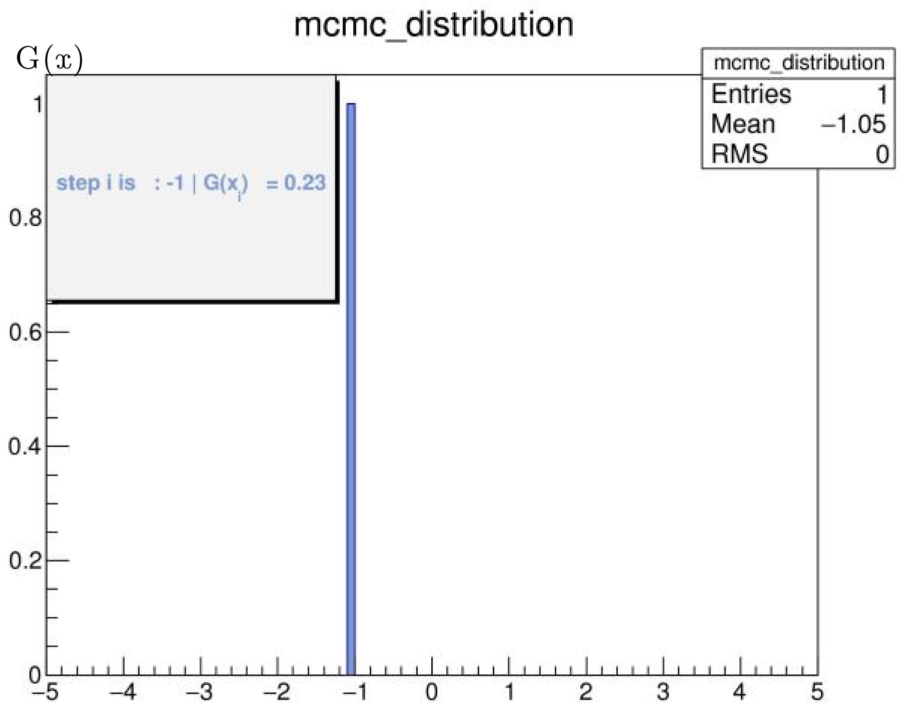
details about systematical uncertainties in in this talk





- MCMC is a semi-random walk in a parameter space
- The chain samples the parameters posterior distribution using the Metropolis Hastings algorithm:

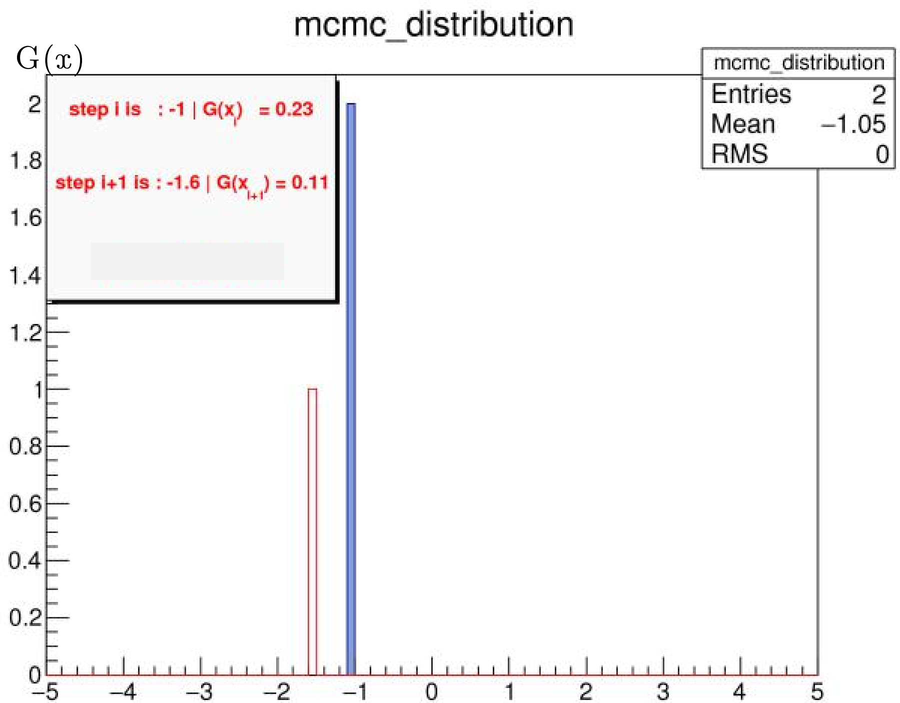
➡ Starts by chosing randomly a starting point in the parameter space.



- MCMC is a semi-random walk in a parameter space
- The chain samples the parameters posterior distribution using the Metropolis Hastings algorithm:

➡ Starts

➡ Propose new step by throwing a random value from the jump function $J(x_i+1 \mid x_i)$



The solution ! Markov Chain Monte-Carlo

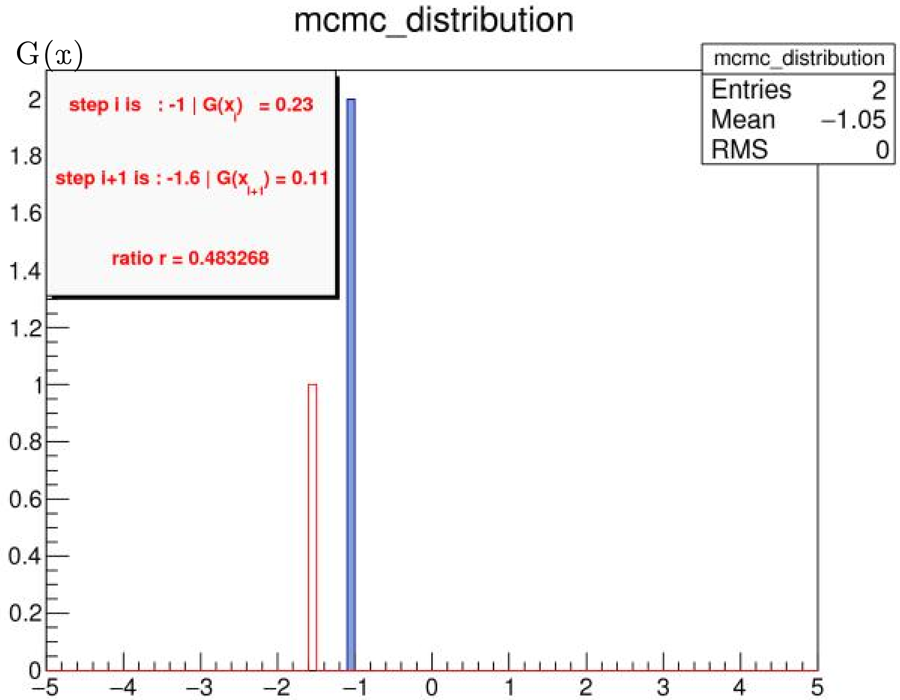
- MCMC is a semi-random walk in a parameter space
- The chain samples the parameters posterior distribution using the Metropolis Hastings algorithm:

➡ Starts

➡ Propose step $i+1$

➡ Compute

$$r = \frac{G(x_i+1)\ \ J(x_i+1 \mid x_i)}{G(x_i)\ \ \ \ J(x_i \mid x_i+1)}$$



- MCMC is a semi-random walk in a parameter space
- The chain samples the parameters posterior distribution using the Metropolis Hastings algorithm:

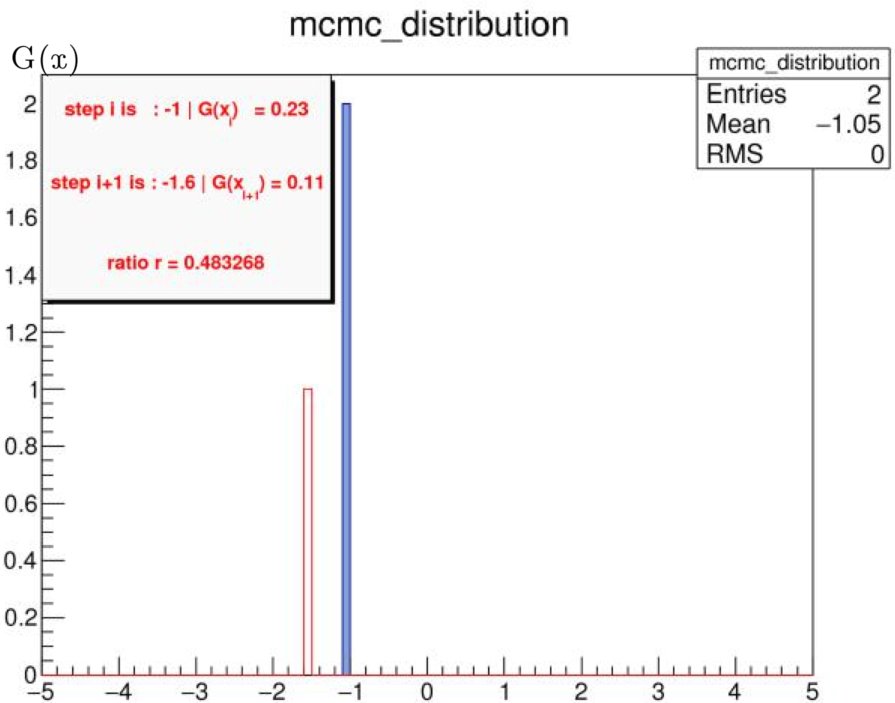
➡ Starts

➡ Propose step $i+1$

➡ Compute r

➡ Reject...

$r<1$ → throw in $U(0,1)$
$\rightarrow$ reject is r < $U(0,1)$
and re-count step i



- MCMC is a semi-random walk in a parameter space
- The chain samples the parameters posterior distribution using the Metropolis Hastings algorithm:

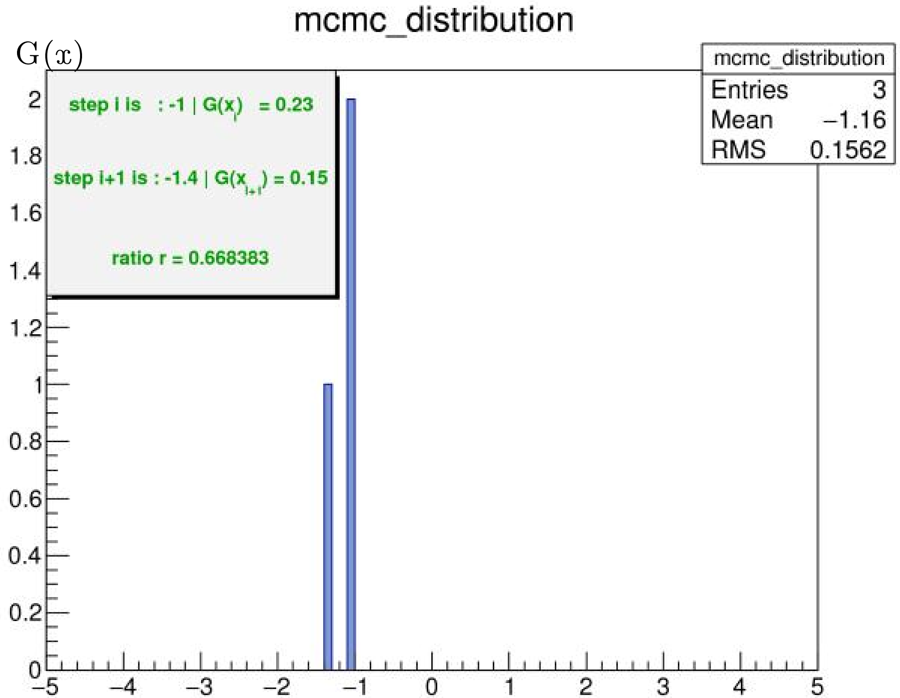
➡ Starts

➡ Propose step $i+1$

➡ Compute r

➡ Reject or accept step

$r<1$ → throw in $U(0,1)$
$\rightarrow$ reject if $r < U(0,1)$ and re-count step i
$\rightarrow$ accept if $r > U(0,1)$
$r>1$ → accept step $i+1$



- MCMC is a semi-random walk in a parameter space
- The chain samples the parameters posterior distribution using the Metropolis Hastings algorithm:

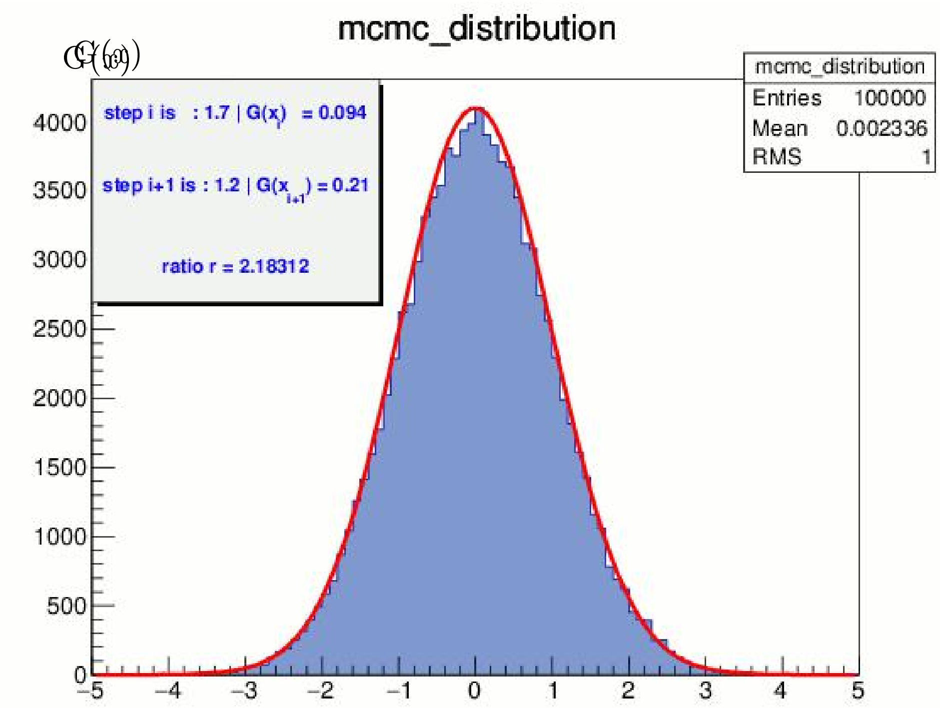
➡ Starts

➡ Propose step $i+1$

➡ Compute r

➡ Reject or accept step

➡ Keep on proposing / accepting / rejecting



correct sampling is ensured by the detailed balance condition (see backup)

The advantages

- Can handle a very high numbers of parameters and samples:
  can fit ND280 and Super-K data at the same time
  no extrapolation from near to far detector

