Neutrino Theory and Phenomenology:

Lecture III

NBIA PhD School: Neutrinos Underground & in the Heavens (Copenhagen, DK, 2016)

Eligio Lisi (INFN, Bari, Italy) eligio.lisi@ba.infn.it

Outline of lectures:

Lecture I
Pedagogical intro + warm-up case study for oscillations

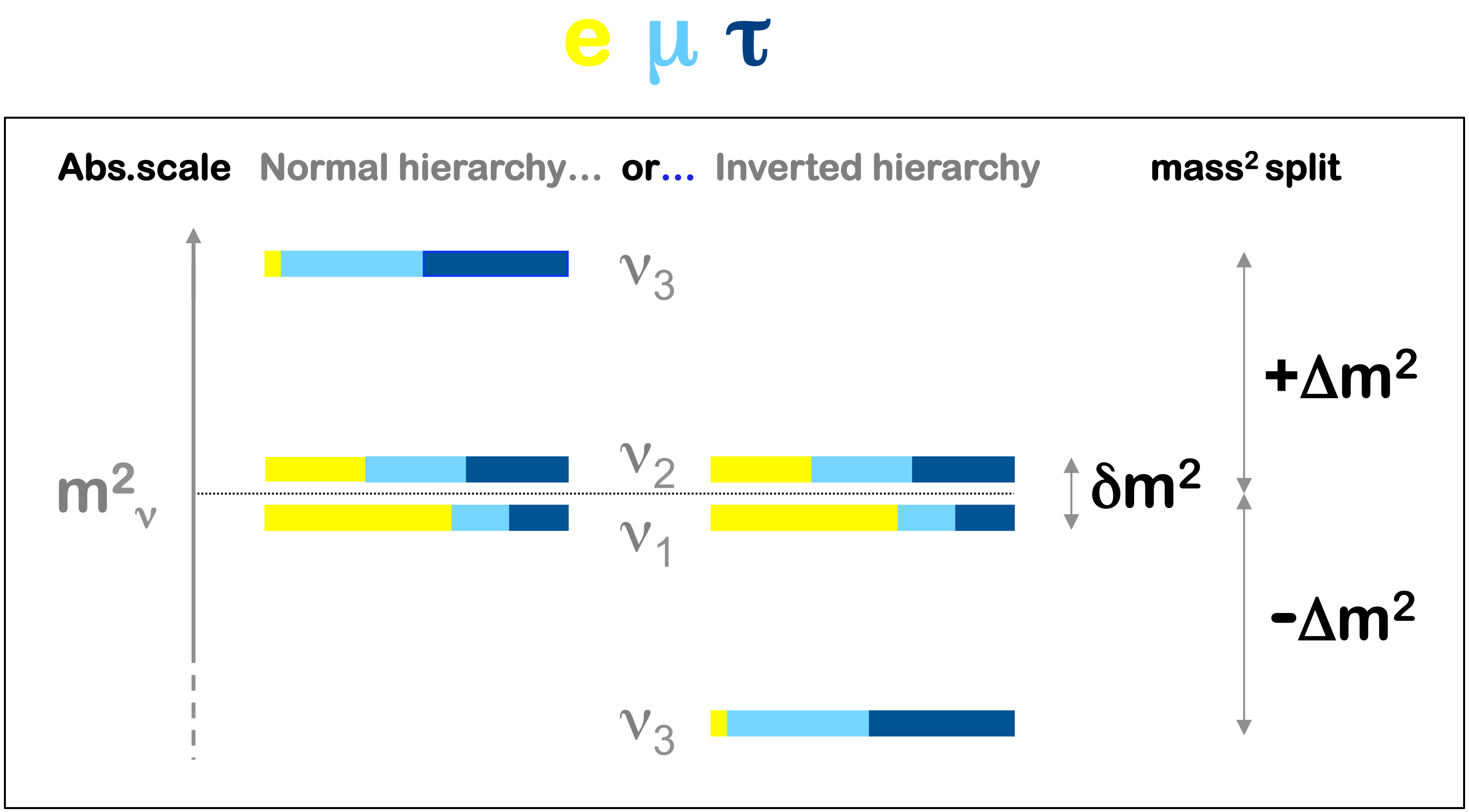
Lecture II
Standard 3ν oscillations: evolution and current status

Lecture III
Neutrino absolute masses + open problems in ν physics

(+ Appendix on statistics and data analysis)

Neutrino absolute masses

Present 3ν knowledge in one slide (with 1-digit accuracy)



We have seen:

$\delta m^2 \sim 7 \times 10^{-5}$ eV^2
$\Delta m^2 \sim 2 \times 10^{-3}$ eV^2
$\sin^2\theta_{12} \sim 0.3$
$\sin^2\theta_{23} \sim 0.5$
$\sin^2\theta_{13} \sim 0.02$

We would like to see:

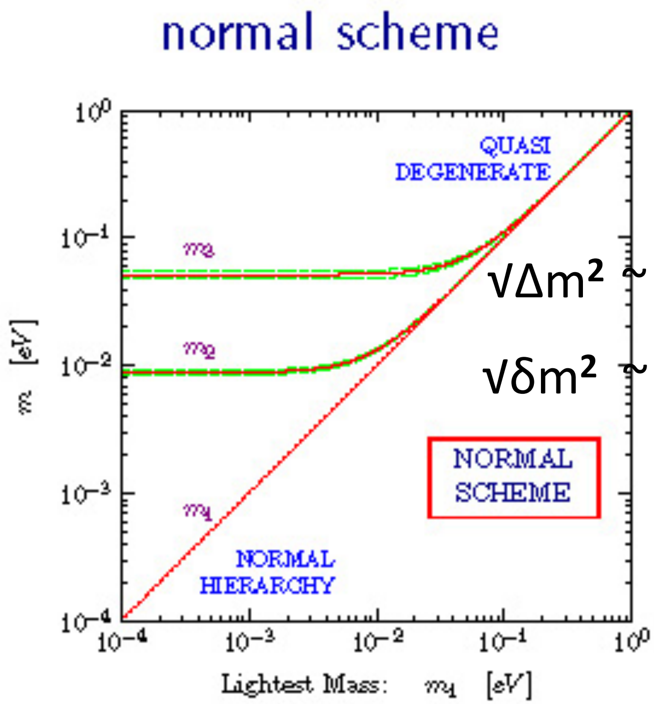
δ (CP)
sign(Δm^2)
octant(θ_{23})
absolute mass scale
Dirac/Majorana nature

+ Physics beyond 3ν?

(anomalies, new states or interactions)

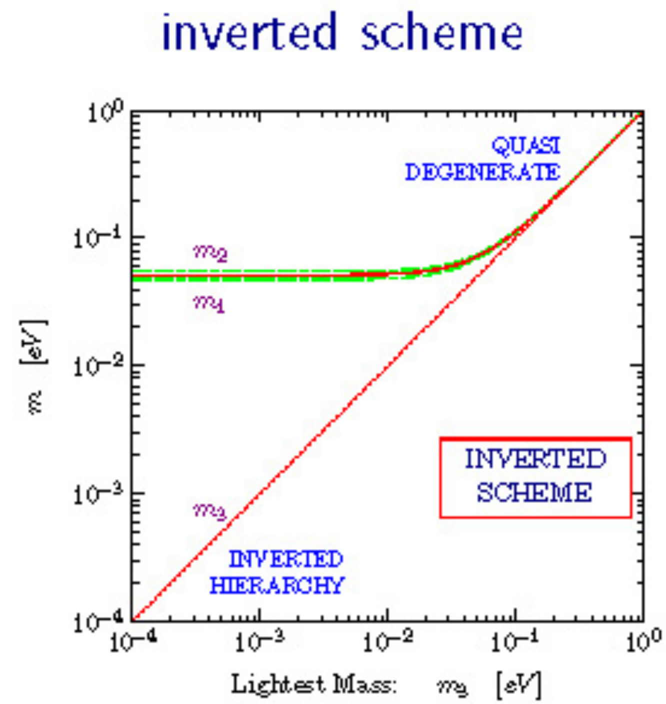
Oscillations constrain neutrino mixings and mass splittings
but not the absolute mass scale.

E.g., can take the lightest neutrino mass as free parameter:



$\sqrt{\Delta m^2} \sim 0.05$ eV

$\sqrt{\delta m^2} \sim 0.009$ eV



However, the lightest neutrino mass is not really an "observable"

We know three realistic observables to probe ν masses →

Absolute mass observables: (m_β, $m_{\beta\beta}$, Σ)

1) **β decay**: $m^2_i \neq 0$ can affect spectrum endpoint. Sensitive to the "effective electron neutrino mass":

$$m_\beta = \left[c_{13}^2 c_{12}^2 m_1^2 + c_{13}^2 s_{12}^2 m_2^2 + s_{13}^2 m_3^2\right]^{\frac{1}{2}}$$

2) **0νββ decay**: Can occur if $m^2_i \neq 0$ and $\nu = \bar{\nu}$ (**Majorana**, not Dirac) Sensitive to the "effective Majorana mass" (and phases):

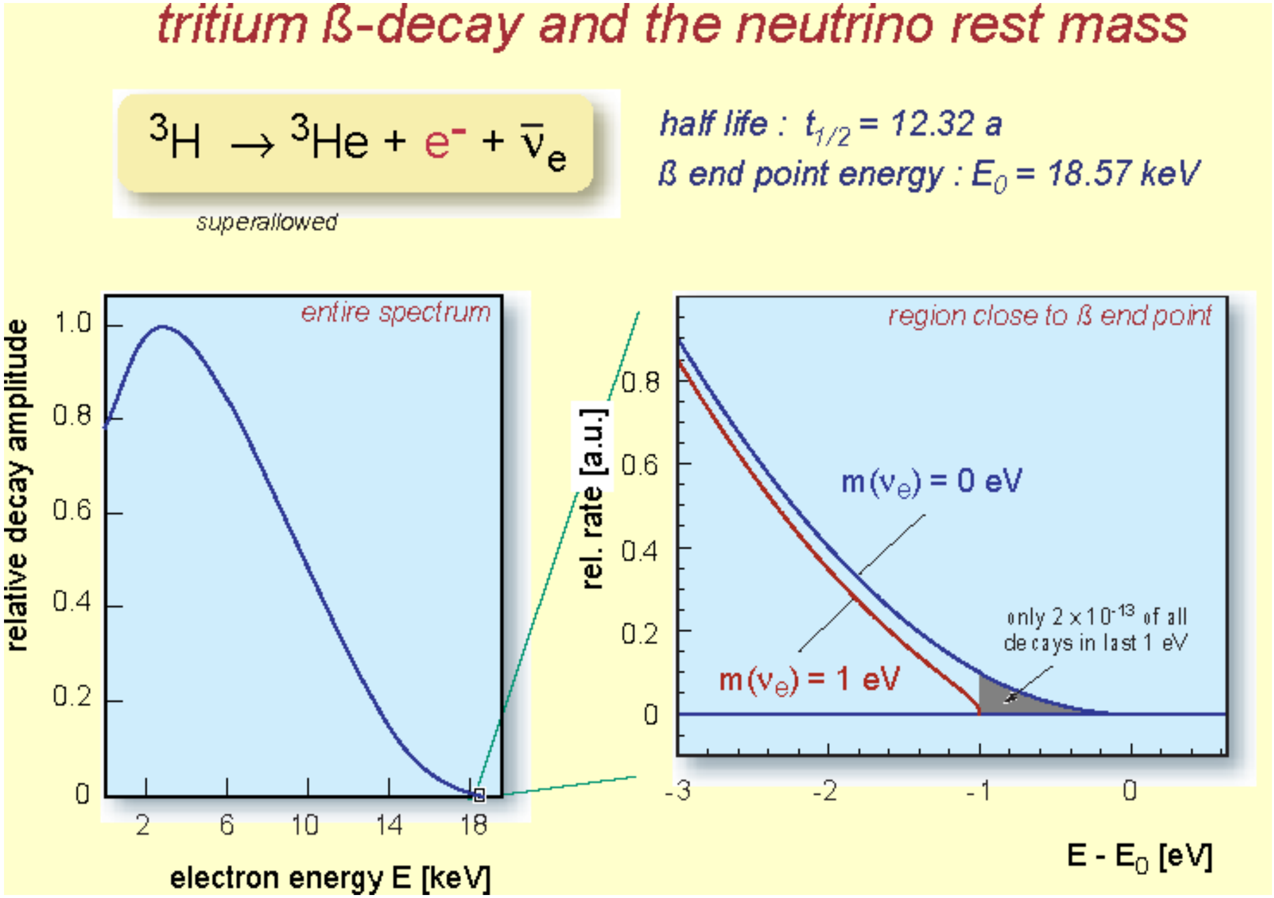
$$m_{\beta\beta} = \left|c_{13}^2 c_{12}^2 m_1 + c_{13}^2 s_{12}^2 m_2 e^{i\phi_2} + s_{13}^2 m_3 e^{i\phi_3}\right|$$

3) **Cosmology**: $m^2_i \neq 0$ can affect large scale structures in (standard) cosmology constrained by CMB + other data. Sensitive to:

$$\Sigma = m_1 + m_2 + m_3$$

Beta decay: Classic kinematic search for neutrino mass. Look at high-E endpoint Q of spectrum (depends on m^2_ν).

Tritium: low-Q, fast decays



Need good energy resolution

What is the "squared neutrino mass" in this context?

For just one (electron) neutrino family: $m^2(\nu_e)$ (obsolete terminology!)

For three neutrino families ν_i, and individual masses experimentally unresolved in beta decay: sensitivity to the sum of $m^2(\nu_i)$, weighted by squared mixings $|U_{ei}|^2$ with the electron neutrino. Observable:

$$m_\beta = \left[c_{13}^2 c_{12}^2 m_1^2 + c_{13}^2 s_{12}^2 m_2^2 + s_{13}^2 m_3^2\right]^{\frac{1}{2}}$$

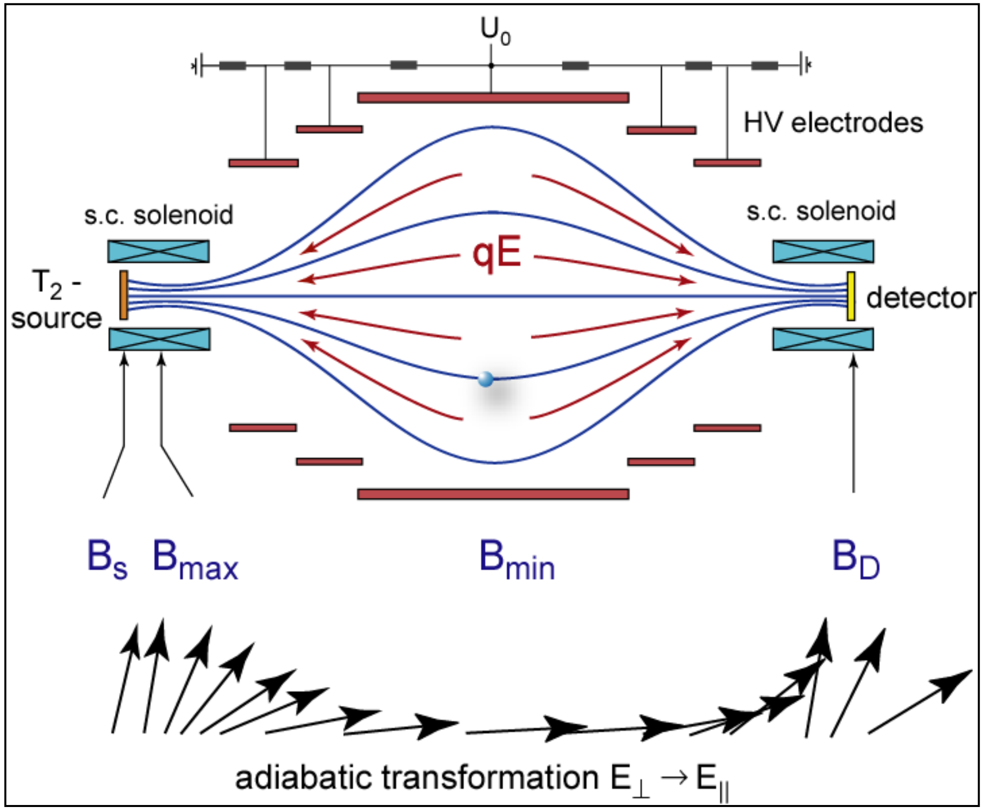
(so-called "effective electron neutrino mass")

Note: mass state with largest electron flavor component is ν_1:

$$|U_{e1}|^2 \approx \cos^2\theta_{12} \approx 0.7$$

... and we can't exclude that ν_1 is ~massless in normal hierarchy.

Tritium experiment in construction: KATRIN



Very good energy measurement and resolution via

Magnetic Adiabatic Collimation with an Electrostatic Filter

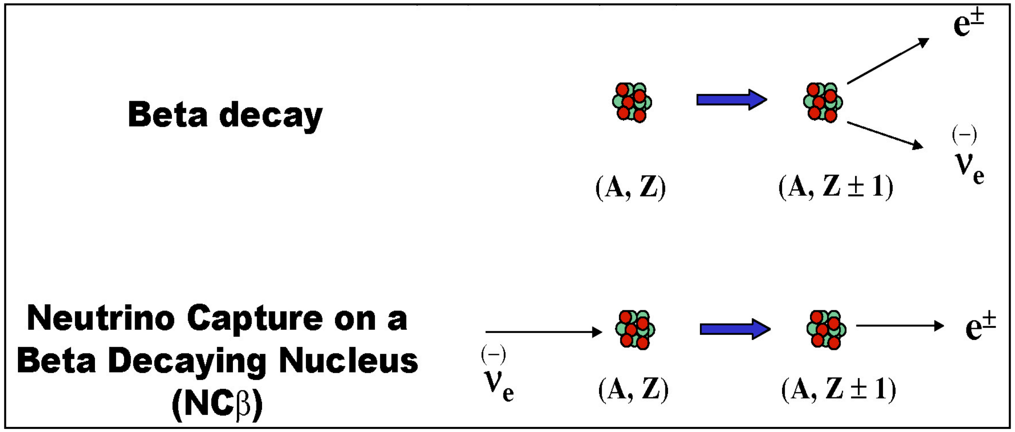
Probably the "ultimate" spectrometer of this kind....

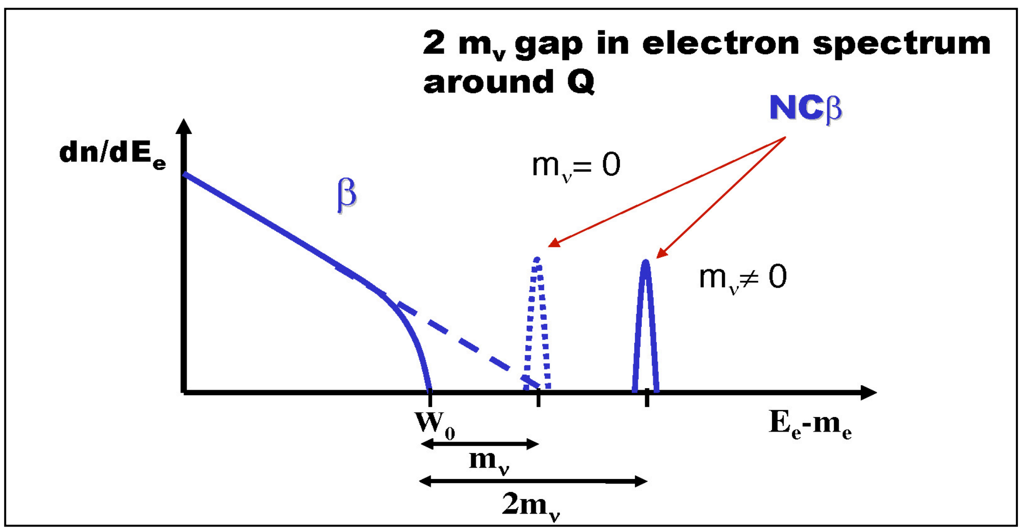
Expected to reach

m_β < 0.2 eV

(~1/10 than current limits)

β decay: need new ideas to go beyond KATRIN (calorimetry?). Very far future … a possible observation of the relic neutrino bkgd ?





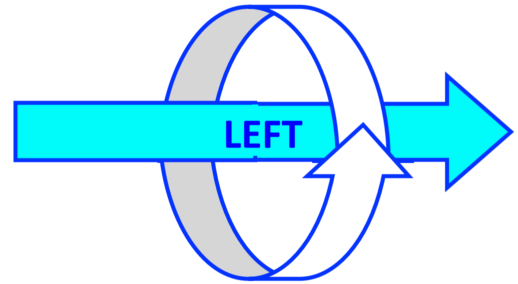
(PTOLEMY)

Neutrinoless double beta decay - Basics

Weak interactions are chiral (= not mirror-symmetric):

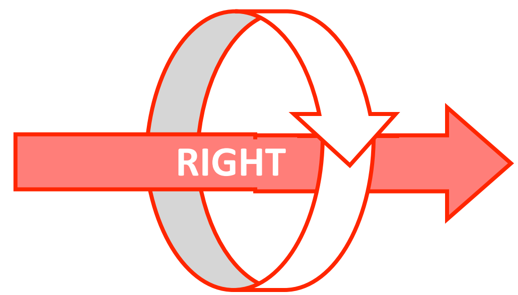
Neutrinos are created in
a left-handed (LH) state

ν

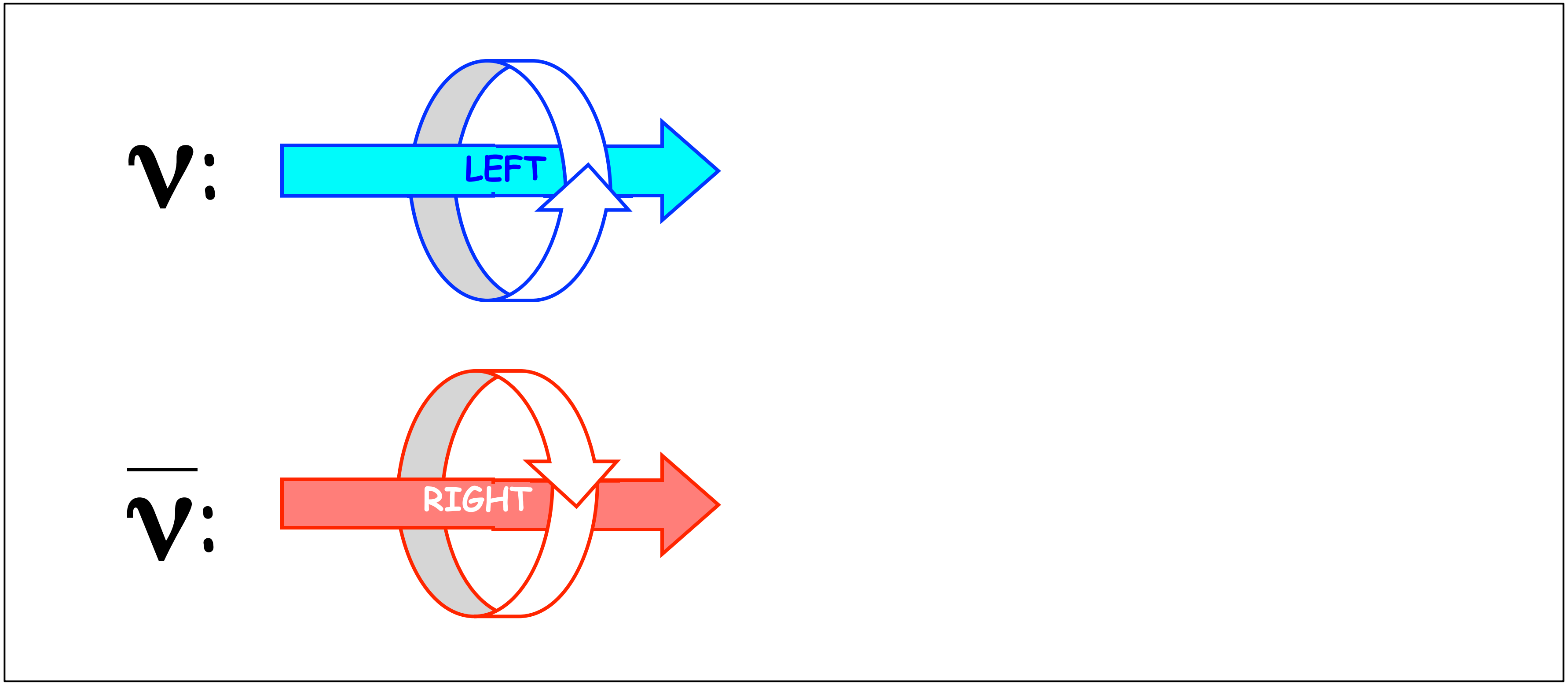


Anti-nus are created in
a right-handed (RH) state

$\bar{\nu}$

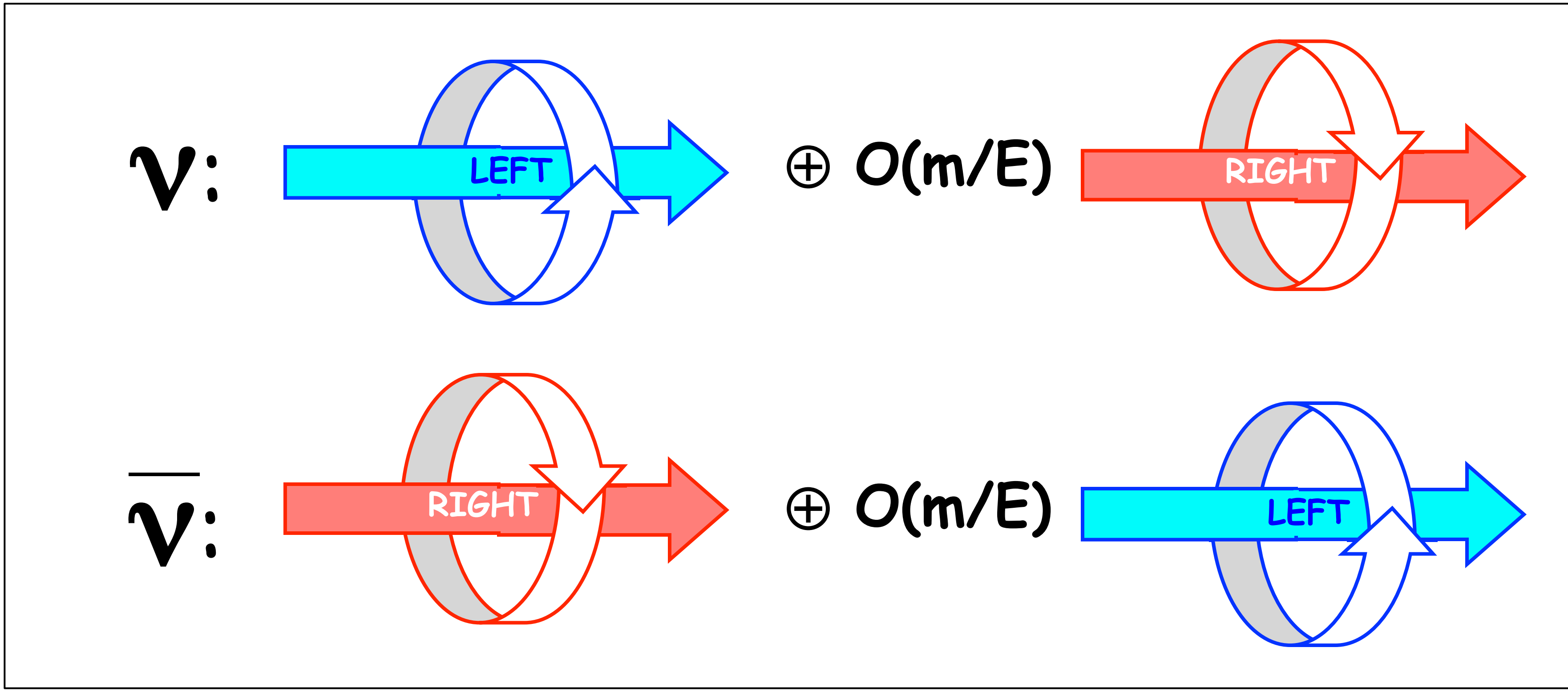


For massless neutrinos: handedness is a constant of motion



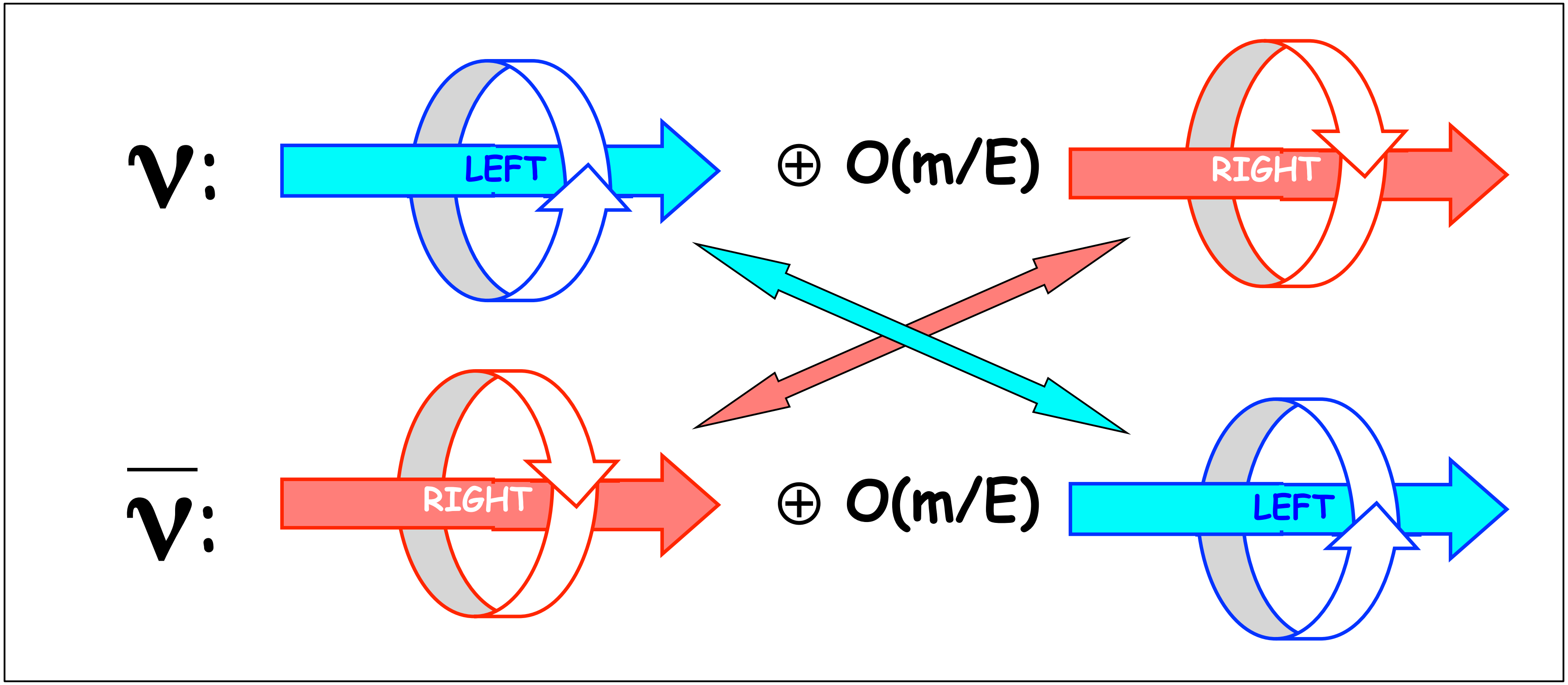
2 independent d.o.f.: massless ("Weyl") 2-spinor

But: massive ν can develop the "wrong" handedness at O(m/E)
(the Dirac equation mixes RH and LH states for $m_\nu \neq 0$):



If these 4 d.o.f. are independent: massive ("Dirac") 4-spinor
[→ Distinction between neutrinos and antineutrinos, as for electrically charged fermions. Can define a "lepton number"]

But, for neutral fermions, 2 components might be identical !



Massive ("Majorana") 4-spinor with 2 independent d.o.f.
[No distinction between nu / antinu, up to a CPV "Majorana phase":
A *very* neutral particle: no electric charge, no leptonic number…]

Exercise: The $\nu = \bar{\nu}$ paradox for Majorana neutrinos.

- We can define ν_e as the particle emitted in β^+ decay: $(A,Z) \to (A,Z-1)\,e^+\,\nu_e$
- We can define $\bar{\nu}_e$ as the particle emitted in β^- decay: $(A,Z) \to (A,Z+1)\,e^-\,\bar{\nu}_e$

- The following reactions have been experimentally observed:

$$\boxed{\nu_e + n \to p + e^-} \qquad \boxed{\bar{\nu}_e + p \to n + e^+}$$

- The following reactions have not been experimentally observed:

$$\boxed{\bar{\nu}_e + n \to p + e^-} \qquad \boxed{\nu_e + p \to n + e^+}$$

- This makes sense if ν's are Dirac, since $\nu_e \neq \bar{\nu}_e$. One can attach a "leptonic number" to the doublets (ν_e, e^-) and $(\bar{\nu}_e, e^+)$, which is conserved in the observed reactions ($\Delta L = 0$) and would be violated in the other two ($\Delta L = 2$).

- Try to make sense of this fact for Majorana ν's ($\nu = \bar{\nu}$)

Solution – For Majorana neutrinos ($\nu = \bar{\nu}$) the unobserved reactions can actually take place! But they are suppressed by many orders of magnitude, $\mathcal{O}(m/E)$.

Indeed, for Majorana ν, we are just naming:

"ν_e" = LH component of ν state,

"$\bar{\nu}_e$" = RH component of ν state.

The initial ν state produced in β^+ decay is LH (weak inter.) and is thus "ν_e". While propagating, it remains dominantly LH, but can develop a small RH component ("$\bar{\nu}_e$") at $\mathcal{O}(m/E)$. Then, also the reaction $\bar{\nu}_e + n \to p + e^-$ can occur in principle, but it is so suppressed to be practically unobservable. In other words, lepton number violation ($\Delta L = 2$) is allowed in principle, but suppressed at $\mathcal{O}(m/E)$ in practice.

Summary of options for neutrino spinor field:

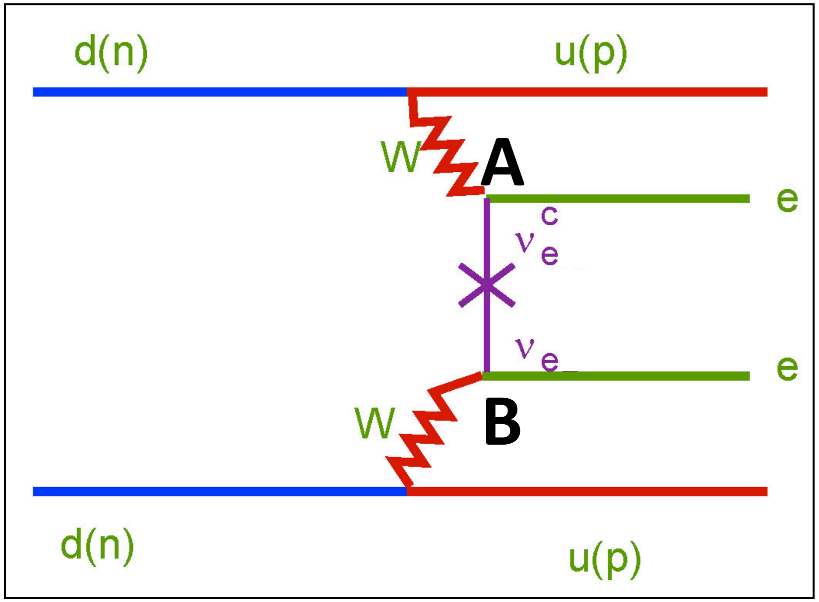
m=0, Weyl:	$\psi = \psi_R$ or $\psi = \psi_L$	massless field with 2 d.o.f.
m≠0, Majorana:	$\psi = \psi_R + \psi_R^c = \psi^c$ or $\psi = \psi_L + \psi_L^c = \psi^c$	massive field with 2 d.o.f.
m≠0, Dirac:	$\psi = \psi_R + \psi_L \neq \psi_c$	massive field with 4 d.o.f.

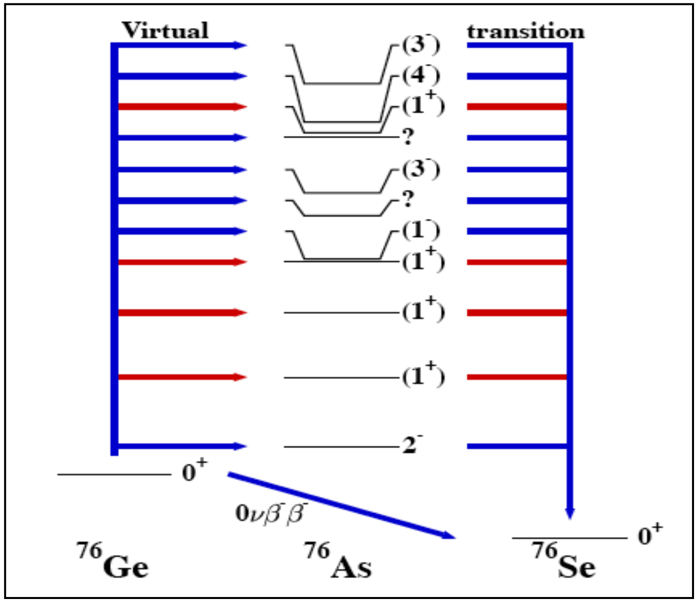
Conjugation operator: $\psi^c = \mathcal{C}(\psi) = i\gamma^2\psi^*$, $\psi_{\text{antiparticle}} = \mathcal{C}(\psi_{\text{particle}})$

[Later: Majorana masses and "see-saw" mechanism to explain their smallness]

"Unique" experimental handle to Majorana neutrinos →

Neutrinoless double beta decay: (A,Z) → (A,Z+2)+2e





Can occur only for Majorana neutrinos. Intuitive picture:

1) A RH antineutrino is emitted at point "**A**" together with an electron
2) If it is massive, at O(m/E) it develops a LH component (not possible if Weyl)
3) If neutrino=antineutrino, this component is a LH neutrino (not possible if Dirac)
4) The LH (Majorana) neutrino is absorbed at "**B**" where a 2nd electron is emitted

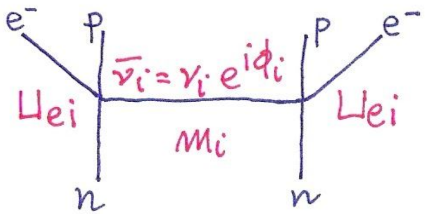
[EW part is "simple". Nuclear physics part is rather complicated and uncertain.]

Exercise: Probability of $0\nu\beta\beta$ decay and $m_{\beta\beta}$ (effective Majorana mass)

Amplitude of $0\nu\beta\beta$ decay:

- Depends on ν_e mixings U_{ei} with ν_i
- Is proportional to ν_i masses m_i (being a m/E effect)
- Depends on generalized Majorana conditions (and phases): $\bar{\nu}_i = \nu_i e^{i\phi_i}$

Probability of decay:

$\left| \sum_{i=1}^{3} \right.$  $\left. \right|^2 \propto \left| \sum_{i=1}^{3} U_{ei}^2 m_i e^{i\phi_i} \right| \equiv m_{\beta\beta}^2$ ← effective Majorana mass in $0\nu\beta\beta$ decay

Since there is an absolute value, only 2 out of 3 Majorana phases are physical ("relative phases"); typical notation:

$$m_{\beta\beta} = \left| c_{13}^2 c_{12}^2 m_1 + c_{13}^2 s_{12}^2 m_2 e^{i\phi_2} + s_{13}^2 m_3 e^{i\phi_3} \right|$$

Note that $e^{i\phi_{2,3}}$ may be also equal to -1 and, in general, may induce cancellations in $m_{\beta\beta}$.

Exercise: Majorana/Dirac "confusion" in ν oscillations

For Majorana ν's, the mixing matrix U is generalized as:

$U \to U \cdot U_M$, where $U_M = \mathrm{diag}(1, e^{i\phi}, e^{i\phi'})$ contains two independent Majorana phases.

In the hamiltonian of ν oscillations (either in vacuum or in matter), the mixing matrix always appears in the form:

$U \frac{\mathcal{M}^2}{2E} U^+$, with $\mathcal{M}^2 = \mathrm{diag}(m_1^2, m_2^2, m_3^2)$.

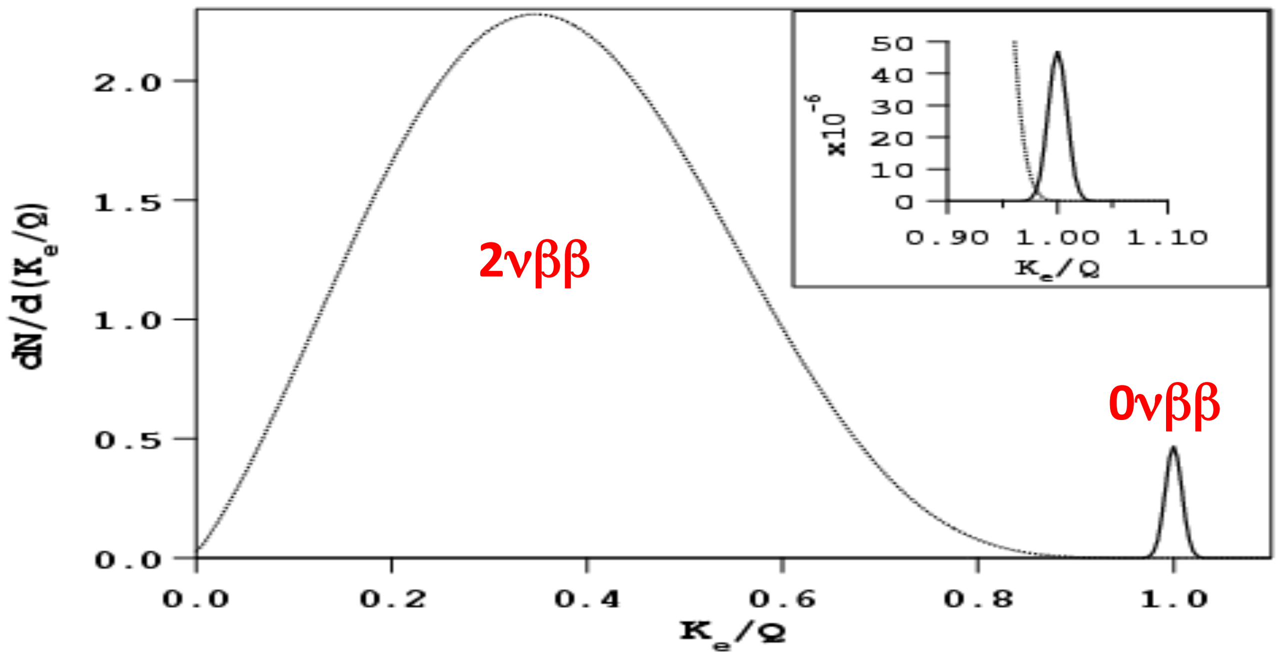
The replacement $U \to U U_M$ is then ineffective:

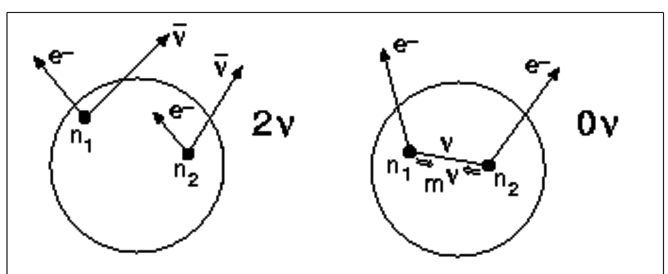
$$U \frac{\mathcal{M}^2}{2E} U^+ \to U U_M \frac{\mathcal{M}^2}{2E} U_M^+ U = U \frac{\mathcal{M}^2}{2E} U^+ .$$

Thus, oscillations do not distinguish Dirac/Majorana neutrinos ("confusion" of the two possibilities).

Experimentally: Look at sum energy of both electrons

Need to see the **0νββ** line emerge above bkgd, at endpoint spectrum of "conventional" **2νββ** decay.



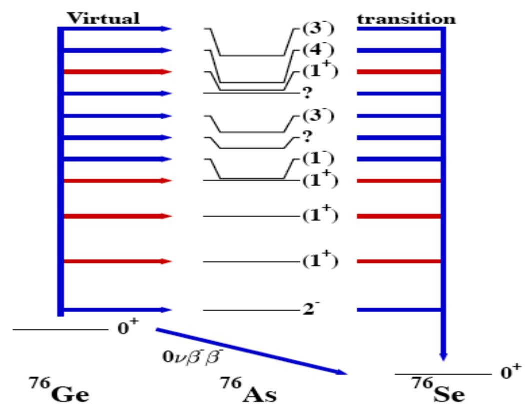


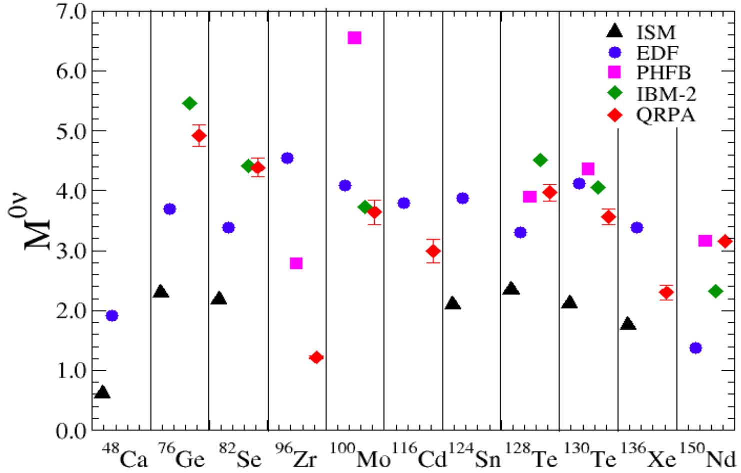
What sets the uncertainty of $m_{\beta\beta}$?

In case of positive signal, a major concern is the accuracy of the **nuclear matrix element |M|**, rather than the expt. uncertainty on the decay half life:

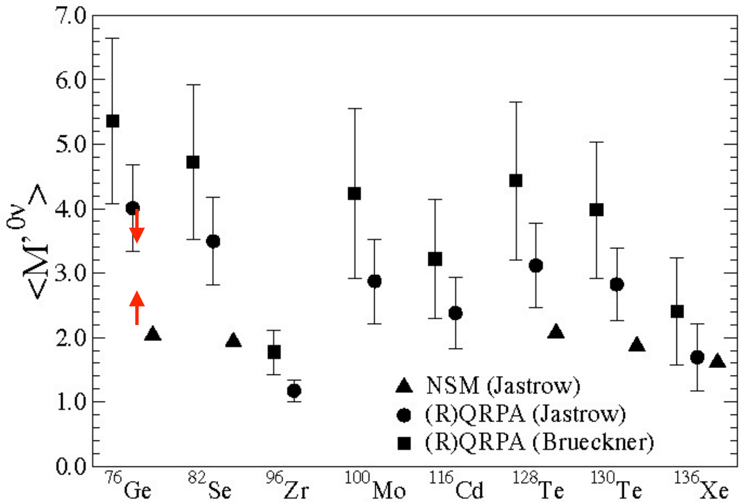
$$T_i^{-1} = G_i \, |M_i'|^2 \, m_{\beta\beta}^2$$

Half-life Phase space Matrix element





Luckily, independent nuclear physics models are slowly converging, better than it could be hoped only a few years ago …



from: Simkovic

… especially when using the same theo. inputs for comparison (e.g, same description of short range nucleon repulsion) and exploiting additional data to constrain models.

BUT: errors remain large for each candidate nucleus.

Many runners in the race to discover $0\nu\beta\beta$ decay...