- Can handle a very high numbers of parameters and samples
- Compute the joint posterior probability

  the sampled function is the posterior probability of all parameters

  $\vec{x} = \vec{o} + \vec{n}$ (o = oscillation ; n = nuisance)

$$G(x) \rightarrow P(\vec{x}|D) = \frac{P(D|\vec{x})\; P(\vec{x})}{P(D)}$$

binned Poisson likelihood: $P(D|\vec{x})$

prior knowledge on parameters: $P(\vec{x})$
→ flat for oscillation (except solar)
→ Gaussian for nuisance

joint posterior probability: $P(\vec{x}|D)$

normalisation parameter: $P(D)$

$\int P(D|\vec{x})P(\vec{x})d\vec{x}$

- Can handle a very high numbers of parameters and samples
- Compute the joint posterior probability
- Can sample distribution of any shape
  - can escape local minima
  - will always eventually find the distribution to sample
  - can sample the two separate distributions of each mass hierarchy
  - by setting a 50% probability of changing $\mathrm{sign}(\Delta m_{23}{}^2)$ at each step

$$r = \frac{G(x_i+1)\ J(x_i+1 \mid x_i)}{G(x_i)\ J(x_i \mid x_i+1)}$$

$r<1$ → throw in $U(0,1)$
→ reject is $r < U(0,1)$ and re-count step i
→ accept if $r > U(0,1)$
$r>1$ → accept step $i+1$

The advantages

- Can handle a very high numbers of parameters and samples
- Compute the joint posterior probability
- Can sample distribution of any shape
- Automatically marginalise the posterior probability

    mainly interested in the posterior distribution of the oscillation parameters

    projecting the posterior distribution includes the distribution of the nuisance parameters

a good lecture about Bayesian statistics and marginalisation

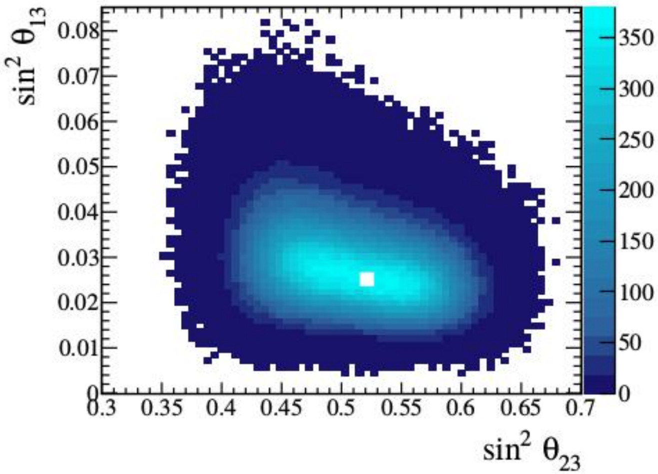
another one

The advantages

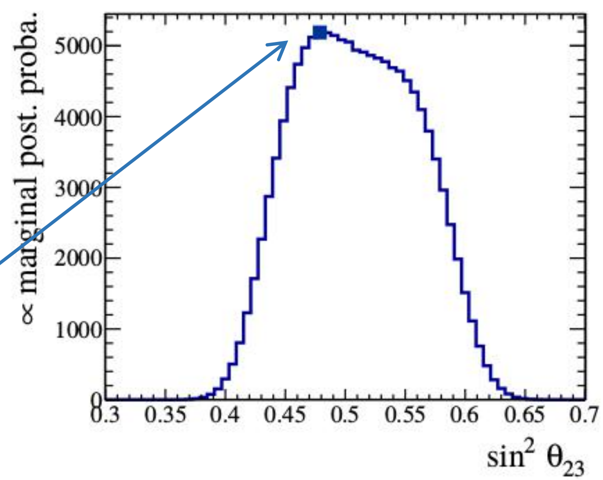
- Can h ... ·ameters and samples
- Comp
- Can s
- Autom ... probability

... distribution of the oscillation parameters

... ɔn includes the distribution of the



mode is shifted

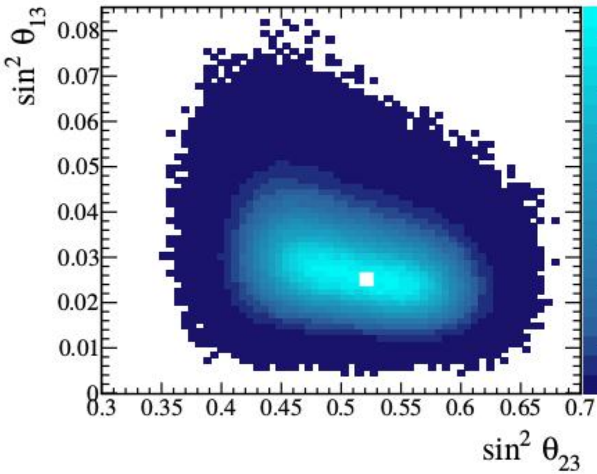


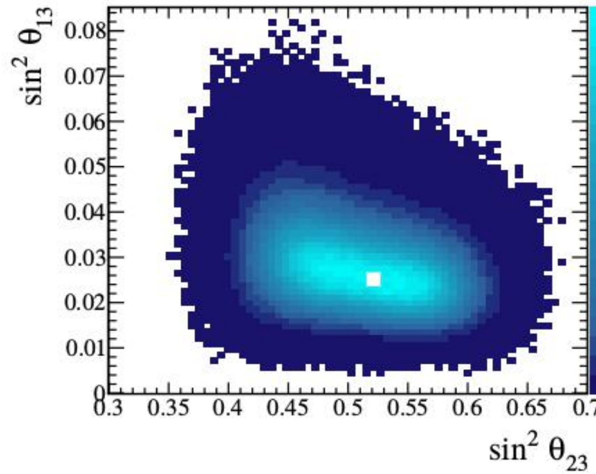
marginalising

$$P_m(\vec{o_i}) = \int P(\vec{o_i}, \vec{o_j}, \vec{n} \mid D)\, d(\vec{o_j}, \vec{n})$$

The advantages

- Can h
- Comp
- Can s
- Autom

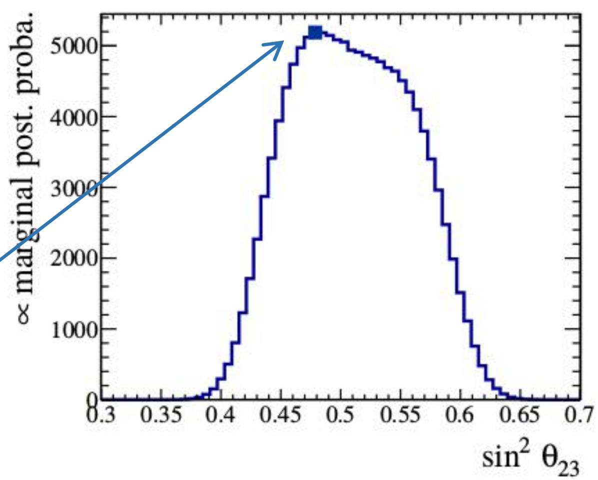


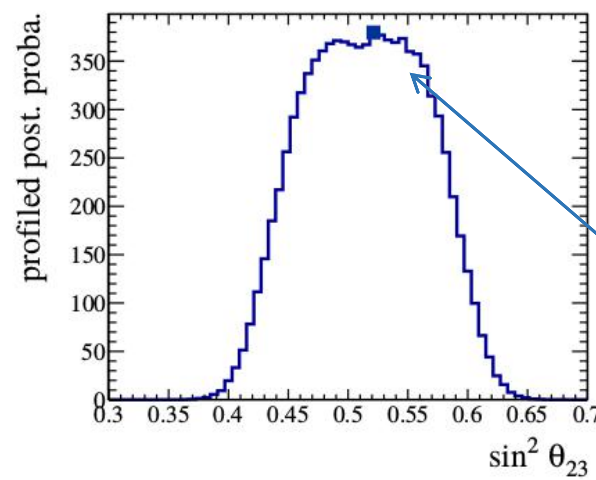


rameters

e

mode is shifted





mode is the same

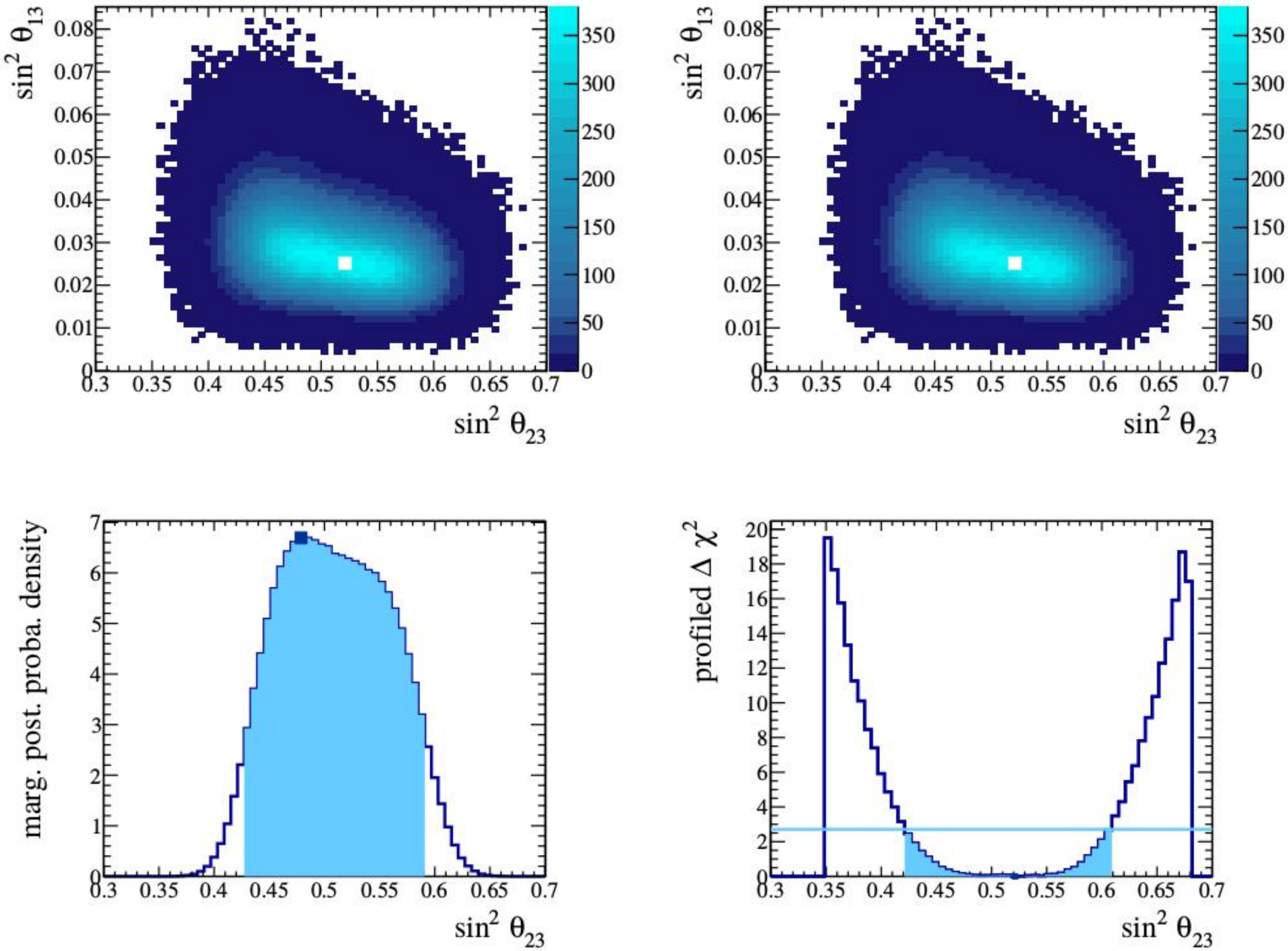
marginalising

$$P_m(\vec{o_i}) = \int P(\vec{o_i}, \vec{o_j}, \vec{n} | D)\ d(\vec{o_j}, \vec{n})$$

profiling

$$P_p(\vec{o_i}) = max_{\{oj,\ n\}}\ P(\vec{o_i}, \vec{o_j}, \vec{n} | D)$$

- Can h
- Comp
- Can s
- Autom



rameters

e

marginalising
credible intervals are $X\%$ of the area with highest probability

profiling
$\Delta\chi^2$ intervals is the area under a certain χ^2 value corresponding to $X\%$

The data analysed

ν-mode: 7.482×10^{20} P.O.T.

$\bar{\nu}$-mode: 7.471×10^{20} P.O.T.

Oscillation parameter	**Oscillation set A**
$\sin^2\theta_{12}$	0.304
$\sin^2\theta_{23}$	0.528
$\sin^2\theta_{13}$	0.0217
Δm^2_{21}	7.53×10^{-5} eV^2
Δm^2_{32}	2.509×10^{-3} eV^2
δ_{CP}	$-\pi/2$

	Total expected event rate per sample				
	FHC 1R$_e$	FHC 1R$_e$ CC-1π^+	FHC 1R$_\mu$	RHC 1R$_e$	RHC 1R$_\mu$
prefit MC	26.163	3.581	129.837	5.815	62.477

The data analysed

ν-mode: 7.482×10^{20} P.O.T.

$\bar{\nu}$-mode: 7.471×10^{20} P.O.T.

Oscillation parameter	**Oscillation set A**
$\sin^2\theta_{12}$	0.304
$\sin^2\theta_{23}$	0.528
$\sin^2\theta_{13}$	0.0217
Δm^2_{21}	7.53×10^{-5} eV^2
Δm^2_{32}	2.509×10^{-3} eV^2
δ_{CP}	$-\pi/2$

	Total expected event rate per sample				
	FHC 1R$_e$	FHC 1R$_e$ CC-1π^+	FHC 1R$_\mu$	RHC 1R$_e$	RHC 1R$_\mu$
prefit MC	26.163	3.581	129.837	5.815	62.477
Data	32	5	135	4	66

The data analysed

ν-mode: 7.482×10^{20} P.O.T.

$\bar{\nu}$-mode: 7.471×10^{20} P.O.T.