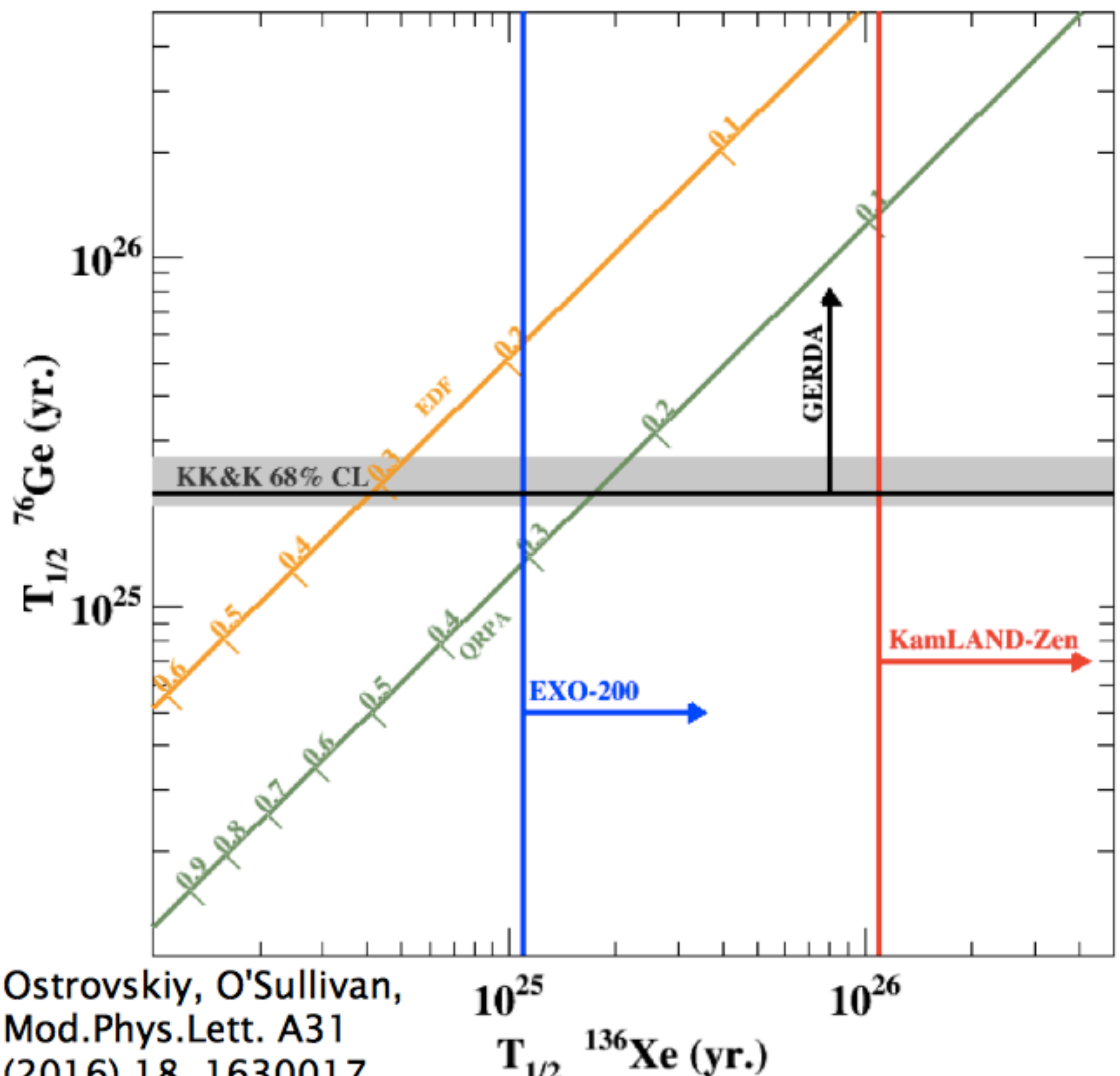
TABLE VII. In this table, the main features and performances of some past, present and future $0\nu\beta\beta$ experiments are listed.

Experiment	Isotope	Techinique	Total mass [kg]	Exposure [kg yr]	FWHM @ $Q_{\beta\beta}$ [keV]	Background [counts/keV/kg/yr]	$S^{0\nu}$ (90% C. L.) [10^{25} yr]
Past							
Cuoricino, [179]	^{130}Te	bolometers	40.7 (TeO_2)	19.75	5.8 ± 2.1	0.153 ± 0.006	0.24
CUORE-0, [180]	^{130}Te	bolometers	39 (TeO_2)	9.8	5.1 ± 0.3	0.058 ± 0.006	0.29
Heidelberg-Moscow, [181]	^{76}Ge	Ge diodes	11 (^{enr}Ge)	35.5	4.23 ± 0.14	0.06 ± 0.01	1.9
IGEX, [182, 183]	^{76}Ge	Ge diodes	8.1 (^{enr}Ge)	8.9	~ 4	$\lesssim 0.06$	1.57
GERDA-I, [167, 184]	^{76}Ge	Ge diodes	17.7 (^{enr}Ge)	21.64	3.2 ± 0.2	~ 0.01	2.1
NEMO-3, [185]	^{100}Mo	tracker + calorimeter	6.9 (^{100}Mo)	34.7	350	0.013	0.11
Present							
EXO-200, [186]	^{136}Xe	LXe TPC	175 (^{enr}Xe)	100	89 ± 3	$(1.7 \pm 0.2) \cdot 10^{-3}$	1.1
KamLAND-Zen, [187, 188]	^{136}Xe	loaded liquid scintillator	348 (^{enr}Xe)	89.5	244 ± 11	~ 0.01	1.9
Future							
CUORE, [189]	^{130}Te	bolometers	741 (TeO_2)	1030	5	0.01	9.5
GERDA-II, [174]	^{76}Ge	Ge diodes	37.8 (^{enr}Ge)	100	3	0.001	15
LUCIFER, [190]	^{82}Se	bolometers	17 (Zn^{82}Se)	18	10	0.001	1.8
MAJORANA D., [191]	^{76}Ge	Ge diodes	44.8 ($^{enr/nat}$Ge)	100 [a]	4	0.003	12
NEXT, [192, 193]	^{136}Xe	Xe TPC	100 (^{enr}Xe)	300	$12.3 - 17.2$	$5 \cdot 10^{-4}$	5
AMoRE, [194]	^{100}Mo	bolometers	200 (Ca^{enr}MoO$_4$)	295	9	$1 \cdot 10^{-4}$	5
nEXO, [195]	^{136}Xe	LXe TPC	4780 (^{enr}Xe)	12150 [b]	58	$1.7 \cdot 10^{-5}$ [b]	66
PandaX-III, [196]	^{136}Xe	Xe TPC	1000 (^{enr}Xe)	3000 [c]	$12 - 76$	0.001	11 [c]
SNO+, [197]	^{130}Te	loaded liquid scintillator	2340 (^{nat}Te)	3980	270	$2 \cdot 10^{-4}$	9
SuperNEMO, [198, 199]	^{82}Se	tracker + calorimeter	100 (^{82}Se)	500	120	0.01	10

[a] our assumption (corresponding sensitivity from Fig. 14 of Ref. [191]).
[b] we assume 3 tons fiducial volume.
[c] our assumption by rescaling NEXT.

from 1601.07512

Examples of half-life limits and their impact on $m_{\beta\beta}$



Ostrovskiy, O'Sullivan, Mod.Phys.Lett. A31 (2016) 18, 1630017

Cosmology: a "modern" probe

Standard big bang cosmology predicts a relic neutrino background with total number density $336/\mathrm{cm}^3$ and temper. $T_\nu \sim 2$ K $\sim 1.7 \times 10^{-4}$ eV $<< \sqrt{\delta m^2}, \sqrt{\Delta m^2}$.

→At least two relic neutrino species are nonrelativistic today (we can't exclude the lightest to be ~massless)

→Their total mass contributes to the normalized energy density as $\mathbf{\Omega_\nu \approx \Sigma / 50\ eV}$, where

$$\Sigma = m_1 + m_2 + m_3$$

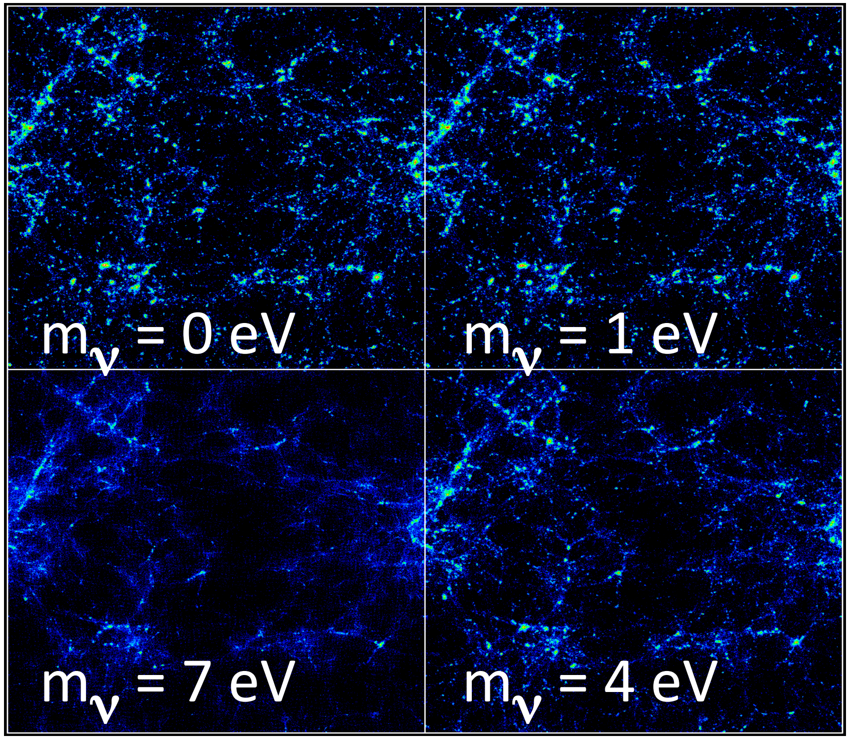
→So, if we just impose that neutrinos do not saturate the total matter density, $\mathbf{\Omega_\nu < \Omega_m \approx 0.25}$, we get

$\mathbf{m_i < 4\ eV}$ **- not bad!**

Much better bounds can be derived from neutrino effects on structure formation.

Massive neutrinos are difficult to cluster because of their relatively high velocities: they suppress matter fluctuations on scales smaller than their mass-dependent free-streaming scale.

→ Get mass-dependent suppression of small-scale structures



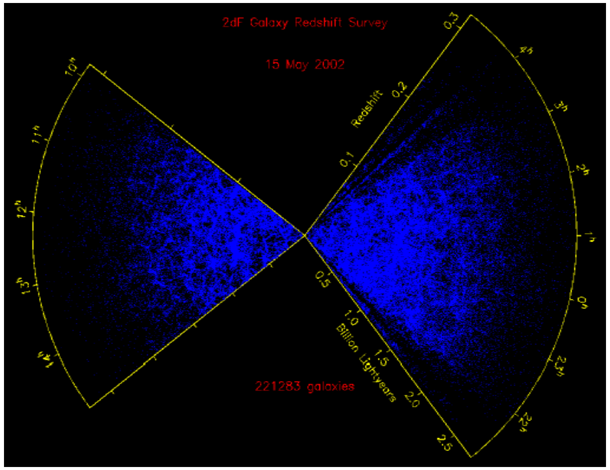
(E..g., Ma 1996)

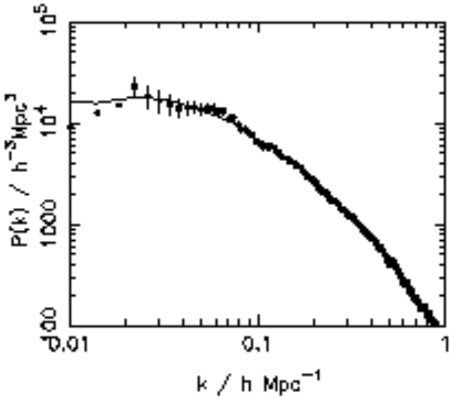
Constraints from CMB also help removing degeneracies.

Observations:

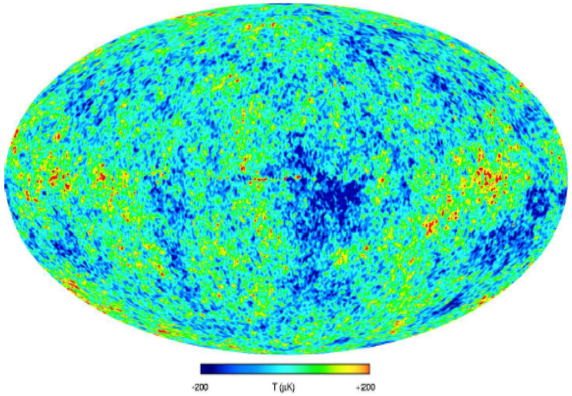
Spectra:

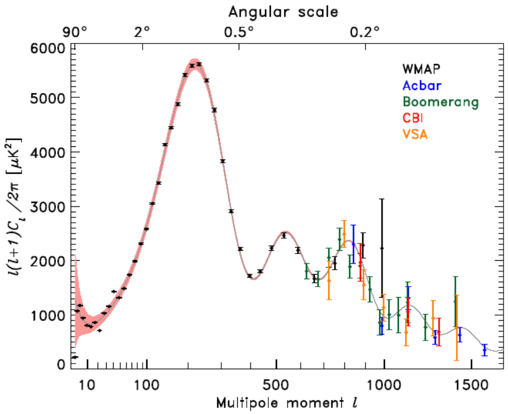
LSS



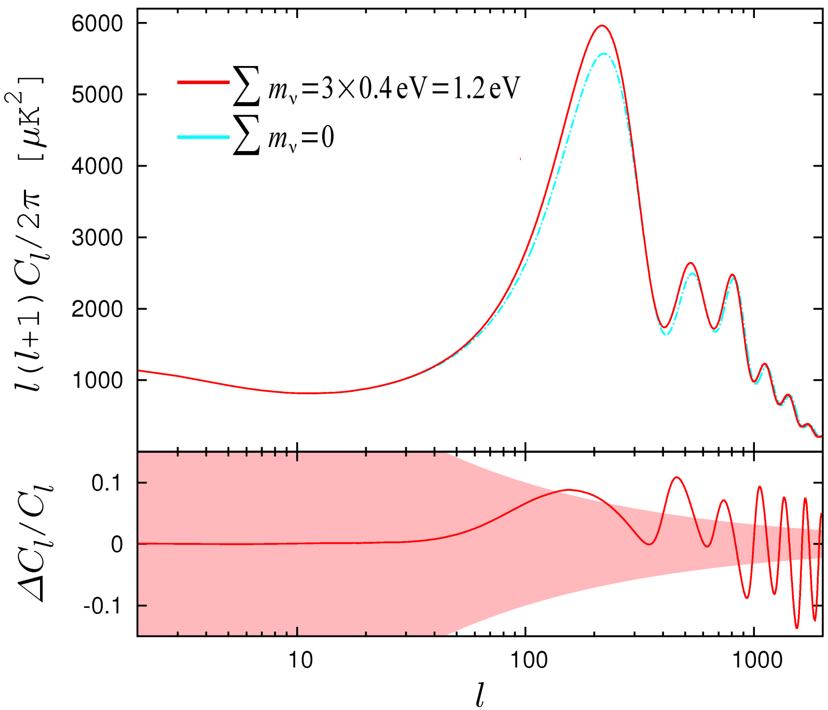


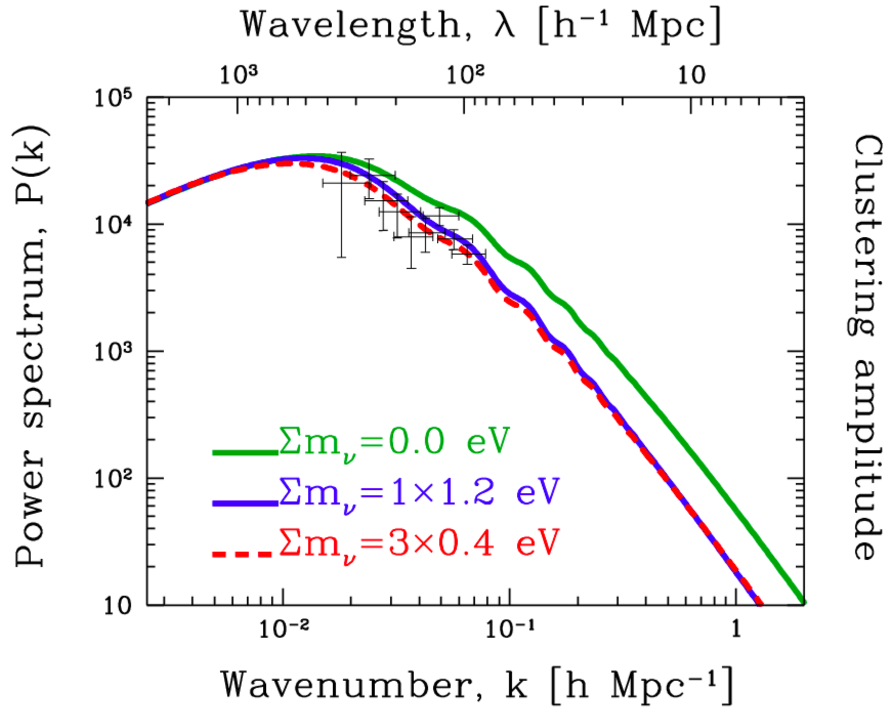
CMB





Spectral effect of massive neutrinos (e.g., from Y.Y.Y. Wong)





Significant progress after WMAP, PLANCK and recent galaxy surveys

In general, typical upper limits from current data can reach

Σ < 0.2 - 0.3 eV,

or worse if model and/or uncertainties are questioned.

Combining constraints from:

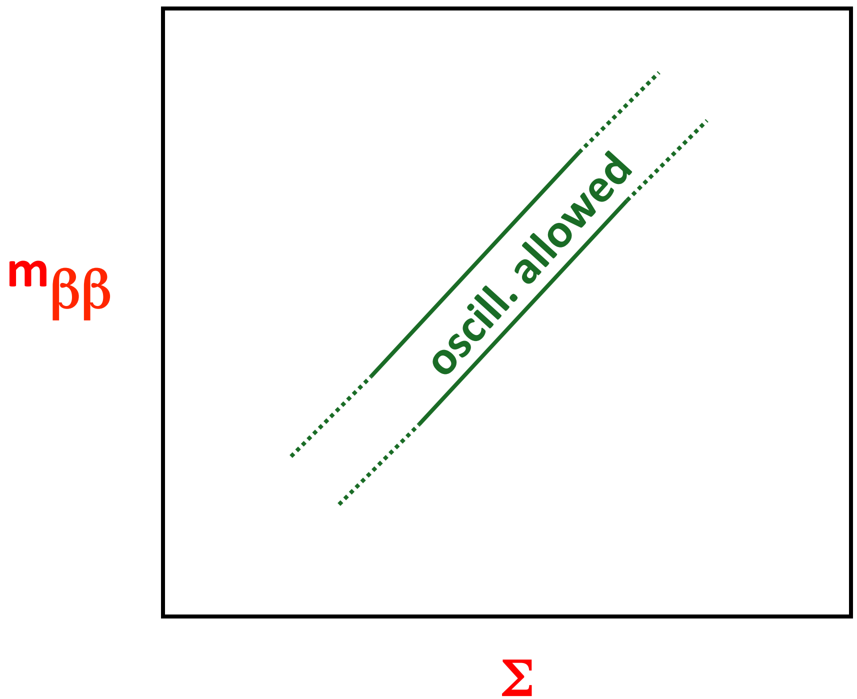
oscillations

$$m_\beta = \left[c_{13}^2 c_{12}^2 m_1^2 + c_{13}^2 s_{12}^2 m_2^2 + s_{13}^2 m_3^2\right]^{\frac{1}{2}}$$

$$m_{\beta\beta} = \left|c_{13}^2 c_{12}^2 m_1 + c_{13}^2 s_{12}^2 m_2 e^{i\phi_2} + s_{13}^2 m_3 e^{i\phi_3}\right|$$

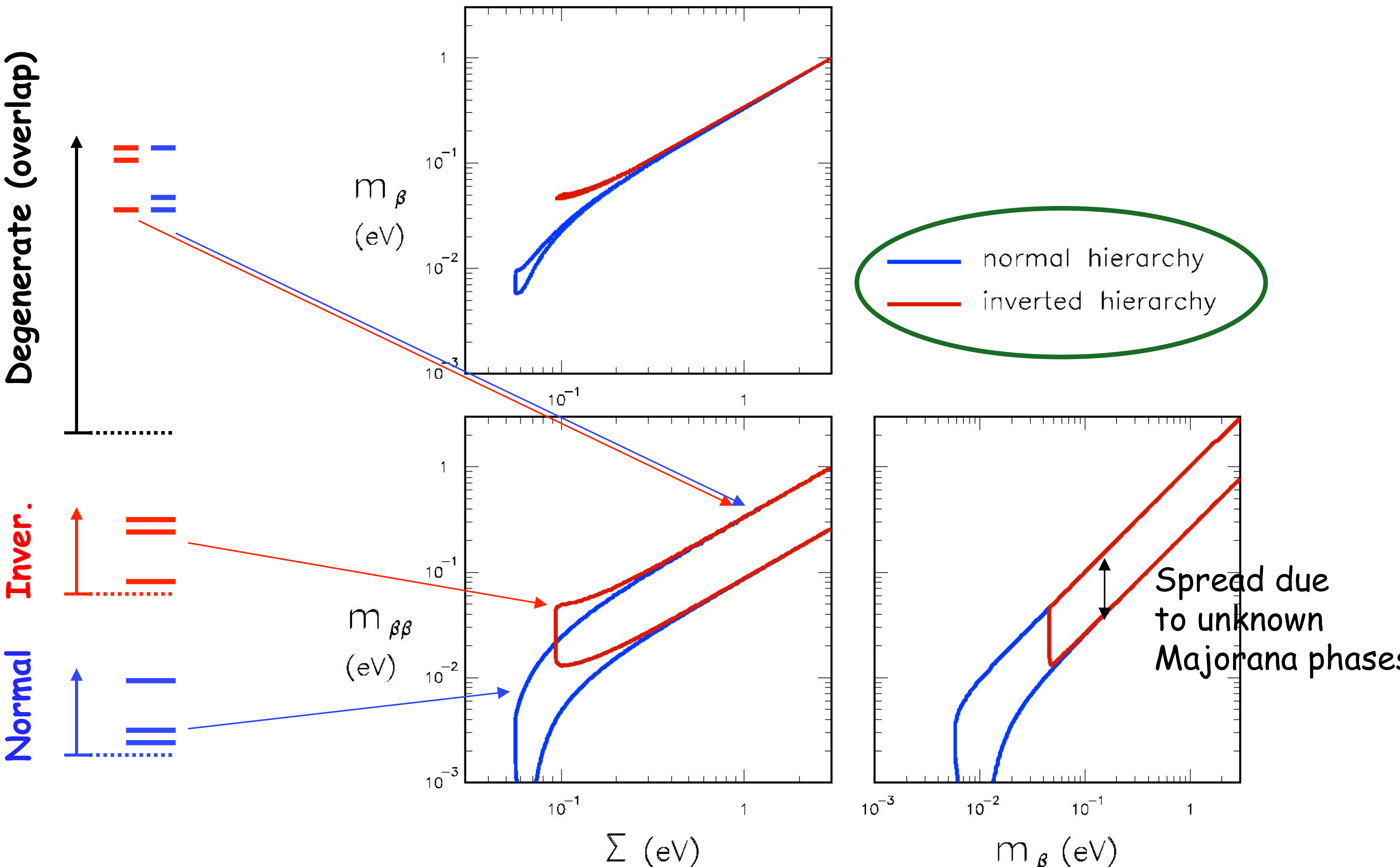
$$\Sigma = m_1 + m_2 + m_3$$

Interplay: **Oscillations** fix the **mass^2 splittings,** and thus induce **positive correlations** between any pair of the three observables **(m_β, $m_{\beta\beta}$, Σ)**, e.g.:

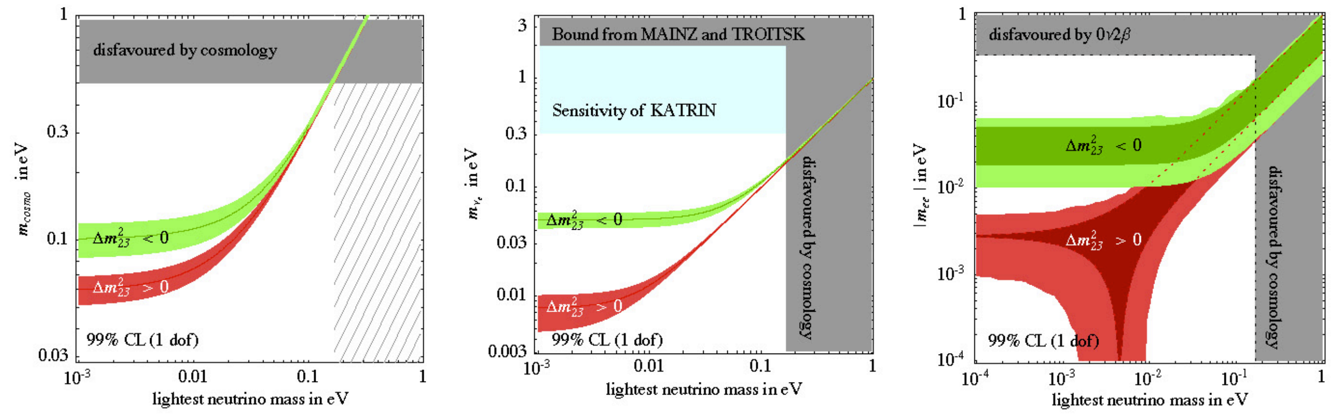


i.e., if one observable increases, the other one (typically) must increase to match mass splitting

Oscillation data constrain the $(m_\beta, m_{\beta\beta}, \Sigma)$ parameters within two bands:

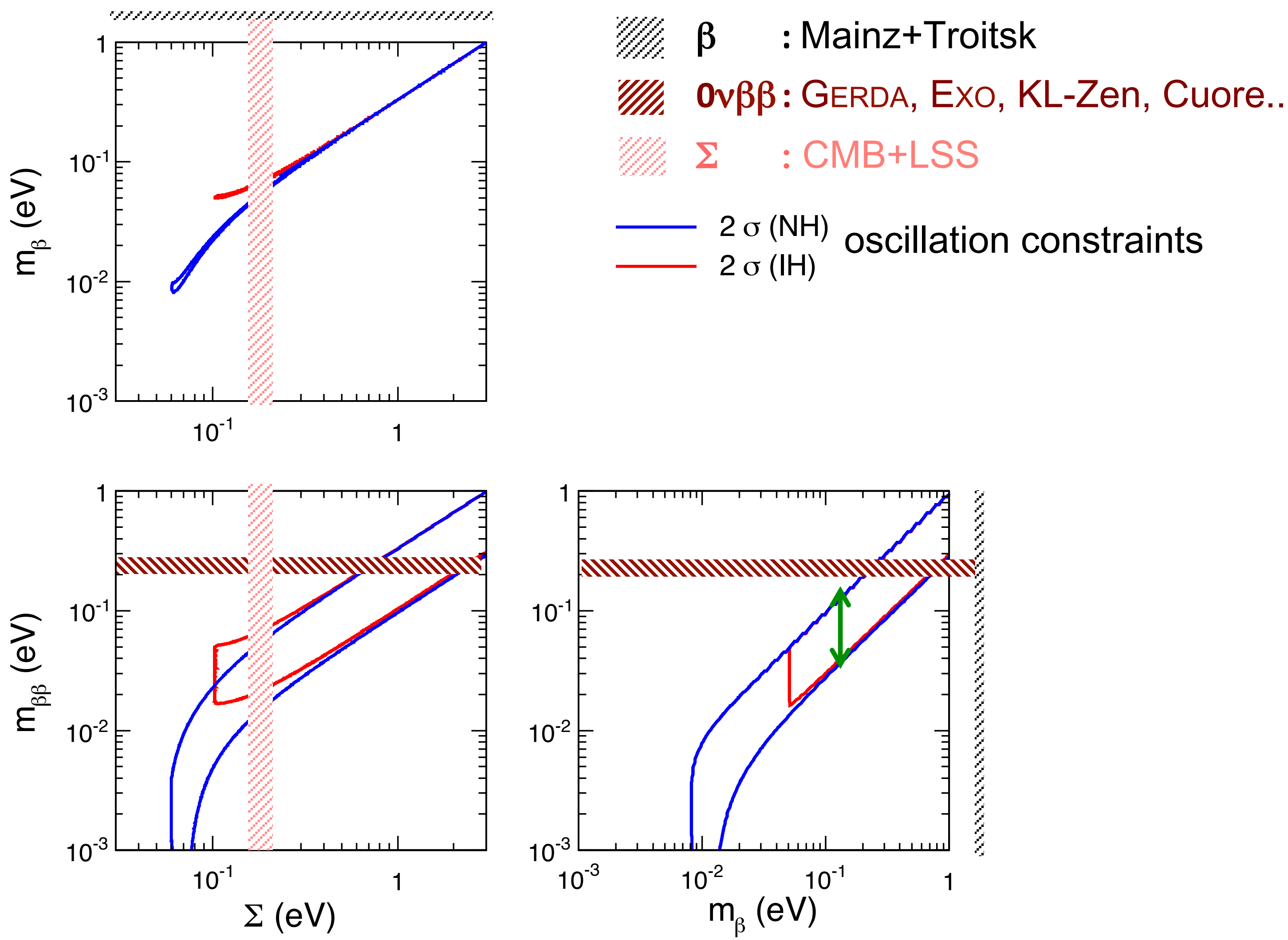


Footnote - Previous plots project away the "unobservable" lightest neutrino mass from graphs like:



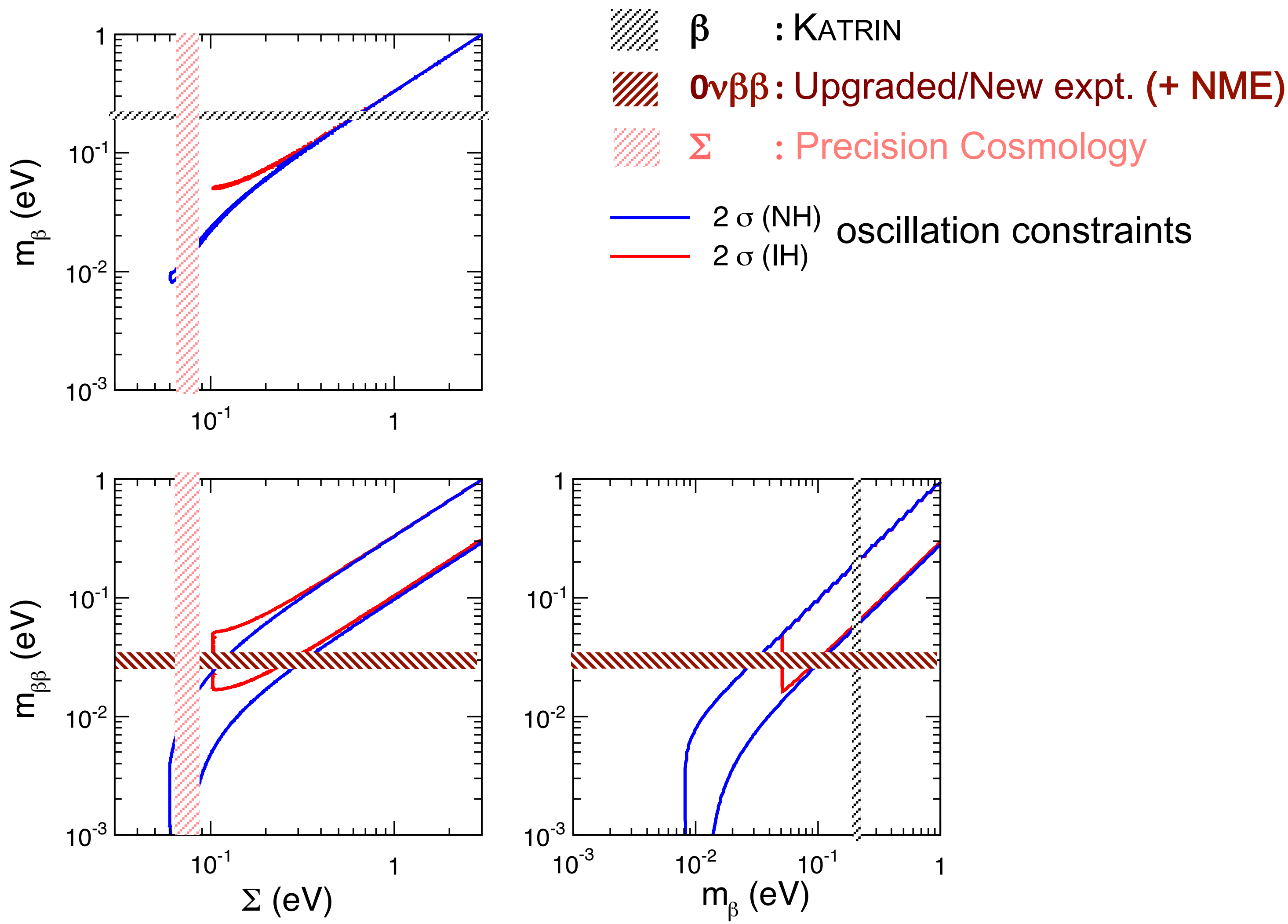
Taken from Strumia and Vissani, 2006

Upper limits on m_β, $m_{\beta\beta}$, Σ (up to some syst.) + osc. constraints



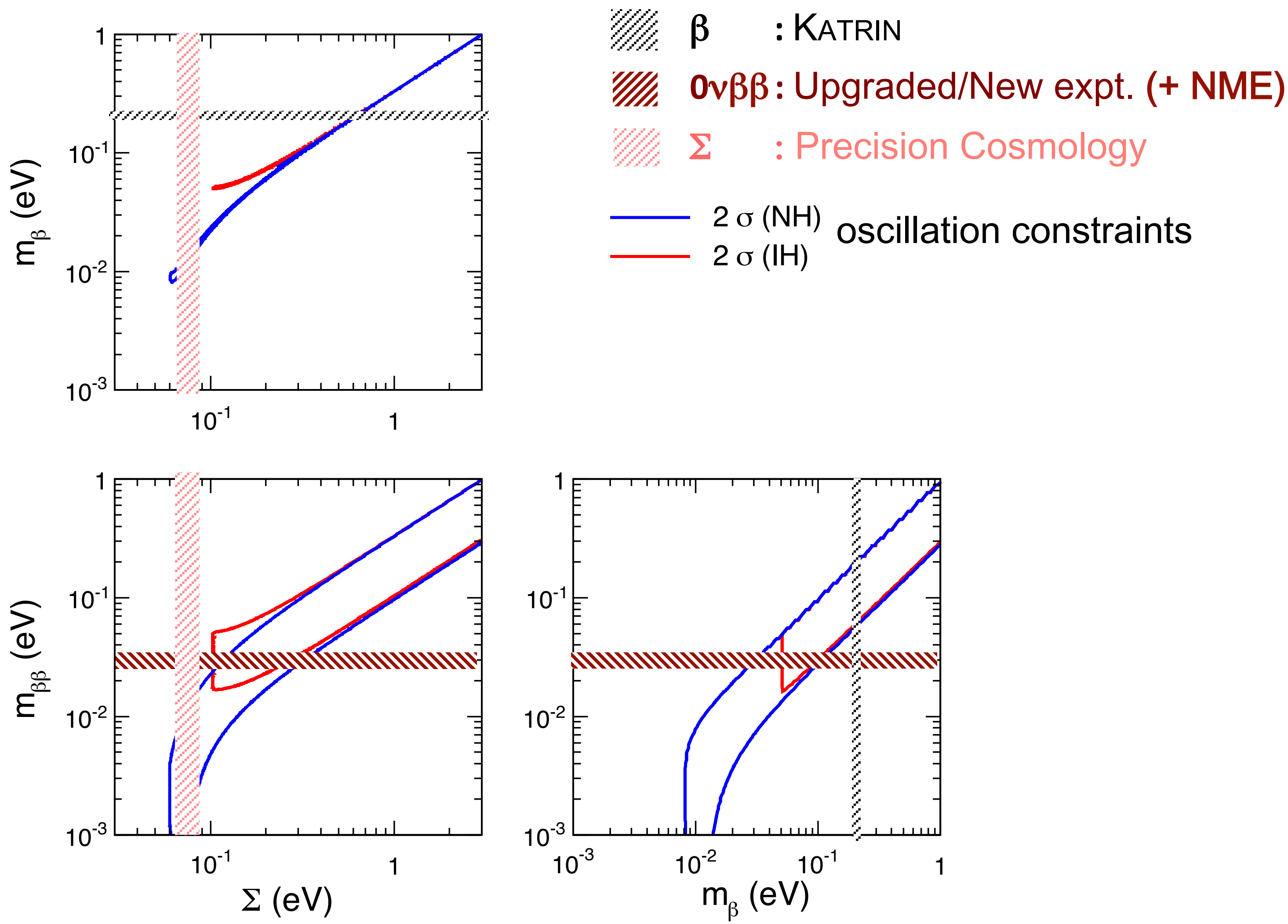
Major improvements expected in the next decade

Upper limits on m_β, $m_{\beta\beta}$, Σ in ~10 years ?



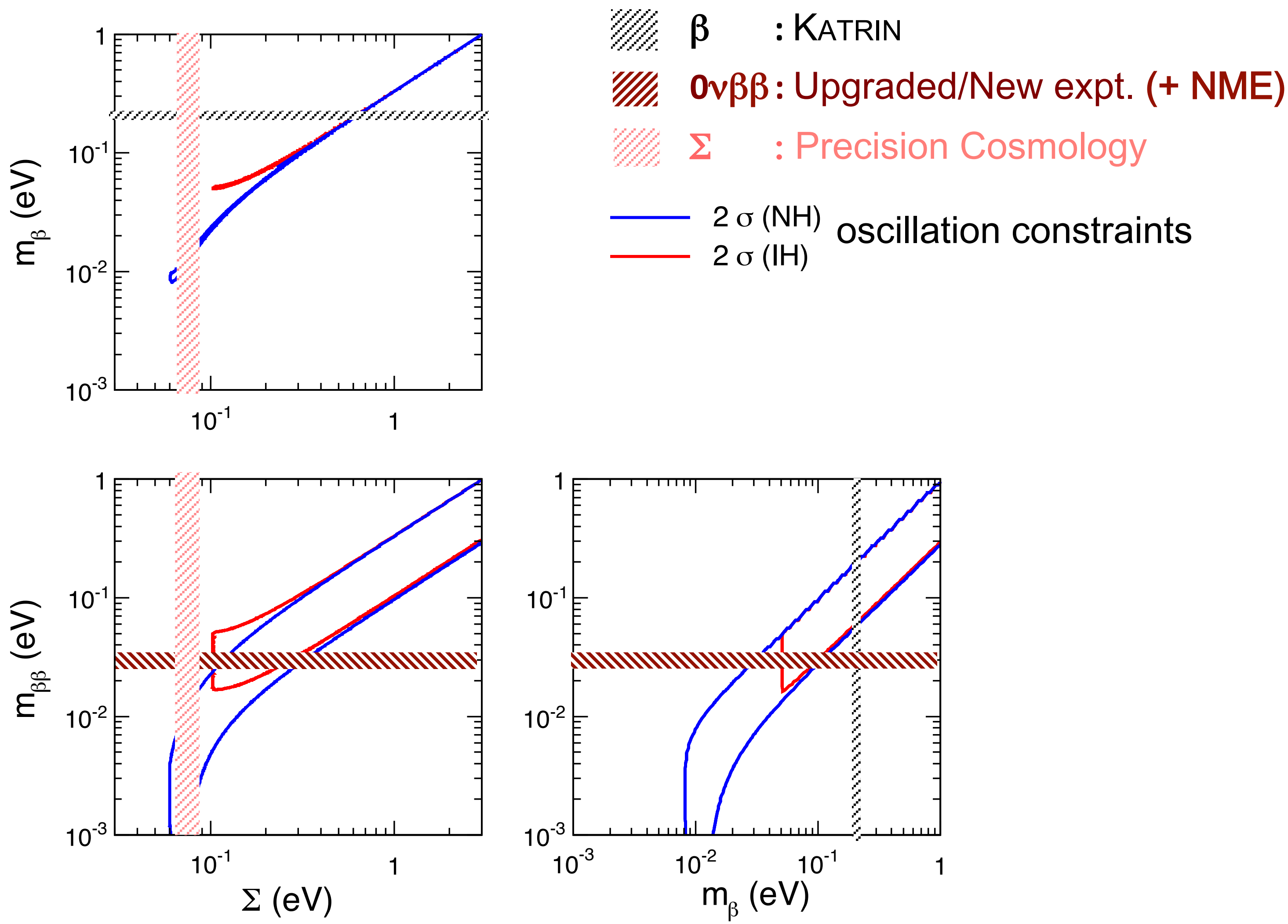
Large phase space for discoveries about ν mass and nature.

Upper limits on m_β, $m_{\beta\beta}$, Σ in ~10 years ?



Cosmology first? Be prepared to $\Sigma>0$ (or IH rejection) claims!

Upper limits on m_β, $m_{\beta\beta}$, Σ in ~10 years ?



[Even now, at face value: Nu2016 osc. data + cosmology → NH favored]

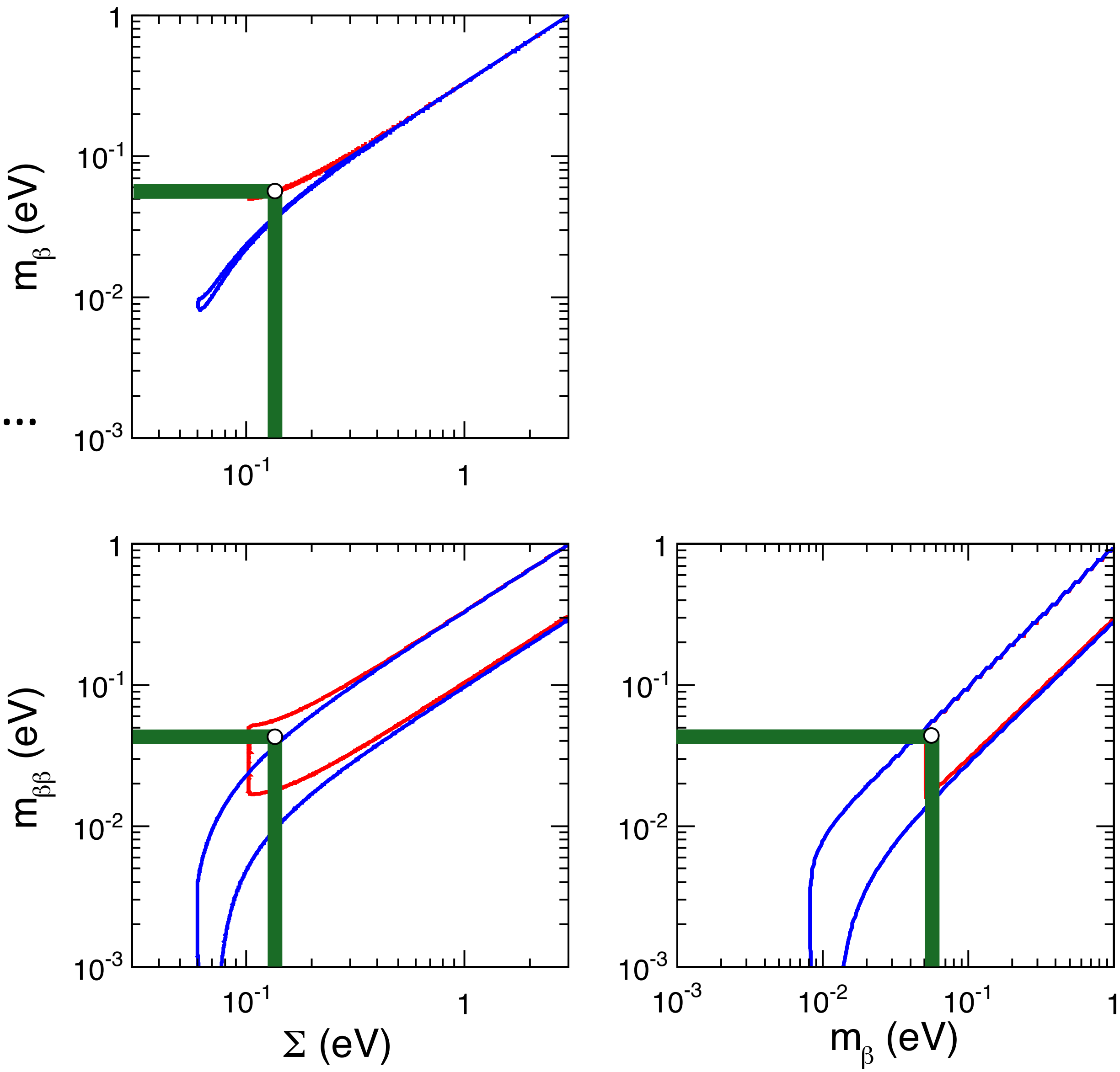
With "dreamlike" and converging data one could, e.g.

Determine the mass scale…

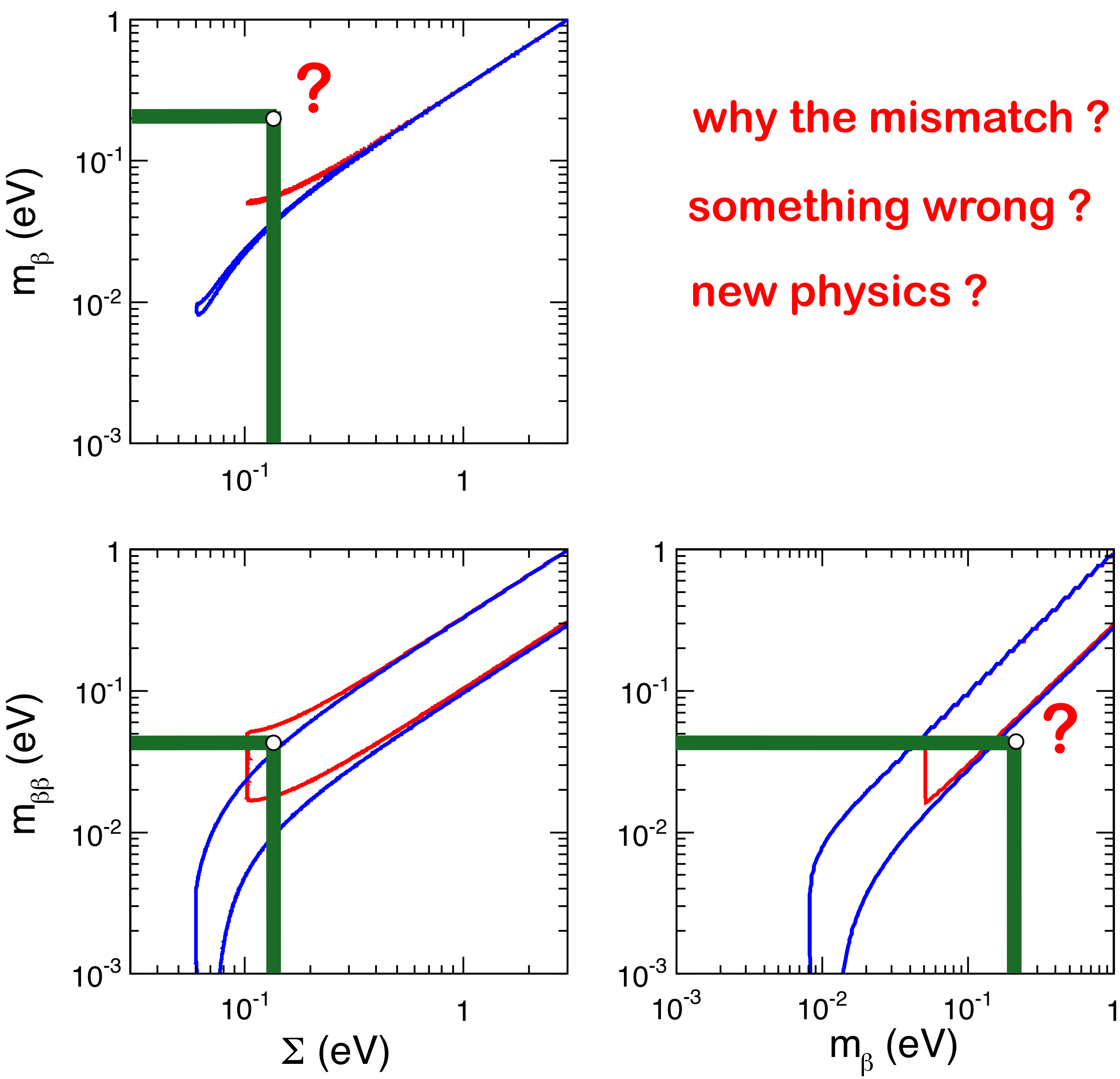
Check 3ν consistency …

Identify the hierarchy …

Probe the Majorana phase(s) …



But alternative situations might also occur….