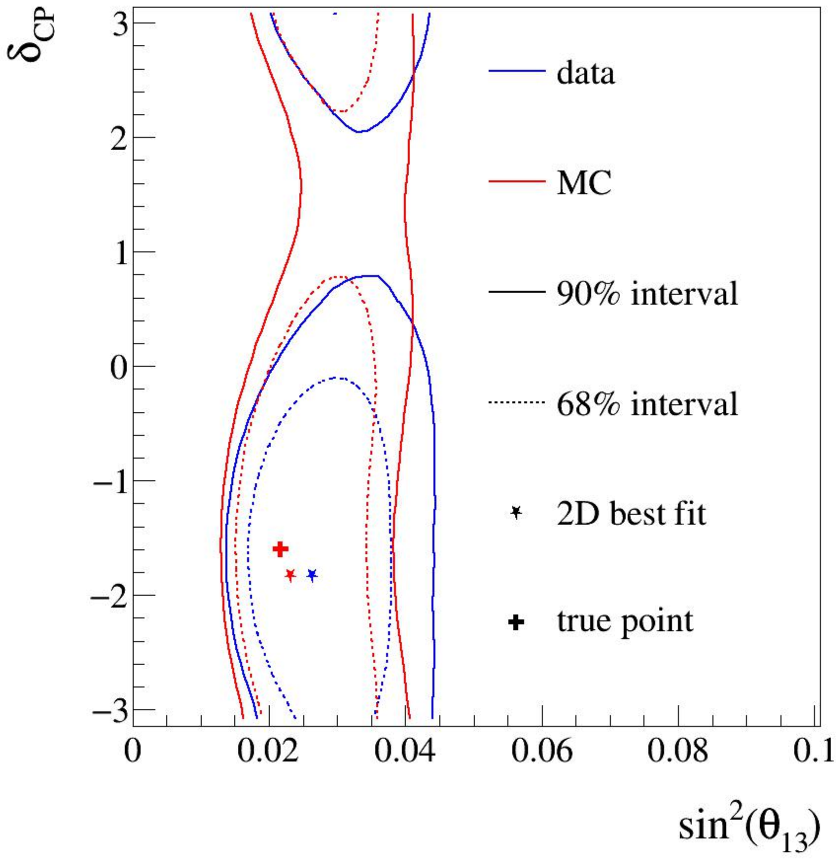
Oscillation parameter	**Oscillation set A**
$\sin^2\theta_{12}$	0.304
$\sin^2\theta_{23}$	0.528
$\sin^2\theta_{13}$	0.0217
Δm^2_{21}	7.53×10^{-5} eV^2
Δm^2_{32}	2.509×10^{-3} eV^2
δ_{CP}	$-\pi/2$

	Total expected event rate per sample				
	FHC 1R$_e$	FHC 1R$_e$ CC-1π^+	FHC 1R$_\mu$	RHC 1R$_e$	RHC 1R$_\mu$
prefit MC	26.163	3.581	129.837	5.815	62.477
Data	32	5	135	4	66
Data / MC (prefit)	1.22	1.40	1.04	0.69	1.06

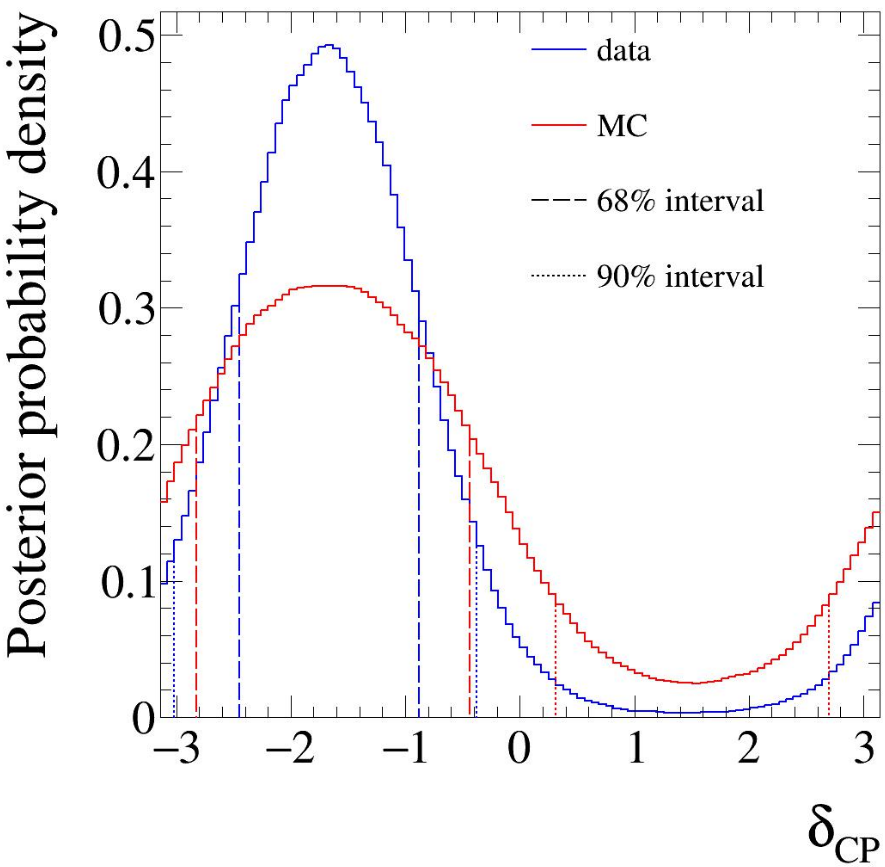
statistical fluctuation ?

Results v.s. sensitivity

fit with flat prior on $sin^2\theta_{13}$

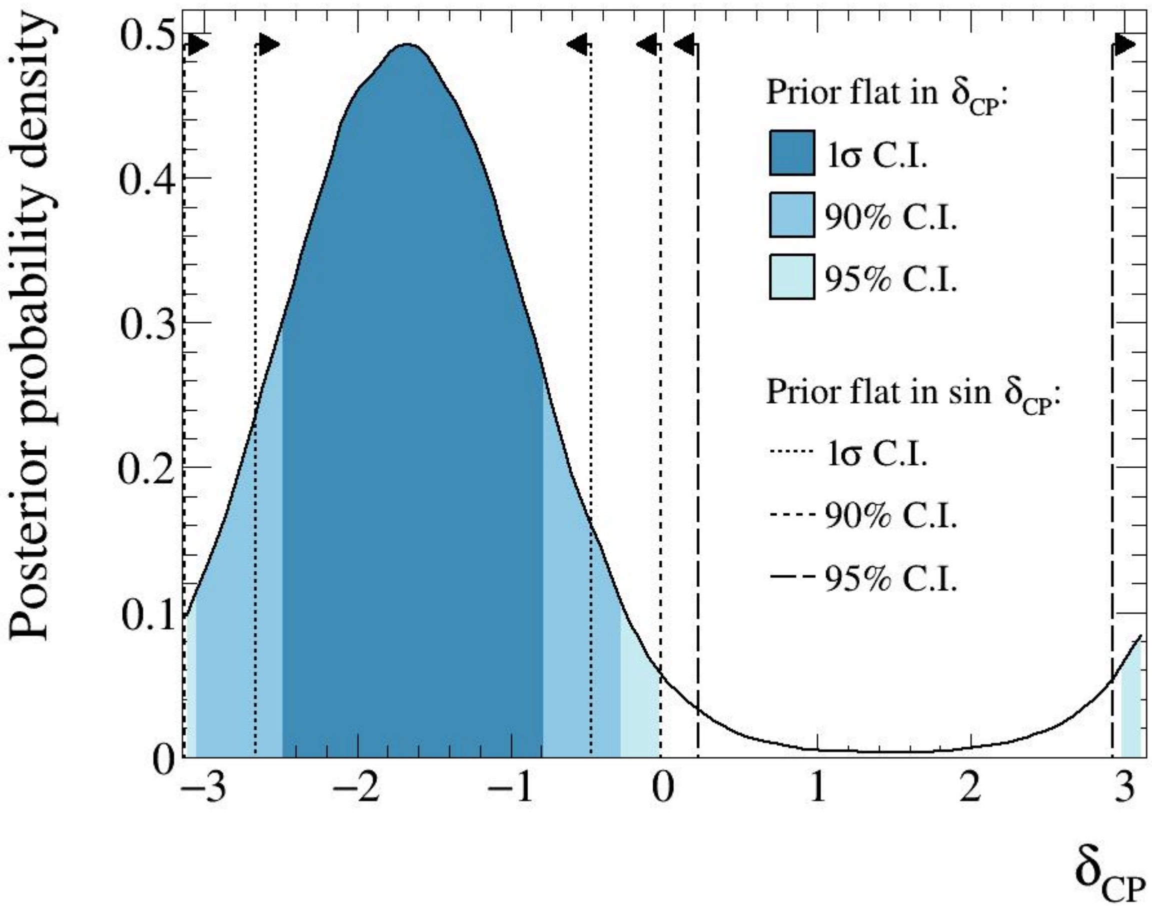


fit with Gaussian prior on $sin^2\theta_{13}$ (PDG 2015 value)



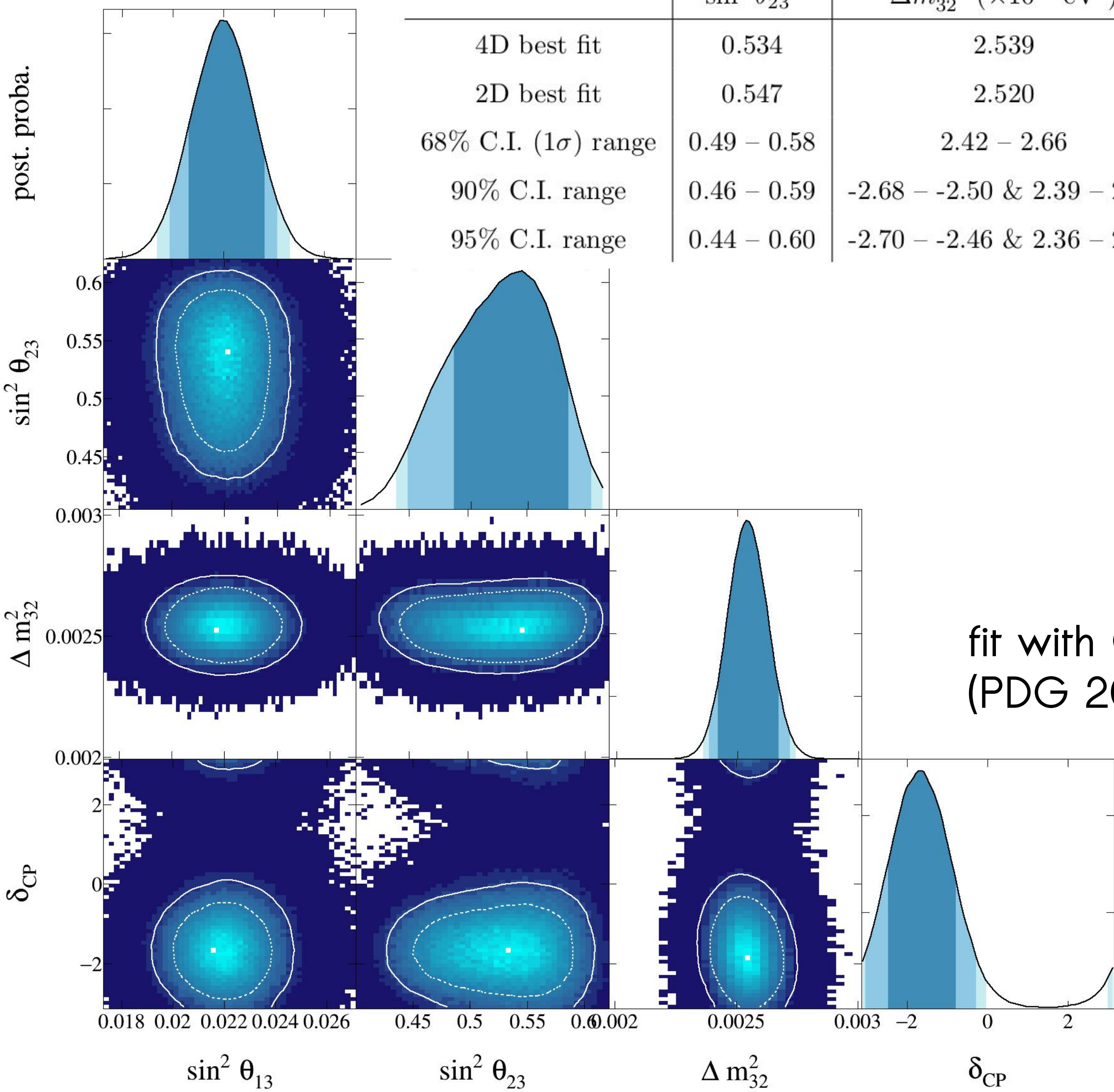
Results v.s. prior

The δ_{CP} credible intervals are different when assigning a flat prior on δ_{CP} or $sin(\delta_{CP})$.
It indicates little power in constraining the parameter with the available data.



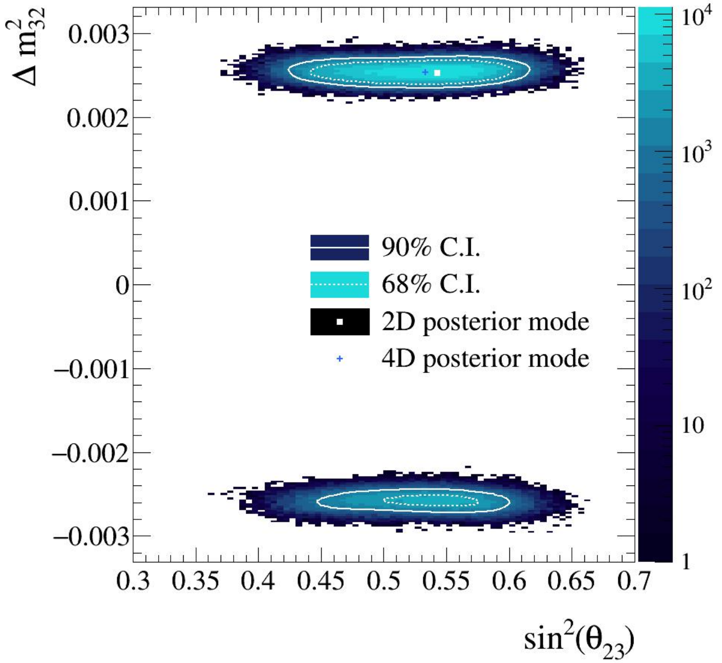
	$\sin^2\theta_{23}$	Δm^2_{32} ($\times 10^{-3}$eV^2)	$\sin^2\theta_{13}$	δ_{CP} (rad.)
4D best fit	0.534	2.539	0.0219	-1.79
2D best fit	0.547	2.520	0.0219	-1.75
68% C.I. (1σ) range	0.49 – 0.58	2.42 – 2.66	0.0208 – 0.0230	-2.45 – -0.88
90% C.I. range	0.46 – 0.59	-2.68 – -2.50 & 2.39 – 2.70	0.0200 – 0.0240	-3.02 – -0.38
95% C.I. range	0.44 – 0.60	-2.70 – -2.46 & 2.36 – 2.72	0.0195 – 0.0242	$-\pi$ – -0.06 & 3.02 – π