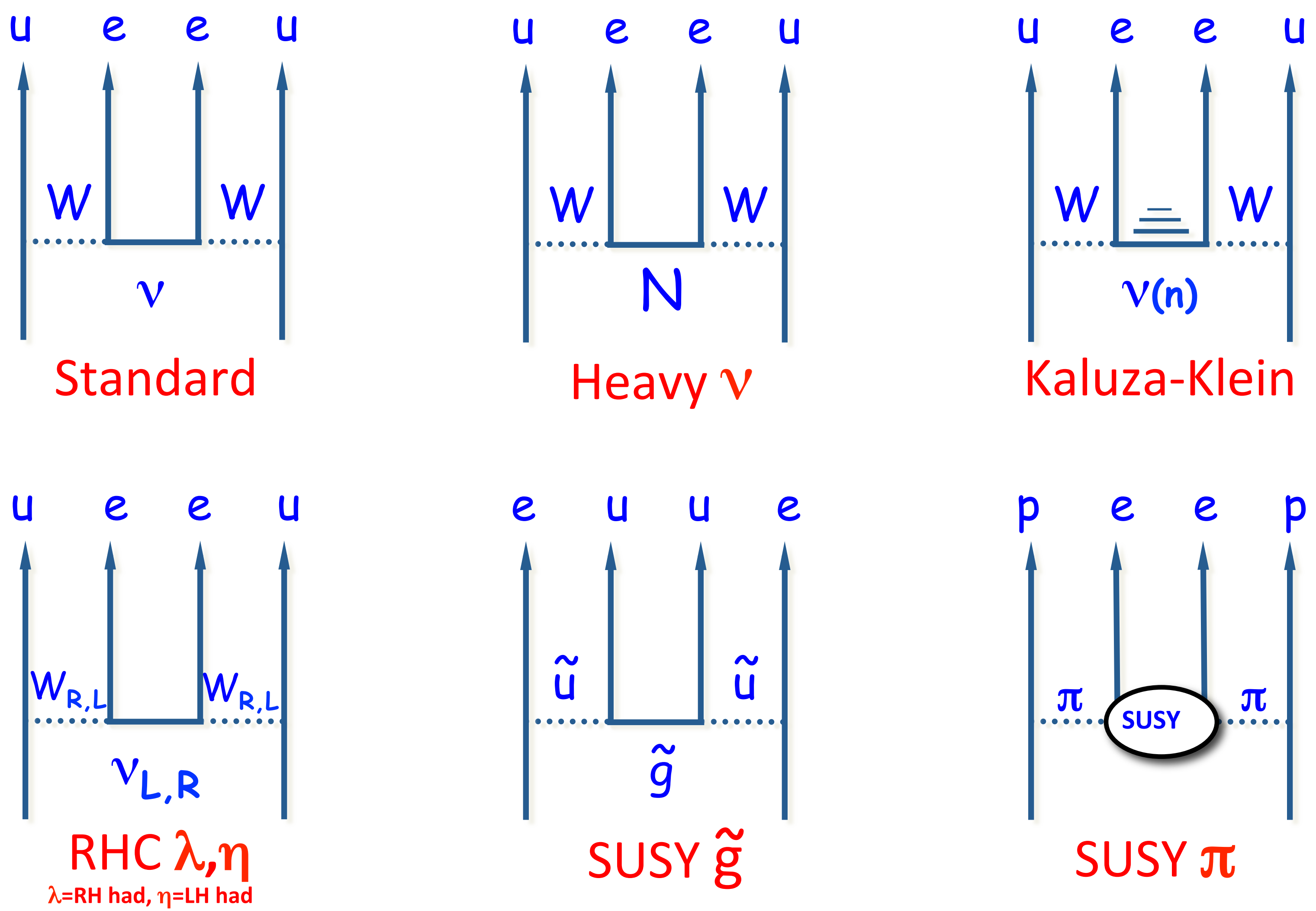


why the mismatch ?

something wrong ?

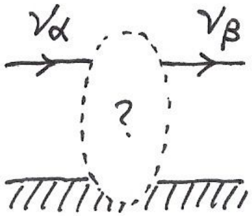
new physics ?

Physics beyond "3 light ν" should always be kept in mind:

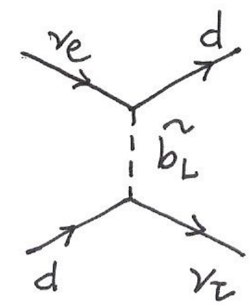


New neutrino interactions

New physics beyond the SM might also be responsible for new interactions of neutrinos, e.g., FCNC:



→ an example: SUSY $\not{R}$:



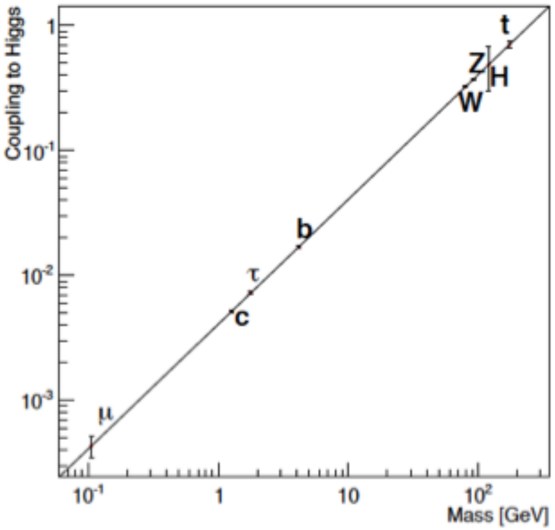
→ May get relevant modifications of the evolution hamiltonian:

$$H = \underbrace{H_{vac} + H_{mat}}_{\text{standard}} + \underbrace{H_{new\,physics}}_{\text{non standard}}$$

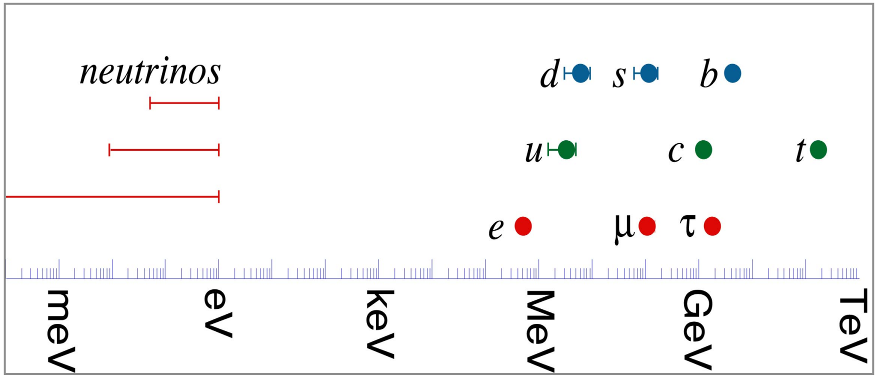
However, new coupling are tipically expected to be $\mathcal{O}(\epsilon G_F)$ with $\epsilon \ll 1$ → effects difficult to disentangle from standard oscillations.

Neutrino mass issue in the larger context of HEP:

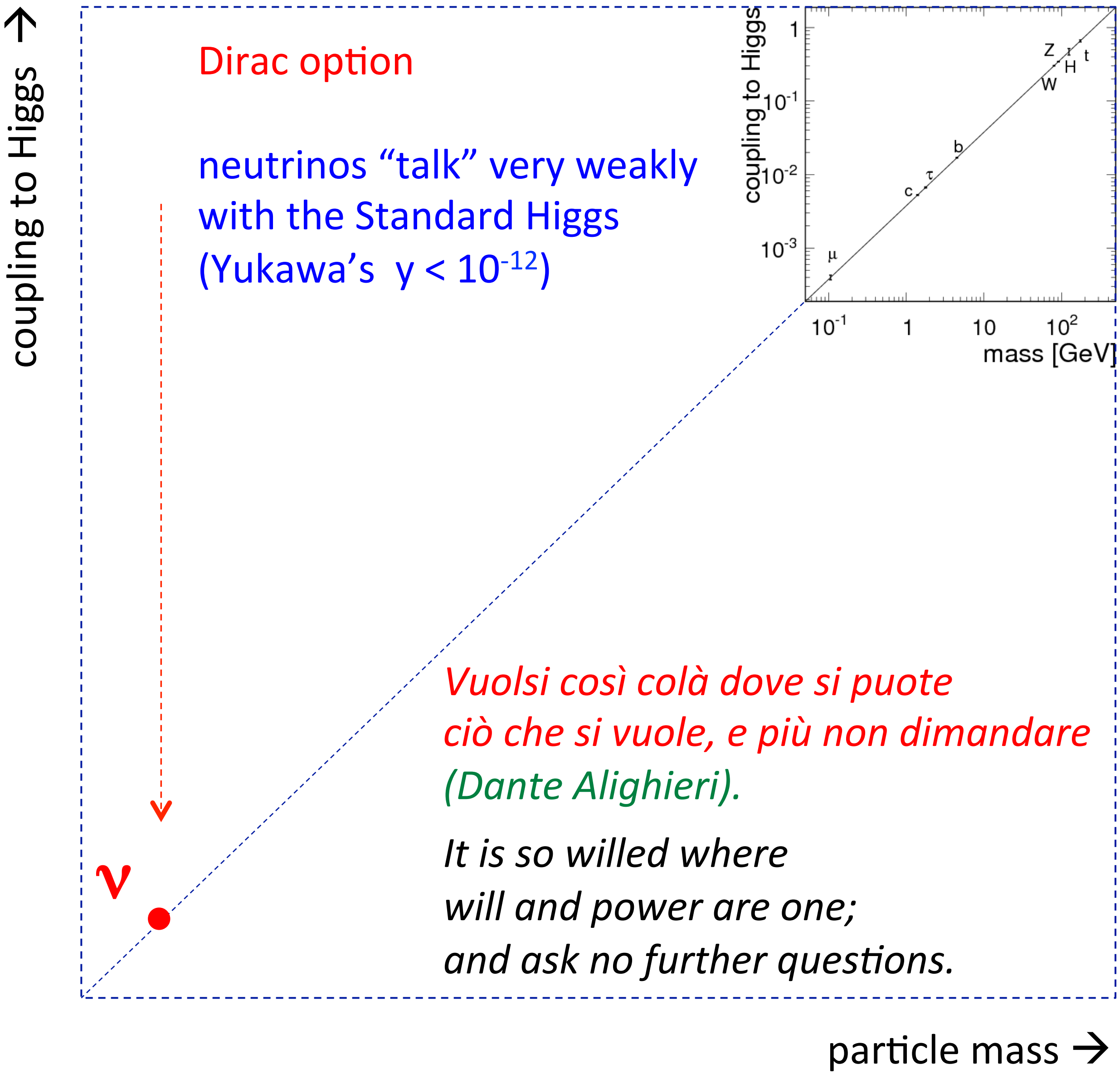
Testing Higgs sector



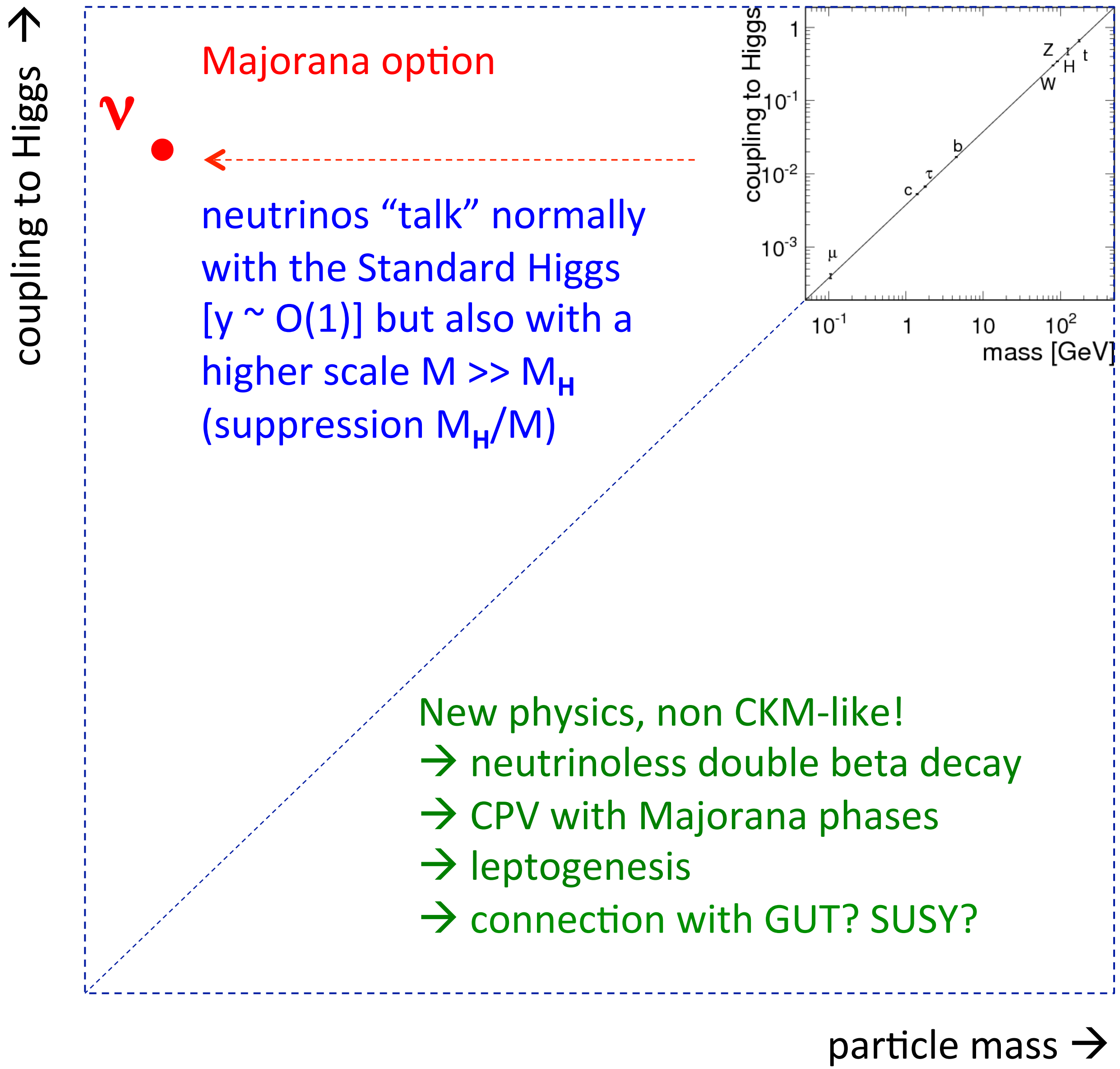
Finding ν masses



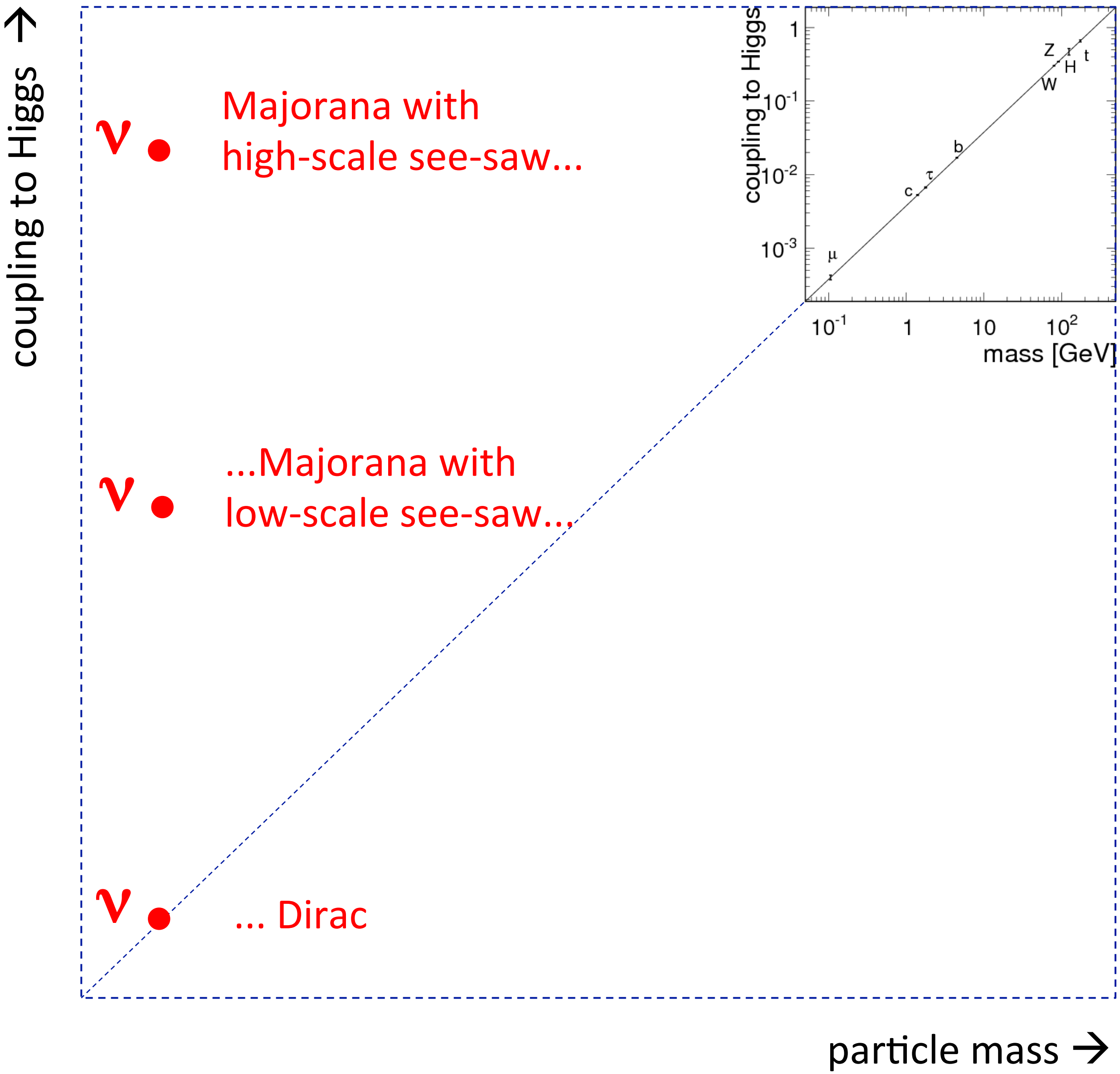
Where are neutrino masses on this plot? Options:



Another option: window to physics beyond the standard Higgs mechanism?



Large "neutrino phase space" where new physics scale(s) may show up...



More on these options...

Dirac and Majorana mass terms + See saw (1 family)

- Dirac mass terms are of the form $m\bar{\psi}\psi$ (4 dof ψ)
- Majorana " " " " " " $\frac{1}{2}m\bar{\psi}\psi$ (2 dof ψ)

Three possibilities:

Dirac : $\psi = \psi_L + \psi_R \rightarrow \bar{\psi}\psi = \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L$

Majorana (L) : $\psi = \psi_L + \psi_L^c \rightarrow \bar{\psi}\psi = \bar{\psi}_L\psi_L^c + \bar{\psi}_L^c\psi_L$

Majorana (R) : $\psi = \psi_R + \psi_R^c \rightarrow \bar{\psi}\psi = \bar{\psi}_R\psi_R^c + \bar{\psi}_R^c\psi_R$

} absent for charged fermions!

Most general mass term for one neutrino family:

$$m_D(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) + \frac{1}{2}m_L(\bar{\psi}_L\psi_L^c + \bar{\psi}_L^c\psi_L) + \frac{1}{2}m_R(\bar{\psi}_R\psi_R^c + \bar{\psi}_R^c\psi_R)$$

$$= \frac{1}{2}\left[\bar{\psi}_L + \bar{\psi}_L^c, \bar{\psi}_R + \bar{\psi}_R^c\right]\begin{bmatrix} m_L & m_D \\ m_D & m_R \end{bmatrix}\begin{bmatrix} \psi_L + \psi_L^c \\ \psi_R + \psi_R^c \end{bmatrix} \quad \text{(matrix form)}$$

In the above eq. in matrix form, the basis fields $\psi_L + \psi_L^c$ and $\psi_R + \psi_R^c$ are Majorana. Therefore, in general, diagonalization will produce mass eigenvectors which are also Majorana, despite the presence of a Dirac mass term (unless special cancellations occur).

Explicit diagonalization of $M = \begin{bmatrix} m_L & m_D \\ m_D & m_R \end{bmatrix}$

- Trace $T = \mathrm{Tr}\, M = m_L + m_R$
- Determ. $D = \det M = m_L m_R - m_D^2$
- Eigenvalues : $m_\pm = \frac{1}{2}\left(T \pm \sqrt{T^2 - 4D}\right)$
- Diagonalization angle (not a mixing angle!) : $\sin 2\theta = \frac{m_D}{\sqrt{T^2 - 4D}}$ $\cos 2\theta = \frac{m_L - m_R}{\sqrt{T^2 - 4D}}$
- Diagonalizing rotation : $\begin{bmatrix} m_+ & 0 \\ 0 & m_- \end{bmatrix} = \begin{bmatrix} c_\theta & s_\theta \\ -s_\theta & c_\theta \end{bmatrix} \begin{bmatrix} m_L & m_D \\ m_D & m_R \end{bmatrix} \begin{bmatrix} c_\theta & -s_\theta \\ s_\theta & c_\theta \end{bmatrix}$
- Eigenvectors : $[v_1, v_2] \begin{bmatrix} m_L & m_D \\ m_D & m_R \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = [v_1', v_2'] \begin{bmatrix} m_+ & 0 \\ 0 & m_- \end{bmatrix} \begin{bmatrix} v_1' \\ v_2' \end{bmatrix}$, $\begin{bmatrix} v_1' \\ v_2' \end{bmatrix} = \begin{bmatrix} c_\theta & s_\theta \\ -s_\theta & c_\theta \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- If one mass eigenvalue < 0, then redefine field : $\psi \to \gamma_5 \psi$, so that $m \to -m$.
- Special case : $M = \begin{bmatrix} 0 & m \\ m & \Lambda \end{bmatrix}$ with $m \ll \Lambda$. Then:

Eigenvectors (ν fields) :	Eigenvalues (masses)
$\nu_{heavy} \simeq (\nu_R + \nu_R^c) + \frac{m}{\Lambda}(\nu_L + \nu_L^c)$	Λ
$\nu_{light} \simeq -(\nu_L + \nu_L^c) + \frac{m}{\Lambda}(\nu_R + \nu_R^c)$	$(-)\,\frac{m^2}{\Lambda} \ll m$

← "**see-saw** mechanism"

The existence of a singlet neutrino ν_R is predicted in many extensions of the EW Standard Model. E.g., in the $\underline{16}$ representation of $SO(10)$:

$$\begin{pmatrix} u_L & u_L & u_L & \nu_L \\ d_L & d_L & d_L & e_L \\ u_R & u_R & u_R & \boldsymbol{\nu_R} \\ d_R & d_R & d_R & e_R \end{pmatrix}$$

→ Can get a Majorana mass term $\sim \Lambda(\bar{\nu}_R^c \nu_R + \bar{\nu}_R \nu_R^c)$, in addition to the "standard" Dirac mass term $\sim m(\bar{\nu}_L \nu_R + \bar{\nu}_R \nu_L)$, where m is naturally associated to the EW scale, while Λ is naturally associated to the scale of new physics (Λ_{GUT} ?) which characterizes the SM extension.