no CP violation excluded at 90%



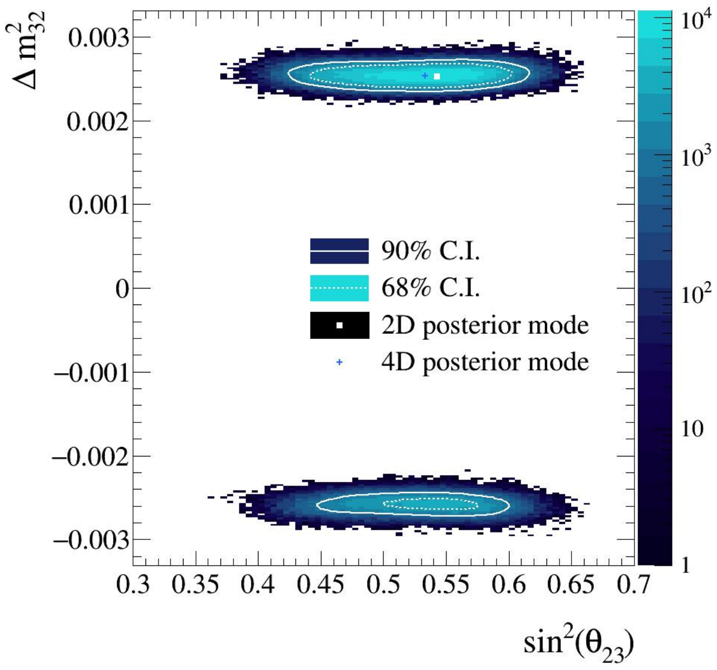
fit with Gaussian prior on $\sin^2\theta_{13}$ (PDG 2015 value)

Mass hierarchy



	$\sin^2\theta_{23} < 0.5$	$\sin^2\theta_{23} > 0.5$	Sum
IH ($\Delta m^2_{32} < 0$)	0.060	0.152	0.212
NH ($\Delta m^2_{32} > 0$)	0.233	0.555	0.788
Sum	0.293	0.707	1

Mass hierarchy

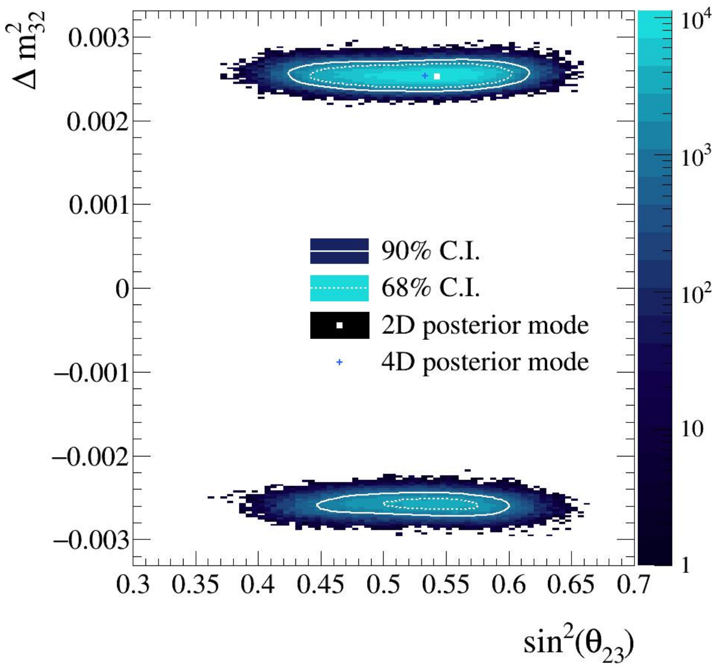


	$\sin^2\theta_{23} < 0.5$	$\sin^2\theta_{23} > 0.5$	Sum
IH ($\Delta m^2_{32} < 0$)	0.060	0.152	0.212
NH ($\Delta m^2_{32} > 0$)	0.233	0.555	0.788
Sum	0.293	0.707	1

posterior odds

$$\frac{P(\Delta m_{32}{}^2 > 0 \mid D)}{P(\Delta m_{32}{}^2 < 0 \mid D)}$$

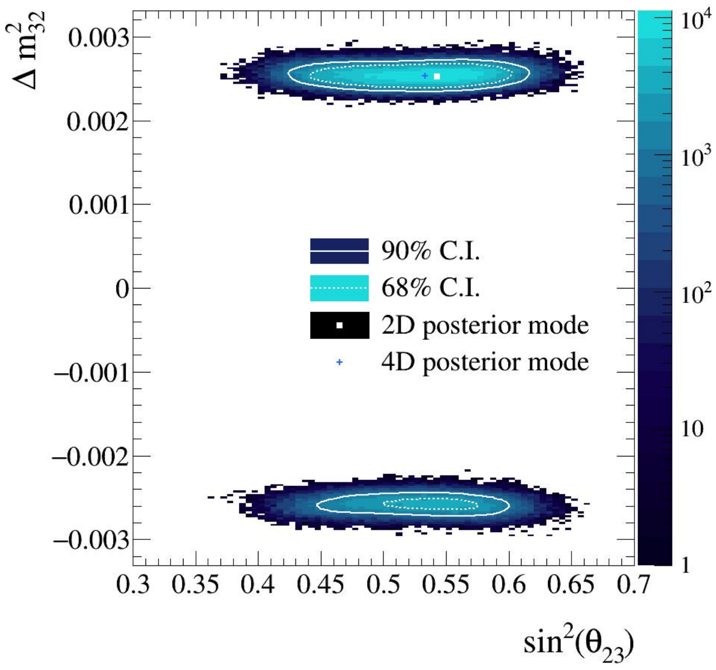
Mass hierarchy



	$\sin^2\theta_{23} < 0.5$	$\sin^2\theta_{23} > 0.5$	Sum
IH ($\Delta m^2_{32} < 0$)	0.060	0.152	0.212
NH ($\Delta m^2_{32} > 0$)	0.233	0.555	0.788
Sum	0.293	0.707	1

$$\underbrace{\frac{P(\Delta m_{32}{}^2>0 \mid D)}{P(\Delta m_{32}{}^2<0 \mid D)}}_{\text{posterior odds}} = \frac{\overbrace{P(D \mid \Delta m_{32}{}^2>0)}^{\text{Bayes factor}}\ \overbrace{P(\Delta m_{32}{}^2>0)}^{\text{1 in this case}}}{P(D \mid \Delta m_{32}{}^2<0)\ P(\Delta m_{32}{}^2<0)}$$

Mass hierarchy



	$\sin^2\theta_{23} < 0.5$	$\sin^2\theta_{23} > 0.5$	Sum
IH ($\Delta m^2_{32} < 0$)	0.060	0.152	0.212
NH ($\Delta m^2_{32} > 0$)	0.233	0.555	0.788
Sum	0.293	0.707	1

for NH: 3.72

for higher octant: 2.41

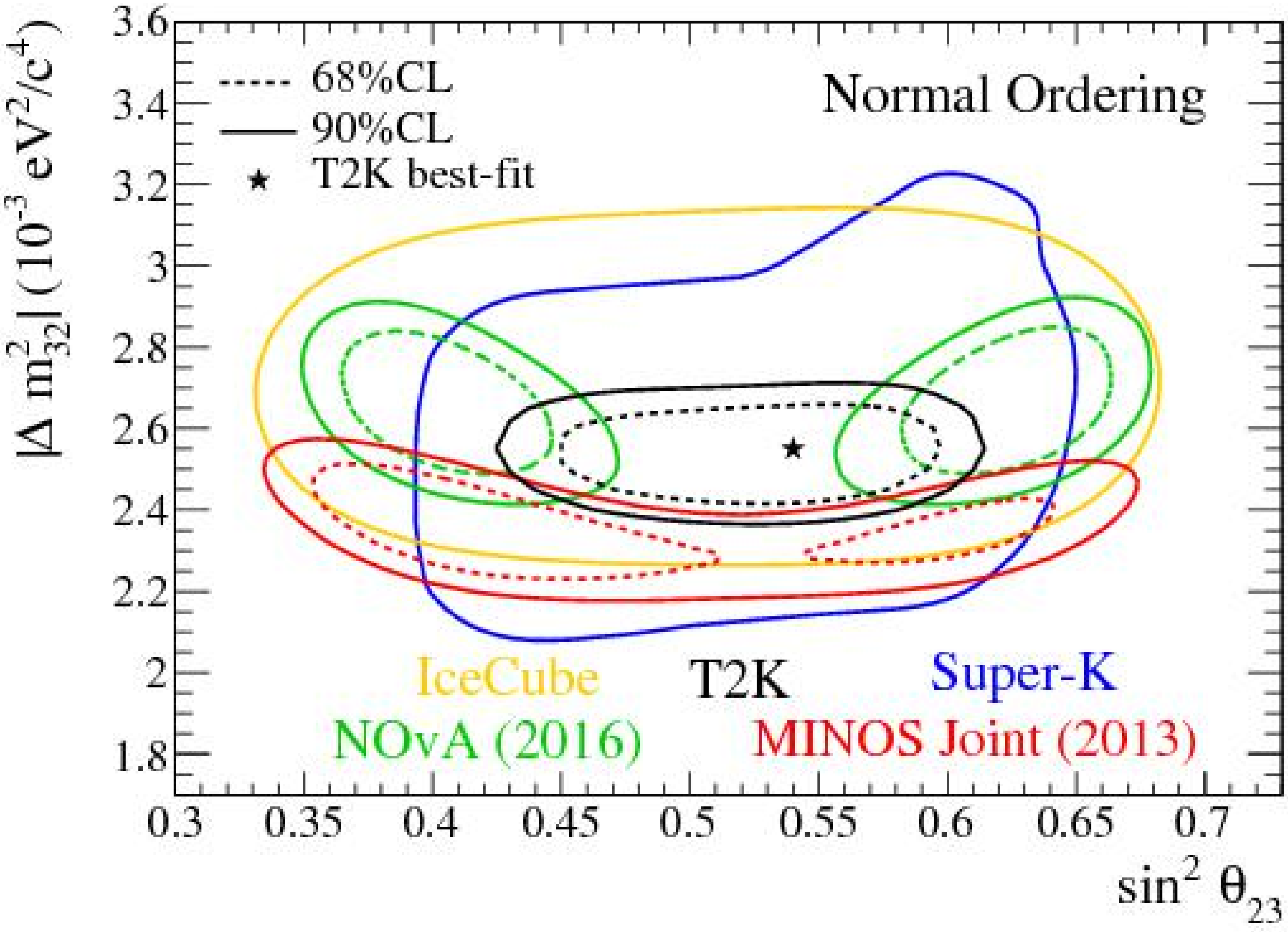
nothing strong under 10

source

posterior odds — Bayes factor — 1 in this case

$$\frac{P(\Delta m_{32}{}^2>0 \mid D)}{P(\Delta m_{32}{}^2<0 \mid D)} = \frac{P(D \mid \Delta m_{32}{}^2>0)\; P(\Delta m_{32}{}^2>0)}{P(D \mid \Delta m_{32}{}^2<0)\; P(\Delta m_{32}{}^2<0)}$$

T2K results v.s. other experiments



In conclusion

- Markov Chain Monte Carlo is a robust technique to sample a marginalised posterior probability distribution.

- It can deal with multidimensionality, local minima, disjoint distributions.

- A famous statistician:
"Frequentists answer the question nobody ask in a way everybody agree.
Bayesians answer the question everbody ask in a way nobody agree."

- You can compare Frequentist and Bayesian output (sometimes):
do both, and understand if the difference lie in the assumption / technique (from the same statistician).

In conclusion

- T2K is one of the leading neutrino oscillation experiments.

- Hints to new physics may rely in neutrino oscillations, and we're on our way to constrain it.

- We are now doing an update of the oscillation analysis, with modification of the event selection, and will release new results in the Summer.

- Undergoing studies of extending P.O.T., upgrading ND280, adding an intermediate detector, making a bigger far detector.

backup

Markov Chain Monte-Carlo

- How does it work :
  - The probability of accepting the proposed step using the Metropolis ratio r can be written : $A(x_{i+1}, x_i) = \min\{1, r\}$

$r > 1 \rightarrow A(x_{i+1}) = 1$ *[step automatically accepted]*

$r < 1 \rightarrow A(x_{i+1}) = r$ *[the chances to accept the step are proportionnal to r]*

  - Defining to transition probability $T(x_{i+1} | x_i) = J(x_{i+1} | x_i) \times A(x_{i+1}, x_i)$ then we can derive the detailed balanced equation (in backup) :

$$G(x_i) \times T(x_{i+1} | x_i) = G(x_{i+1}) \times T(x_i | x_{i+1})$$

$$r = \frac{G(x_{i+1}) \; J(x_i \mid x_{i+1})}{G(x_i) \; J(x_{i+1} \mid x_i)}$$

jump function $J(x_{i+1} \mid x_i)$: a Gaussian

$$G(x_i) \times T(x_{i+1} | x_i) = G(x_i) \times J(x_{i+1} | x_i) \times A(x_{i+1}, x_i)$$

$$= G(x_i) \times J(x_{i+1} | x_i) \times min\{ 1, r \}$$

$$= G(x_i) \times J(x_{i+1} | x_i) \times min\{ 1, \frac{G(x_{i+1}) \times J(x_i | x_{i+1})}{G(x_i) \times J(x_{i+1} | x_i)} \}$$

$$= min\{ G(x_i) \times J(x_{i+1} | x_i), G(x_{i+1}) \times J(x_i | x_{i+1}) \}$$

$$= G(x_{i+1}) \times J(x_i | x_{i+1}) \times A(x_i, x_{i+1})$$

$$= G(x_{i+1}) \times T(x_i | x_{i+1})$$

- What does it mean ?

$$G(x_i) \times T(x_{i+1}|x_i) = G(x_{i+1}) \times T(x_i|x_{i+1})$$

- Let's say we proposed a step with $G(x_{i+1}) > G(x_i)$ then $r > 1$ so the step is accepted : $T(x_{i+1}|x_i)$ as the transition probability is certain.

- Then we get $T(x_i|x_{i+1}) = \dfrac{G(x_i)}{G(x_{i+1})}$

→ the probability to go back to the step we were is equal to the ratio of the values of $G(x)$

→ the stepping frequency is proportionnal to the value of $G(x)$ relative to the current step !