→ Diagonalizing $\begin{bmatrix} 0 & m \\ m & \Lambda \end{bmatrix}$ as in the previous page, one gets a light Majorana neutrino with mass $\sim \frac{m^2}{\Lambda}$, and a very heavy Majorana neutrino with mass $\sim \Lambda$. The light one (in the presence of more than one generations) enters in neutrino oscillations. Can the heavy one be "useful" as well? Yes!

→

→ There is the possibility to have **leptogenesis**.

- CP violation at high energy scales Λ might be responsible for different decay rates of heavy ν_R into charged leptons in the early universe :

$$\Gamma(\nu_R \to \ell^+ + \ldots) \neq \Gamma(\nu_R \to \ell^- + \ldots)$$

- This would provide an unbalance of leptons / antileptons, possibly at the origin of the matter-antimatter asymmetry of the universe.

→ Discovery of Majorana neutrinos + CP violation in the ν sector would make the see-saw + leptogenesis mechanisms more plausible (although it would be very difficult, if not impossible, to prove them).

Dirac and Majorana mass terms (more families)

- Most general case:
  - **3 LH gauge doublets** $\nu_{\alpha L}$ ($\alpha = e\mu\tau$) "ACTIVE" in EW interactions
  - **N_s RH gauge singlets** ν_{sR} ($s = 1, \dots, N_s$) "STERILE" in EW interactions

  where N_s can be any number, at any mass scale (light or heavy).

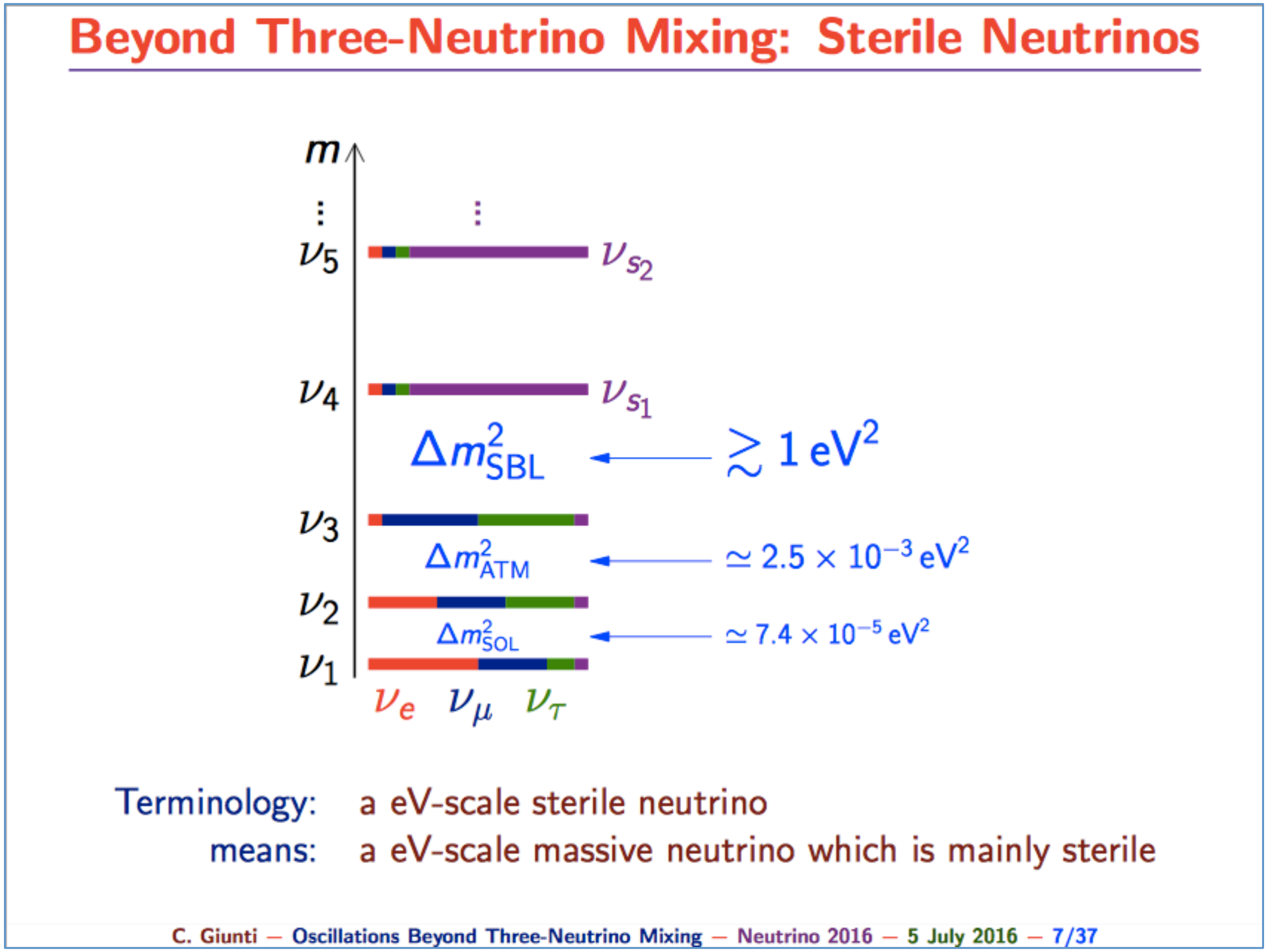
- Most general mass matrix has a block form,
  $M_L = 3\times 3$, $M_D = 3\times N_s$, $M_R = N_s \times N_s$

$$M = \left[\begin{array}{c|c} M_L & M_D \\ \hline M_D' & M_R \end{array}\right]$$

- After diagonalization, generic eigenvectors (ν fields) will be Majorana.
  → expect $0\nu\beta\beta$ decay allowed
  → expect Majorana phases besides CP̸ phase

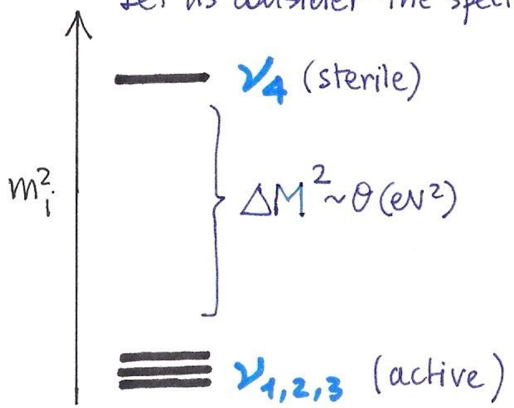
- After diagonalization, active and sterile neutrinos will be mixed
  → expect see-saw suppression of such mixing, but ... who knows?
  → U_{PMNS} not precisely unitary in general
  → active/sterile ν oscillations

Jargon for light (eV) sterile neutrinos...



Exercise : One-dominant-mass-scale oscillations for a 4th, sterile neutrino ν_s

Let us consider the spectrum:



Then, the PMNS 3x3 matrix is not unitary, being part of a larger 4x4 mixing matrix:

$$U = \begin{pmatrix} \left[\quad U_{PMNS} \quad \right] & \begin{matrix} U_{e4} \\ U_{\mu 4} \\ U_{\tau 4} \end{matrix} \\ U_{s1} \; U_{s2} \; U_{s3} & U_{s4} \end{pmatrix}, \quad \text{with } |U_{s4}|^2 \simeq 1 - \text{epsilon},$$

in order not to alter too much the established 3ν phenomenology $\rightarrow |U_{\alpha 4}|^2_{\alpha = e,\mu,\tau} \ll 1$

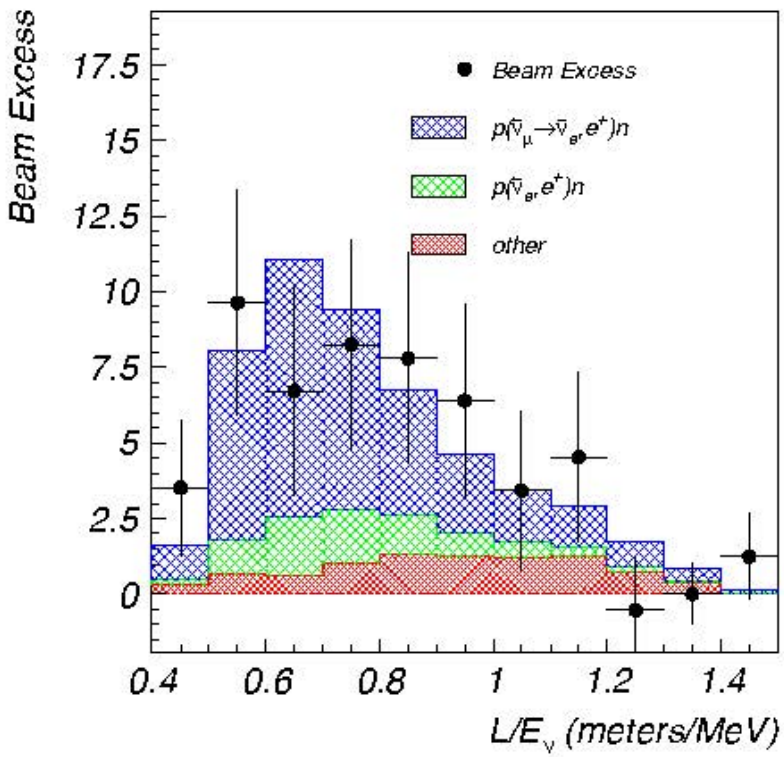
For experiments sensitive mainly to $\Delta M^2 \sim \mathcal{O}(\mathrm{eV}^2)$, one can take the limits $\delta m^2 \to 0$ and $\Delta m^2 \to 0$, and apply the same logic as for one dominant mass scale in $3\nu \cong (2\nu) \oplus (1\nu) \Rightarrow 4\nu \simeq (3\nu) \oplus (1\nu)$. Similarly, one gets (for $\alpha, \beta = e\mu\tau$)

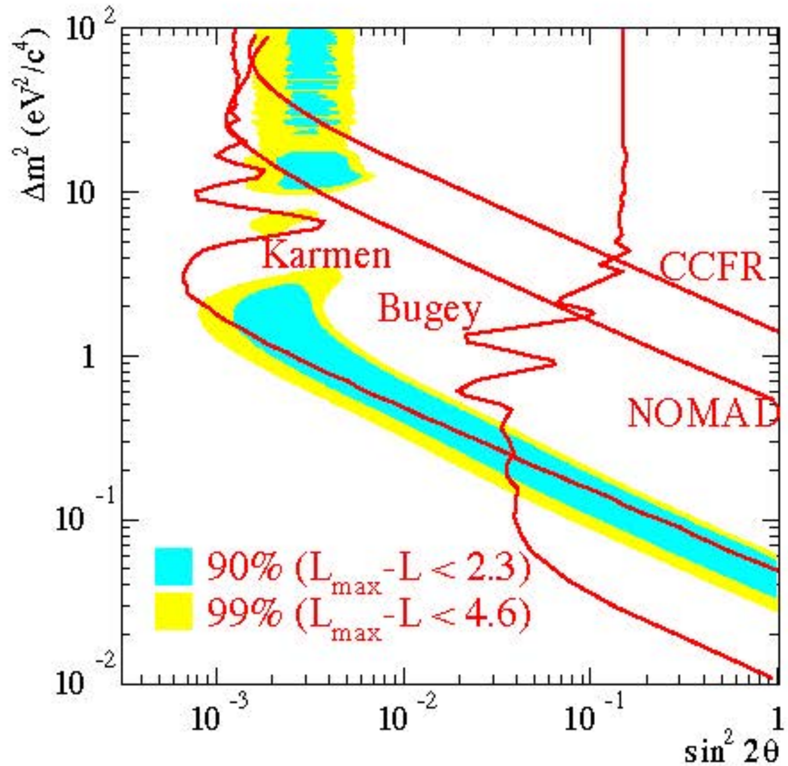
Disappearance $(\alpha = \beta)$: $1 - P_{\alpha\alpha} \simeq 4|U_{\alpha 4}|^2 (1 - |U_{\alpha 4}|^2) \sin^2\left(\frac{\Delta M^2 x}{4E}\right)$ ← "singly" suppressed by $|U_{\alpha 4}|^2 \ll 1$

Appearance $(\alpha \neq \beta)$: $P_{\alpha\beta} \simeq 4|U_{\alpha 4}|^2 |U_{\beta 4}|^2 \sin^2\left(\frac{\Delta M^2 x}{4E}\right)$ ← "doubly" suppressed by both $|U_{\alpha 4}| \ll 1$ and $|U_{\beta 4}| \ll 1$

One short-baseline accelerator experiment (LSND) claimed in the 90's flavor oscillation appearance at the O(eV) scale:

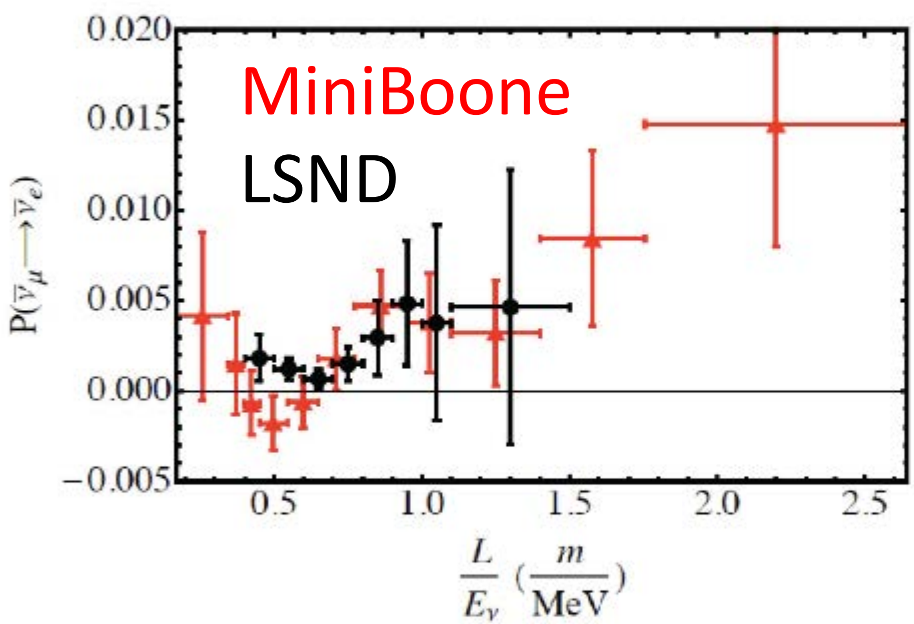
$\bar{\nu}_\mu \to \bar{\nu}_e$ $L \simeq 30$ m $20\,\text{MeV} \leq E \leq 200\,\text{MeV}$





Unfortunately, after >20 years, this result has not been either confirmed or ruled out conclusively, even by dedicated appearance experiments (e.g., MiniBoone).

You may or may not see an oscillation pattern here...



... especially if you exclude the two rightmost data points at lowest energy and highest background.

Moreover, if there is a new $\nu_\mu \rightarrow \nu_e$ appearance signal, there must be larger $\nu_\mu \rightarrow \nu_\mu$ and $\nu_e \rightarrow \nu_e$ disappearance signals at the same scale, $\Delta M^2 \sim O(eV^2)$:

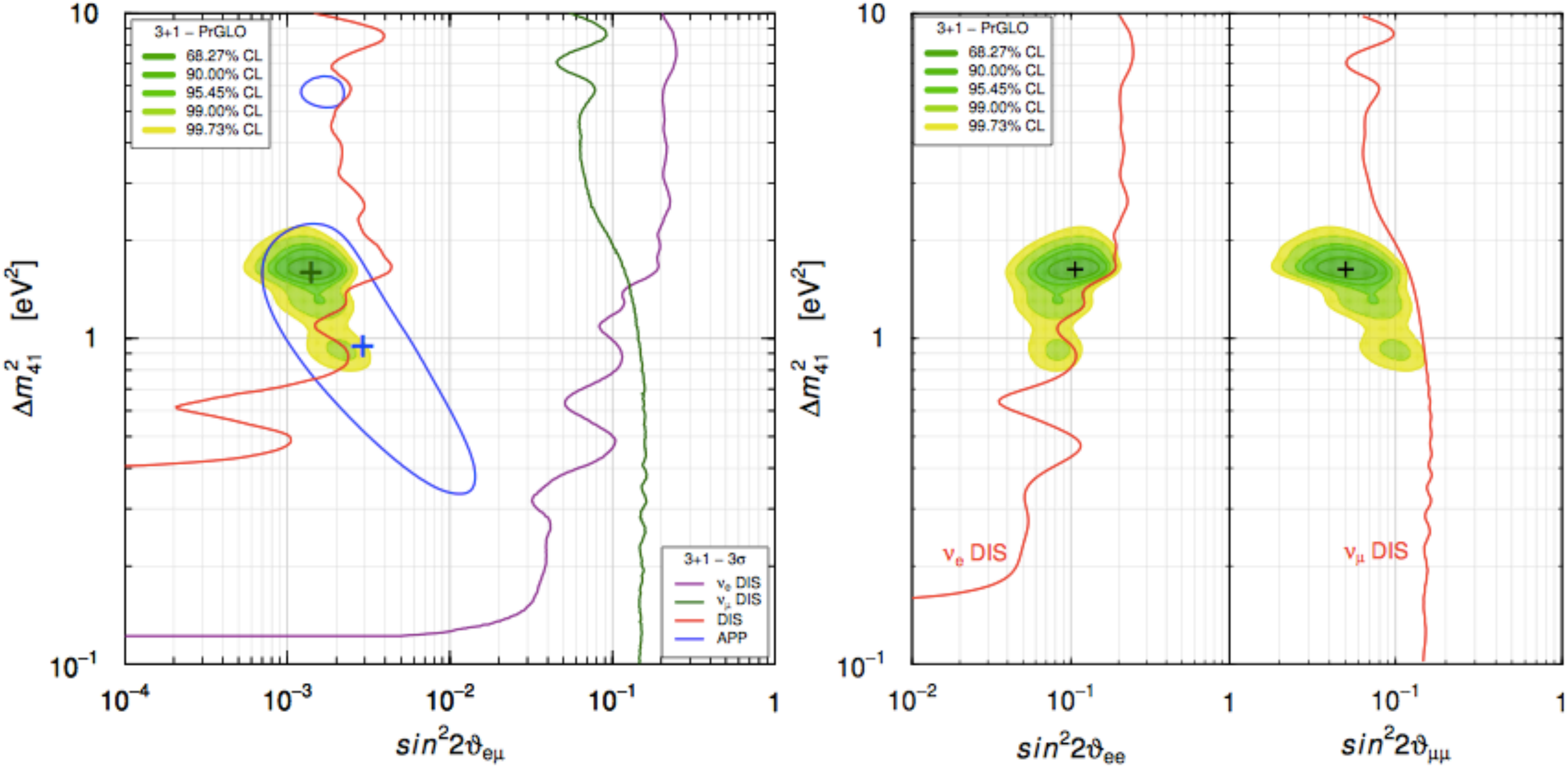
Disappearance ($\alpha=\beta$): $1-P_{\alpha\alpha} \simeq 4|U_{\alpha 4}|^2(1-|U_{\alpha 4}|^2)\sin^2\left(\frac{\Delta M^2 x}{4E}\right)$ ← "singly" suppressed by $|U_{\alpha 4}|^2 \ll 1$

Appearance ($\alpha\neq\beta$): $P_{\alpha\beta} \simeq 4|U_{\alpha 4}|^2|U_{\beta 4}|^2\sin^2\left(\frac{\Delta M^2 x}{4E}\right)$ ← "doubly" suppressed by both $|U_{\alpha 4}| \ll 1$ and $|U_{\beta 4}| \ll 1$

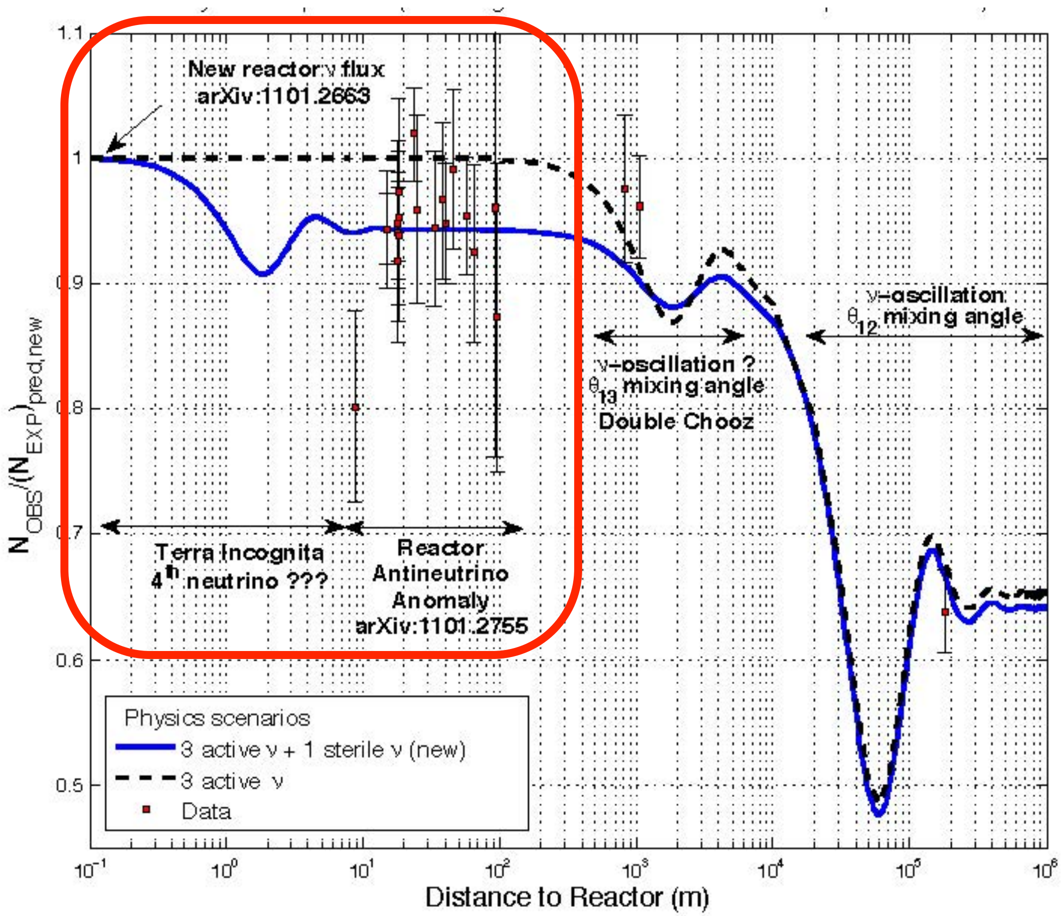
However, no unambiguous disappearance signal has been seen, especially in $\nu_\mu \rightarrow \nu_\mu$ mode.

This fact makes it difficult to reconcile positive LSND results with (mainly) negative results from other epts., not only in models with 1 additional sterile neutrino (3+1) but even with 2 steriles (3+2) →

Sterile neutrinos: Appearance vs Disappearance... [from Giunti+ 2015]

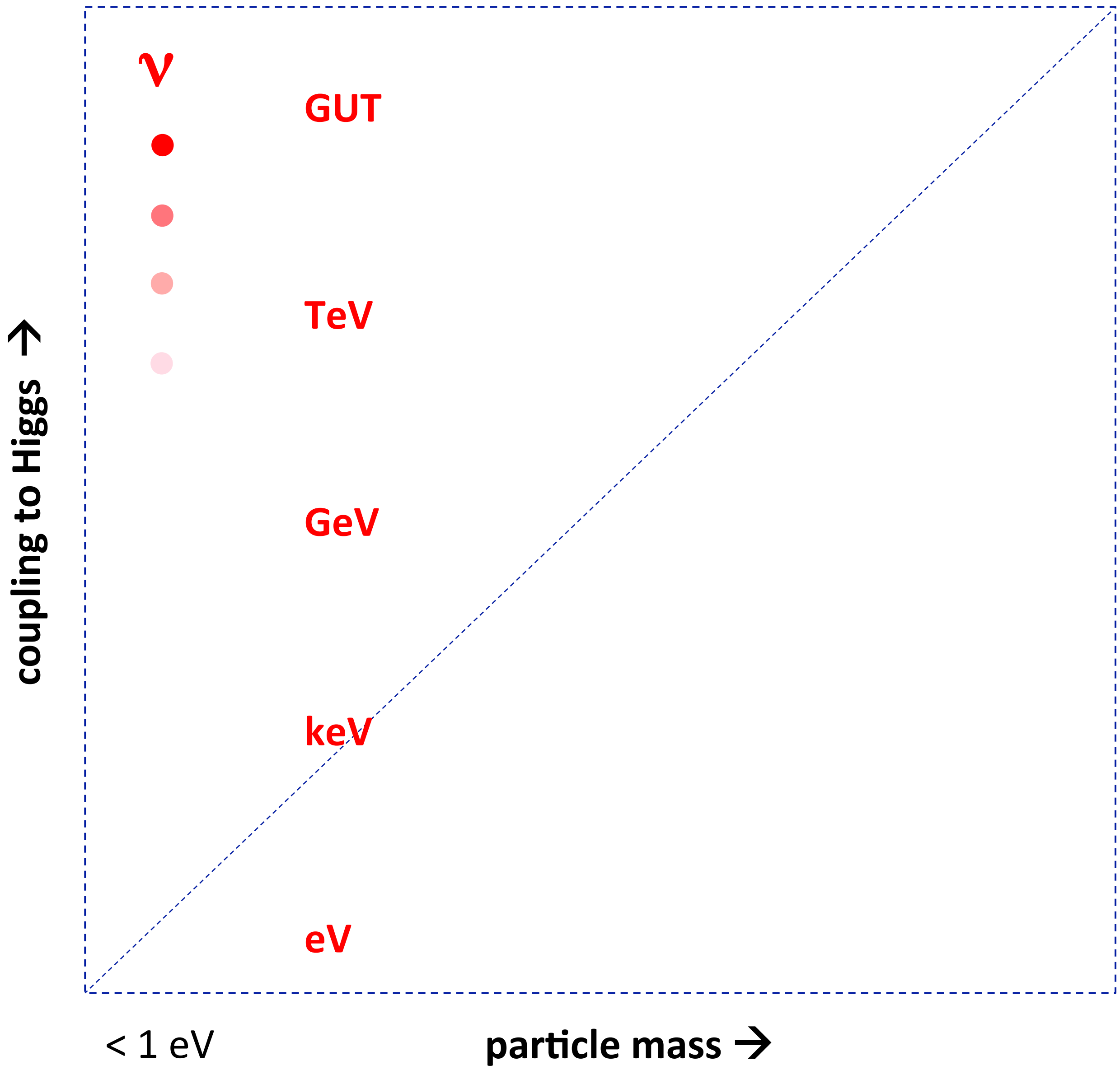


in 2011, a possible hint for $\nu_e \rightarrow \nu_e$ disappearance due to sterile n has been claimed ("reactor neutrino anomaly"), by a reanalysis of new fluxes and old reactor experiments at L<O(100) m. In addition to the known disappearance due to 3n oscillations at L>O(100 m), there seems to be an extra deficit at small L:

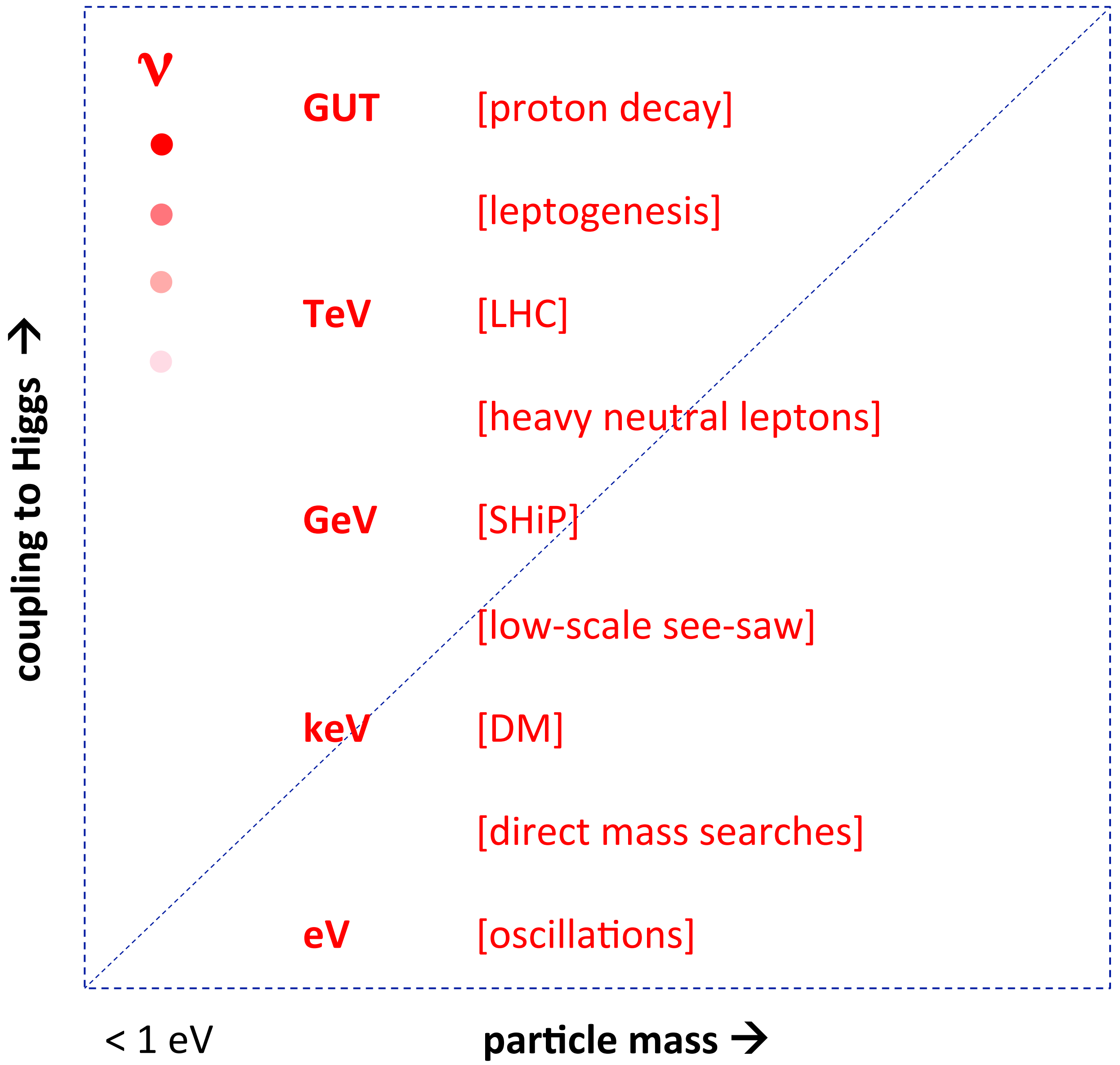


Many ongoing experiments to test this anomaly

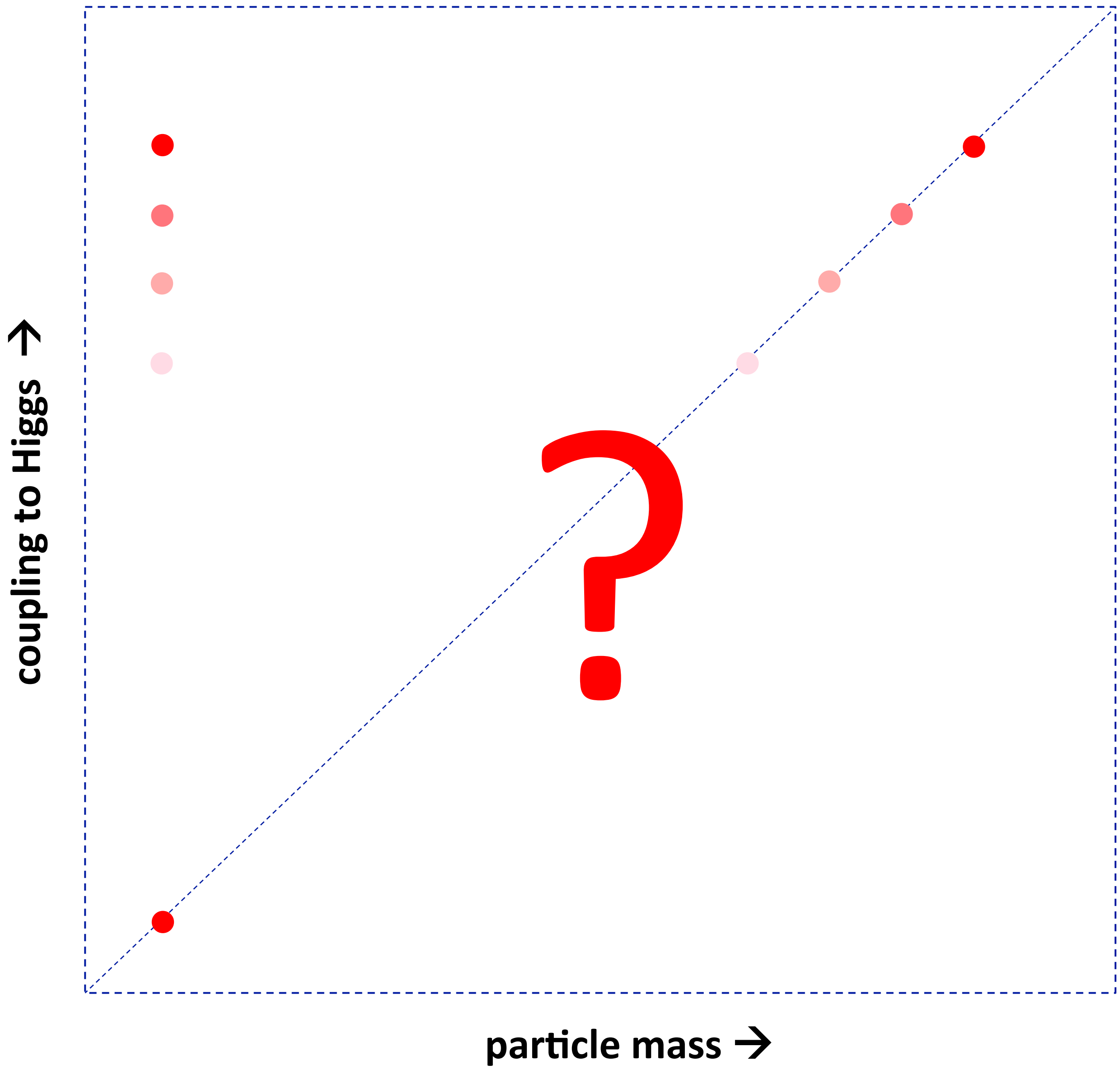
New mass states could actually emerge at (different) new scales ...



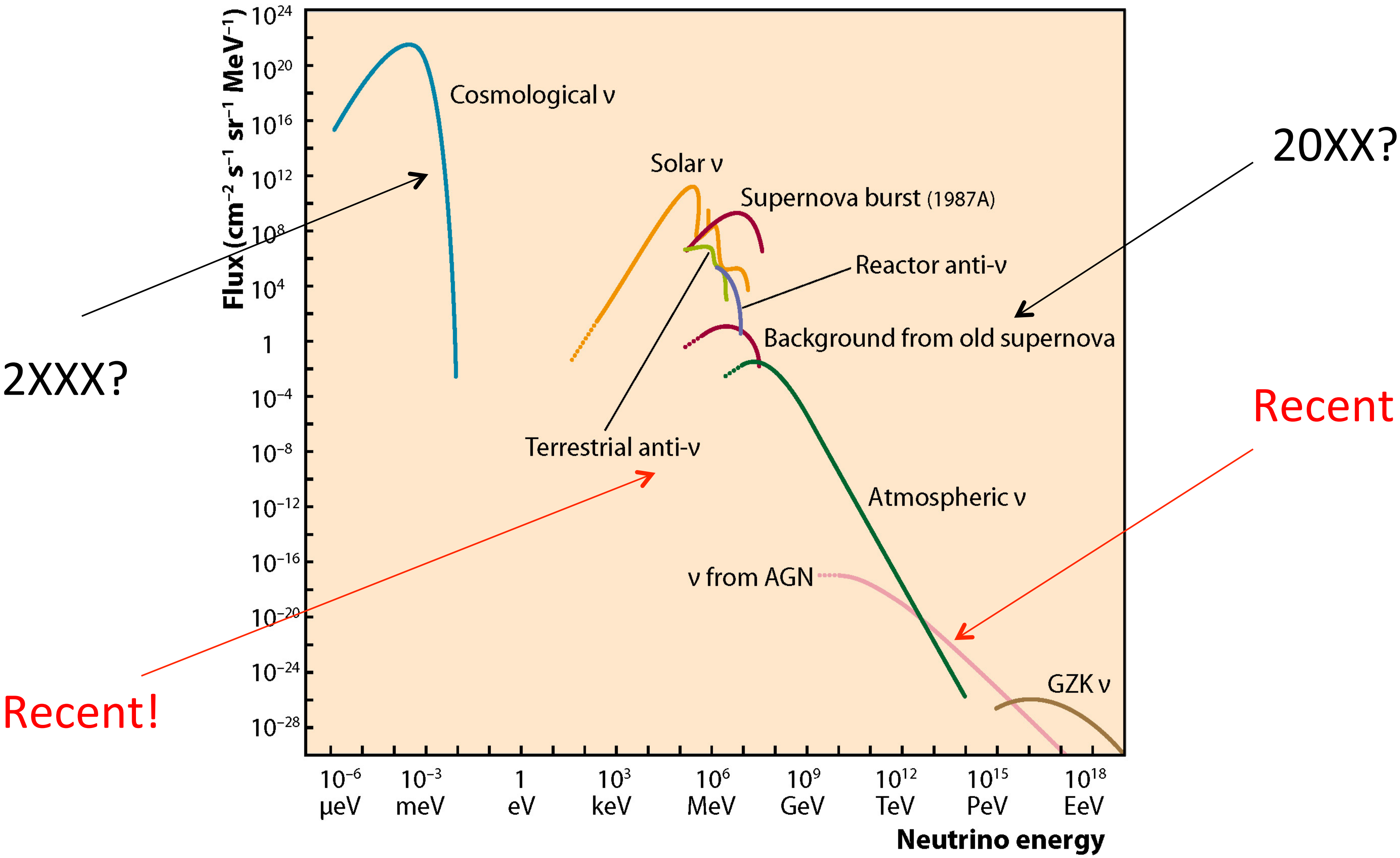
… and contribute to a rich phenomenology, e.g.,



Let us remain open-minded: new physics may emerge at any scale!



Finally, we should never forget that there are still vast lands to be explored in the neutrino world...



A synoptic view of neutrino fluxes. (from ASPERA roadmap)

Conclusions and Open Problems

Great progress in recent years …

Neutrino mass & mixing: established fact
Determination of $(\delta m^2, \theta_{12})$ and $(\Delta m^2, \theta_{23})$
Determination of θ_{13} at reactors (+ accel.)
Observation of (half)-period of oscillations
Direct evidence for solar ν flavor change
Evidence for matter effects in the Sun
ν_e, ν_τ appearance at accelerators
Upper bounds on ν masses in (sub)eV range
…………

Leptonic CP violation
Absolute m_ν from β-decay and cosmology
$0\nu\beta\beta$ decay and Dirac/Majorana ν
Matter effects in the Earth, Supernovae…
Normal vs inverted hierarchy
Octant of θ_{23}
Sterile neutrinos in oscillations and cosmology
New neutrino interactions
Deeper theoretical understanding
See-saw and leptogenesis scenarios
…………

… and great challenges for the future!

- E P I L O G U E -

An old Latin saying:

Nomen [est] Omen

"Name [is] Destiny"

Neutrino – What is the root of this name?

Language	Word tree	*…Some branches*	Meaning
Physics (Fermi 1934)	**NEUTR-INO**		**Little neutral one**
Italian	**NEUTRO**		**Neutral**
Latin	**NE-UTER**		**Not either; neutral**
Latin	**UTER**		**Either**
Greek	↑	*OUDETEROS*	*Neutral*
Old High German	↗	*HWEDAR*	*Which of two; whether*
Phonetic change/loss	[K]**UOTER**[US]		**Which of the two?**
Ionic Greek	**KOTEROS**		**Which of the two?**
Sanskrit	**KATARAS**		**Which of the two?**
Latin	↑	*QUANTUS*	*How much?*
Sanskrit		*KATAMAS*	*Which out of many?*
Sanskrit		*KATHA*	*How?*
Sanskrit	↗	*KAS*	*Who?*
Indo-European root	**KA** *or* **KWA**		**Interrogative base**

The root of the name [neutrino] ... is a [kwa]stion

Language	Word tree	...*Some branches*	Meaning
Physics (Fermi 1934)	NEUTR-INO		Little neutral one
Italian	NEUTRO		Neutral
Latin	NE-UTER		Not either; neutral
Latin	UTER		Either
Greek	↑	*OUDETEROS*	*Neutral*
Old High German		↗ *HWEDAR*	*Which of two; whether*
Phonetic change/loss	[K]UOTER[US]		Which of the two?
Ionic Greek	KOTEROS		Which of the two?
Sanskrit	KATARAS		Which of the two?
Latin	↑	*QUANTUS*	*How much?*
Sanskrit		*KATAMAS*	*Which out of many?*
Sanskrit		*KATHA*	*How?*
Sanskrit		↗ *KAS*	*Who?*
Indo-European root	KA *or* KWA		Interrogative base

If “name is destiny,” then …

Neutrino’s destiny
is to raise questions!

Thank you for your attention

Additional slides

APPENDIX

Elements of statistics and data analysis

Purpose: become familiar with the treatment of correlated uncertainties, which are often a key ingredient of (neutrino) data analyses.

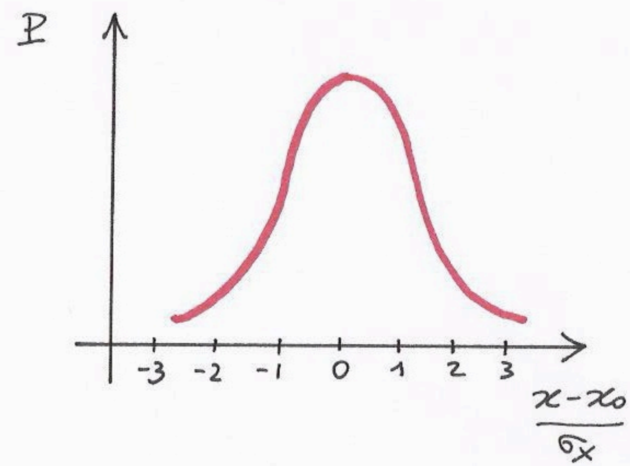
Error distribution for 1 variable

Here we consider only gaussian errors. Also Poisson fluctuations ($\propto \sqrt{N}$) are assumed to be nearly gaussian. [Some comments on small-N cases and on asymmetric errors will be presented at the end.]

Distribution for a single variable:

$$P(x, x_0) = \frac{1}{\sqrt{2\pi} \cdot \sigma_x} \, e^{-\frac{1}{2}\left(\frac{x - x_0}{\sigma_x}\right)^2} ,$$

corresponding to quote $x = x_0 \pm \sigma_x$



Area within: $\pm 1\sigma = 68.27\,\%$

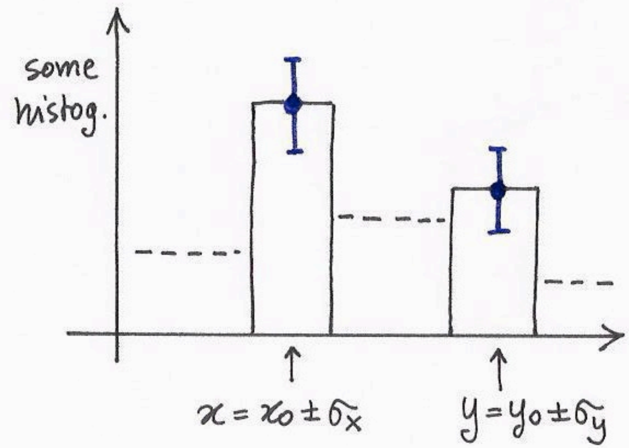
$\pm 2\sigma = 95.45\%$

$\pm 3\sigma = 99.73\%$

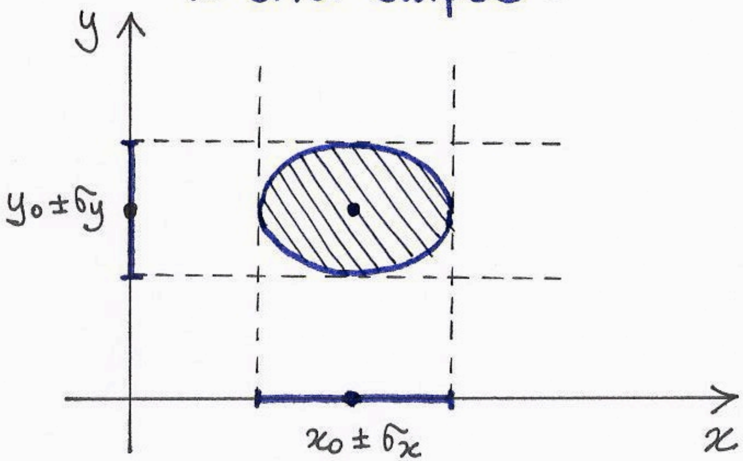
Error distribution for 2 variables (uncorrelated)

Let us consider 2 quantities x & y with errors which have no relation with each other (e.g. statistical errors of two bins):

$$P(x,y;x_0,y_0) = P(x,x_0)\,P(y,y_0)$$



1σ error ellipse:



Ellipse equation:

$$\Delta\chi^2 = 1$$

where $\Delta\chi^2 = \left(\frac{x-x_0}{\sigma_x}\right)^2 + \left(\frac{y-y_0}{\sigma_y}\right)^2$

$[\Delta\chi^2 = 0$ at $(x,y) = (x_0,y_0)]$

Note: The probability of finding (x,y) within the 1σ error ellipse is not 68.27% : it is 39.35% !

68% = probability of finding x in $x_0 \pm \sigma_x$, independently on y

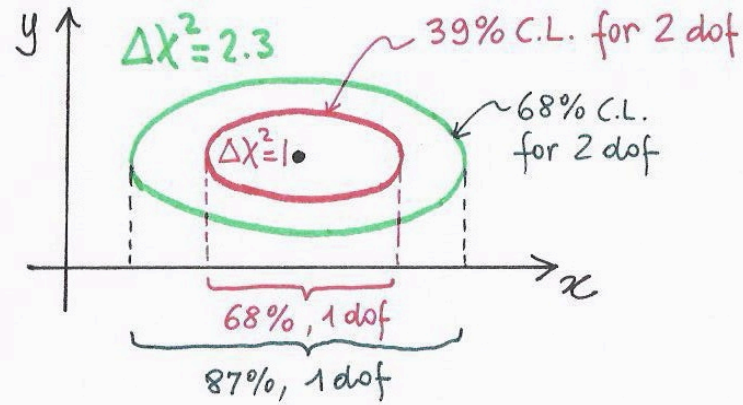
68% = " " " y in $y_0 \pm \sigma_y$, " " x

39% = joint probability of finding (x,y) within the 1σ ellipse

---

If you really want an error ellipse containing 68% joint probab. (68% C.L. for 2 d.o.f.), then you should use $\Delta\chi^2 = 2.3$. Its projections define 87% C.L. for each variable (1 dof).

This is not usually called a "1σ" ellipse.



Confusion may arise if a C.L. is quoted without the corresponding # d.o.f.

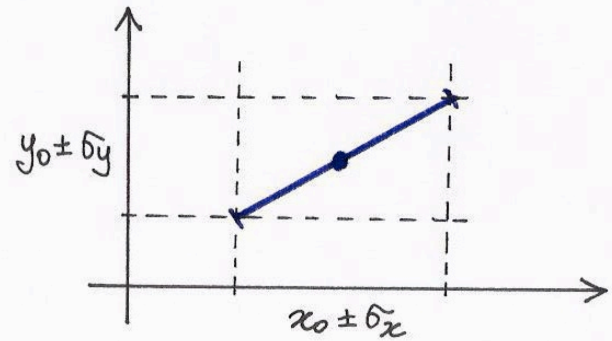
Error distribution for 2 variables (fully correlated)

Consider two variables x and y with errors in one-to-one correspondence, e.g., two bins affected by a common normalization error:

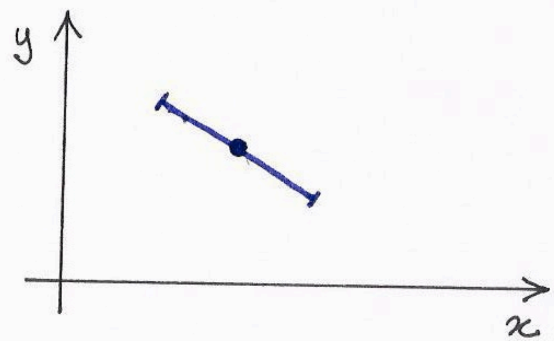
Then the errors go both "up" or "down":

if $x = x_0 + \sigma_x$ then $y = y_0 + \sigma_y$
if $x = x_0 - \sigma_y$ then $y = y_0 - \sigma_y$

The error ellipse is degenerate →
(fully correlated errors)



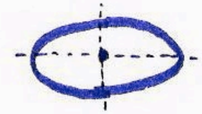
Analogously for full "anticorrelation":
E.g., two bins whose sum is constant;
then, $x = x_0 + \sigma_x$ implies $y = y_0 - \sigma_y$,
and the degenerate ellipse has a negative slope →



Recap: Eqs. for 1σ error ellipses (limit cases)

- No correlation

$$1 = (x-x_0, y-y_0)\underbrace{\begin{pmatrix}\sigma_x^2 & 0 \\ 0 & \sigma_y^2\end{pmatrix}^{-1}}_{\det \neq 0}\begin{pmatrix}x-x_0 \\ y-y_0\end{pmatrix}$$



- Full correlation

$$1 = (x-x_0, y-y_0)\underbrace{\begin{pmatrix}\sigma_x^2 & \sigma_x\sigma_y \\ \sigma_x\sigma_y & \sigma_y^2\end{pmatrix}^{-1}}_{\det = 0}\begin{pmatrix}x-x_0 \\ y-y_0\end{pmatrix}$$

- Full anticorr.

$$1 = (x-x_0, y-y_0)\underbrace{\begin{pmatrix}\sigma_x^2 & -\sigma_x\sigma_y \\ -\sigma_x\sigma_y & \sigma_y^2\end{pmatrix}}_{\det = 0}\begin{pmatrix}x-x_0 \\ y-y_0\end{pmatrix}$$

In general we expect: $1 = (x-x_0, y-y_0)\begin{pmatrix}\text{squared} \\ \text{error} \\ \text{matrix}\end{pmatrix}^{-1}\begin{pmatrix}x-x_0 \\ y-y_0\end{pmatrix}$

More general 1σ error ellipses

Let us consider two variables (x, y) and two sources of uncertainties:

- statistical (s_x, s_y) with no correlation,
- systematic (c_x, c_y) with full correlation,

namely, $\begin{cases} x = x_0 \pm s_x \,(\text{stat}) \pm c_x \,(\text{syst}) \\ y = y_0 \pm s_y \,(\text{stat}) \pm c_y \,(\text{syst}) \end{cases}$

The errors sum up in quadrature <u>at matrix level</u>:

$$\sigma^2 = \begin{bmatrix} s_x^2 & 0 \\ 0 & s_y^2 \end{bmatrix} + \begin{bmatrix} c_x^2 & c_x c_y \\ c_x c_y & c_y^2 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$$

← squared error matrix

where

$$\sigma_x^2 = s_x^2 + c_x^2$$

$$\sigma_y^2 = s_y^2 + c_y^2$$

$$\rho = \frac{c_x c_y}{\sigma_x \sigma_y} = \begin{cases} 0 & \text{for } c_x \text{ or } c_y = 0 \quad \text{(no correlation)} \\ 1 & \text{for } s_x = s_y = 0 \quad \text{(full correlation)} \end{cases}$$

In general, the correlation ρ obeys: $0 \leq |\rho| \leq 1$

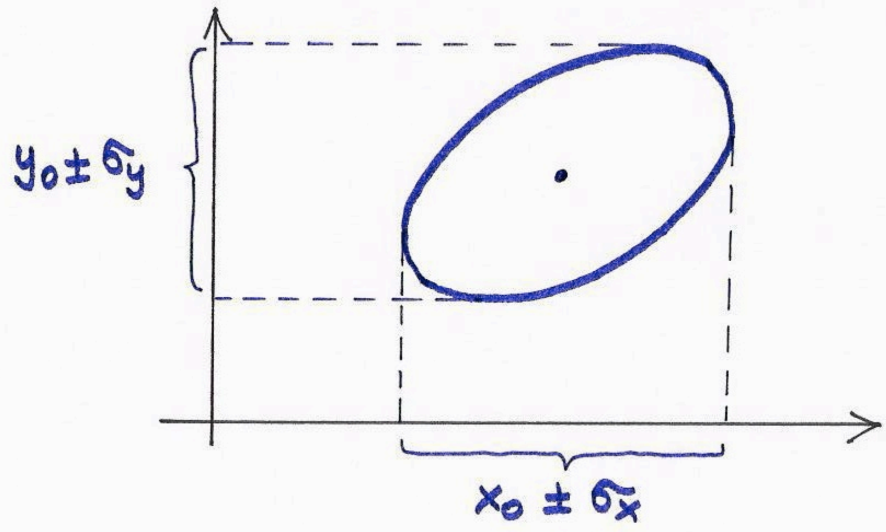
Qualitative shape of 1σ error ellipse:

(s_x, s_y)	(c_x, c_y)	(σ_x, σ_y)
$\rho = 0$	$\rho = +1$	$0 < \rho < 1$

Analogously, adding a fully anticorrelated error source:

(s_x, s_y)	(c_x, c_y)	(σ_x, σ_y)
$\rho = 0$	$\rho = -1$	$-1 < \rho < 0$

Projections of 1σ error ellipse coincide with $\pm 1\sigma$ ranges for the x and y variables:



Equation of 1σ ellipse:

$$1 = \Delta\chi^2$$

$$= (x-x_0,\ y-y_0)\begin{pmatrix}\sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2\end{pmatrix}^{-1}\begin{pmatrix}x-x_0 \\ y-y_0\end{pmatrix}$$

$$= \frac{1}{1-\rho^2}\left[\left(\frac{x-x_0}{\sigma_x}\right)^2 + \left(\frac{y-y_0}{\sigma_y}\right)^2 - 2\rho\,\frac{(x-x_0)(y-y_0)}{\sigma_x\sigma_y}\right]$$

Probability distribution: $$P = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\,e^{-\frac{1}{2}\Delta\chi^2}$$

Positively correlated variables and their ratio

Let us consider two variables with positively correlated errors : $\begin{cases} x = x_0 \pm \sigma_x \\ y = y_0 \pm \sigma_y \end{cases}$ $\quad \rho > 0$

If the correlation is sizable, we expect a significant "cancellation" of errors in the ratio : $r = \frac{x}{y} = r_0 \pm \sigma_r$ with "small σ_r"

The error of the ratio can be evaluated as :

$$\sigma_r^2 = \left(\frac{\partial r}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial r}{\partial y}\right)^2 \sigma_y^2 + 2\rho\left(\frac{\partial r}{\partial x}\right)\left(\frac{\partial r}{\partial y}\right)\sigma_x\sigma_y$$

$$\rightarrow \frac{\sigma_r^2}{r_0^2} = \frac{\sigma_x^2}{x_0^2} + \frac{\sigma_y^2}{y_0^2} - 2\rho\frac{\sigma_x}{x_0}\frac{\sigma_y}{y_0} \; ;$$

the "$-2\rho\ldots$" term is responsible for error cancellation.

Note that, for $\rho = +1$ (full correlation) : $\frac{\sigma_r^2}{r_0^2} = \left(\frac{\sigma_x}{x_0} - \frac{\sigma_y}{y_0}\right)^2$;
then, if $\frac{\sigma_x}{x_0} = \frac{\sigma_y}{y_0}$ (e.g., for a common normalization uncertainty) the cancellation is complete ($\sigma_r = 0$) as expected.

In general, it is preferable to use correlated variables (x,y) whenever possible, rather than their ratio $r = x/y$.

The main reason is that, if y is distributed as a Gaussian, then $1/y$ is distributed as a Lorentzian ("Breit-Wigner"), with a formally infinite variance. In practice, this may be problematic if σ_y is large and/or one is probing the distribution tails.

Therefore, if we measure $x = x_0 \pm \sigma_x$ and $y = y_0 \pm \sigma_y$, and if we know that $r = r_0 \pm \sigma_r$, it is convenient to keep (x,y) in the fit, together with the correlation

$$\rho = \left(\frac{\sigma_x^2}{x_0^2} + \frac{\sigma_y^2}{y_0^2} - \frac{\sigma_r^2}{r_0^2} \right) \Big/ \left(2 \frac{\sigma_x}{x_0} \frac{\sigma_y}{y_0} \right)$$

(instead of using r directly).

Historically, several "ratios" have been progressively abandoned in neutrino data fits.

Estimates of correlations and covariances

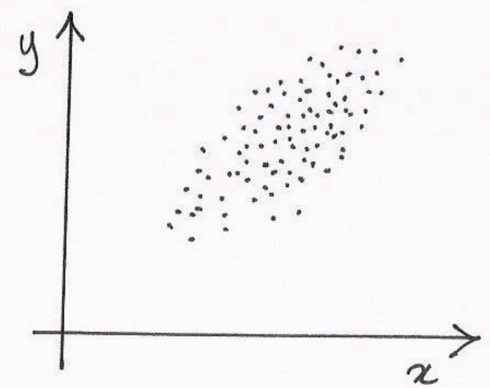
The squared error matrix $\sigma^2 = \begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{bmatrix}$

is also called "covariance matrix":

$\sigma_{xy}^2 = \varrho\, \sigma_x \sigma_y$ = covariance of (x,y)

σ_x^2 = variance of x

σ_y^2 = variance of y



If repeated measurements (or simulations) of the variables x and y are available, →

then:

$$x_0 = \frac{1}{n}\sum_{i=1}^{n} x_i \quad , \quad \sigma_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - x_0)^2$$

$$y_0 = \frac{1}{n}\sum_{i=1}^{n} y_i \quad , \quad \sigma_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - y_0)^2$$

$$\sigma_{xy}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - x_0)(y_i - y_0) \; ; \quad \varrho = \sigma_{xy}^2 / \sigma_x \sigma_y$$

→ Get: $\begin{cases} x = x_0 \pm \sigma_x \\ y = y_0 \pm \sigma_y \\ \varrho \end{cases}$

However, the ideal situation is to identify and break down all possible error sources to the two main categories of "uncorrelated" errors s and "fully correlated" errors c:

$$\sigma^2 = \begin{pmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{pmatrix} = \underbrace{\begin{pmatrix} s_x^2 & 0 \\ 0 & s_y^2 \end{pmatrix} + \cdots}_{\text{uncorrelated}} \quad + \underbrace{\begin{pmatrix} c_x^2 & c_x c_y \\ c_x c_y & c_y^2 \end{pmatrix} + \cdots}_{\text{fully correlated}}$$

[fully anticorrelated errors become fully correlated by changing the sign of one variable].

In this case, one is really "summing in quadrature" all possible, independent (known) error sources.

Generalization to N variables $\{x_i\}_{1 \le i \le N}$

- $\Delta\chi^2 = 1$ error ellipsoid in N-dimensional space is defined by:

$$1 = \Delta x^T (\sigma^2)^{-1} \Delta x$$

where Δx is a column vector, $\Delta x = \begin{pmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \\ \vdots \\ x_N - x_N^0 \end{pmatrix}$

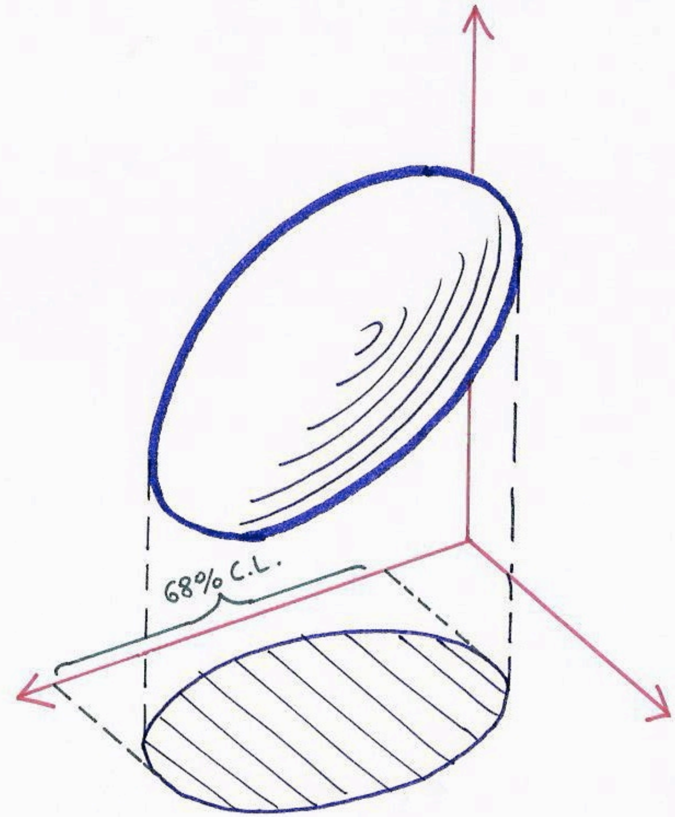
and $(\sigma^2)^{-1}$ is the inverse of the covariance matrix (symmetrical):

$$\sigma^2 = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \cdots \\ & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \cdots \\ & & \sigma_3^2 & \cdots \\ & & & \ddots \end{pmatrix}$$

- Probability distribution (multivariate Gaussian):

$$P = \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{\det \sigma^2}} \, e^{-\frac{1}{2}\Delta\chi^2}$$

- Projection of the $\Delta\chi^2 = 1$ ellipsoid onto one axis x_i gives the 1σ (68% CL) range on the x_i variable ($x_i = x_i^0 \pm \sigma_i$). This holds for any N. Variables $x_j \neq x_i$ are said to be "marginalized" or "projected away"



- However, the joint probability of $(x_1, x_2, \dots, x_N)$ being inside the ellipsoid decreases with N:

N = 1	68%	
N = 2	39%	$< (68\%)^2$
N = 3	20%	$< (68\%)^3$
N = 4	9%	$< (68\%)^4$
⋮	⋮	⋮

- Note : n-σ ellipsoids defined by $\Delta\chi^2 = n^2$

Fitting data with a model

In this case we compare theoretical predictions x_i^{theo} with N experimental data x_i^{exp}. Predictions will depend on N_p ($< N$) parameters $\vec{p}$: $x_i^{theo} = x_i^{theo}(\vec{p})$. Recipe:

- Build $\chi^2 = \Delta x^T (\sigma^2)^{-1} \Delta x$, Δx = column vector of $(x_i^{theo} - x_i^{exp})$
  $\sigma^2 = \sigma^2_{theo} + \sigma^2_{exp}$

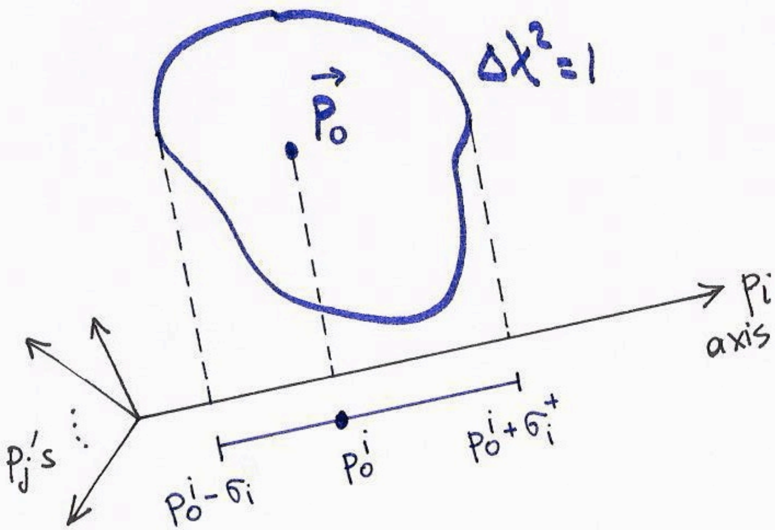
- Find $\chi^2_{min} = \min_{\vec{p}} \chi^2(\vec{p})$ at some $\vec{p} = \vec{p}_0$

- Check that $\chi^2_{min} \sim \underbrace{N - N_p}_{\text{dof for test of hypothesis}} \pm \sqrt{2(N-N_p)}$ ← see, e.g., PDG

- Check not ok : model wrong (χ^2_{min} too high) or "too good" (χ^2_{min} too low). Verify model, underestimated errors....

- Check OK →

Parameter estimation



The N_p-dimensional manifold defined by:

$$\Delta\chi^2 = \chi^2(\vec{p}) - \chi^2_{min} = 1$$

represents the "1σ allowed region" of parameters. Projection onto one axis p_i provides $\pm 1\sigma$ ranges:

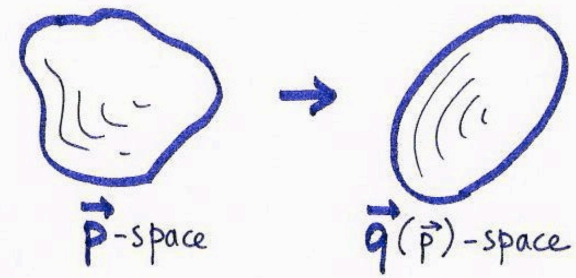
$$p = p_0^i \begin{smallmatrix} + \sigma_i^+ \\ - \sigma_i^- \end{smallmatrix} \quad \text{(generally asymmetric)}$$

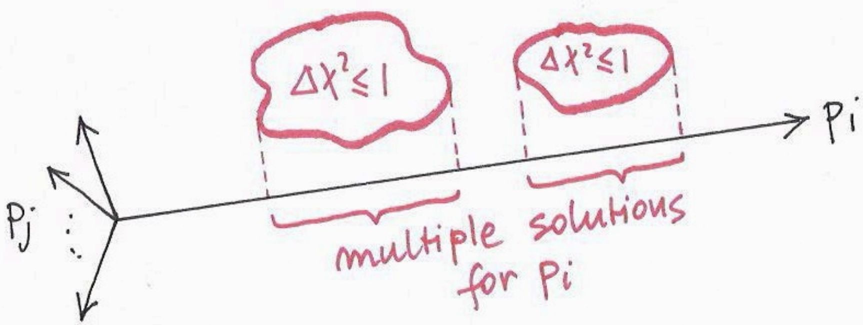
Projection onto p_i ($\equiv$ marginalization over $p_j \neq p_i$) in practice means setting $\Delta\chi_i^2 = 1$, where the "reduced" $\Delta\chi_i^2$ is:

$$\Delta\chi_i^2 = \min_{p_j \neq p_i} \left(\chi^2(\vec{p}) - \chi^2_{min}\right)$$

Analogously, n-σ ranges are defined by $\Delta\chi_i^2 = n^2$

- This procedure can be justified, as far as the allowed region is simply connected. The basic idea is that, through mapping, this volume can be transformed into an ellipsoid (→ gaussian machinery ...)





- However, if there are several (disconnected) allowed region, the statistical interpretation is problematic. No "consensus" approach in this case.

- In practice, most people keep using the $\Delta\chi^2 = \text{const}$ recipe also for multiple solutions, with some cautionary remarks.

- There is no other way than waiting for new experiments/data to solve the ambiguity ("degeneracy of solutions").

- Sometimes, one is interested not only in $\pm n\sigma$ limits on each variable separately ($\equiv \Delta\chi^2 = n^2$ projections), but on the joint probability of $\vec{p}$ being in a volume defined by $\Delta\chi^2 =$ const.

Relevant tables of $\Delta\chi^2$ level cuts (from PDG):

C.L.	N=1	N=2	N=3
68	1.00	2.30	3.53
90	2.71	4.61	6.25
95	3.84	5.99	7.82
99	6.63	9.21	11.34
99.73	9.00	11.83	14.16

E.g. the joint 95% C.L. region for two variables (p_i, p_j) is defined by $\Delta\chi^2_{ij} = 5.99$, where $\Delta\chi^2_{ij} = \min\limits_{p_k \neq p_{i,j}} [\chi^2(\vec{p}) - \chi^2_{min}]$.

Analyzing χ^2 contributions.

- The χ^2 is a global quantity. By itself, it does not necessarily "detect" single observables which might be badly fitted
  → need to split the total χ^2 into "pieces"

- One possibility is to look at the "residuals" or "pulls" after the fit:

$$(\text{pull})_i = \frac{x_i^{theo}(\vec{p}_0) - x_i^{exp}}{\sigma_i} \qquad \left(\sigma_i = \sqrt{\sigma_{ii}^2}\right)$$

- Then, a large pull (say, $\gtrsim 3\sigma$) signals a potential problem in the data and/or in the model.

The Standard Model fit to LEP data is often shown in terms of such pulls.

- With the previous definitions, however,

$$\chi^2 \neq \Sigma (\text{pull})^2$$

since σ^2_{ij} is not diagonal in general.

- It is possible to re-write the same χ^2 in a form $\Sigma (\text{pull})^2$ with a somewhat different approach, which also brings some technical advantages.

Previous χ^2 approach : "covariance method"
Alternative " : "pull method"

Covariance method (recap)

- Consider N observables $\{R_n\}_{n=1\dots N}$

  $\{R_n^{theo}\}$ = theoretical predictions

  $\{R_n^{exp}\}$ = experimental measurements

$$(R_n^{theo} - R_n^{exp}) \pm u_n \pm c_n^1 \pm c_n^2 \pm \dots \pm c_n^K$$

u_n: uncorrel. error

$c_n^1 \dots c_n^K$: set of K systematics produced by independ. sources

with $\rho(u_n, u_m) = 0$ (always uncorrelated)

$\rho(c_n^k, c_m^k) = 1$ (fully correlated for the same k-th source)

$\rho(c_n^k, c_m^h) = 0$ ($h \neq k$, uncorrelated from different sources)

- Then: Build $\sigma_{nm}^2 = \delta_{nm} u_n u_m + \sum_{k=1}^{K} c_n^k c_m^k$

  and evaluate $\chi^2_{cov} = \sum_{n,m=1}^{N} (R_n^{exp} - R_n^{theo}) \left[\sigma_{nm}^2\right]^{-1} (R_m^{exp} - R_m^{theo})$

  as discussed previously

Pull method

- Shift the theoretical predictions linearly in the systematics:

$$R_n^{theor} \rightarrow R_n^{theo} + \sum_{k=1}^{K} \xi_k C_n^k$$

where ξ_k = univariate gaussian random variable $\left(\langle \xi_k \rangle = 0,\ \langle \xi_k^2 \rangle = 1 \right)$

- Minimize over ξ_k the following sum of squared residuals:

$$\chi^2_{pull} = \min_{\{\xi_k\}} \left[\sum_{n=1}^{N} \left(\frac{R_n^{exp} - \left(R_n^{theo} + \sum_{k=1}^{K} \xi_k C_n^k \right)}{u_n} \right)^2 + \sum_{k=1}^{K} \xi_k^2 \right]$$

↑ squared residuals ↑ penalty term

- At minimum $\left(\xi_k \stackrel{def}{=} \bar{\xi}_k \right)$: $\bar{R}_n^{theo} = R_n^{theo} + \sum_{k=1}^{K} \bar{\xi}_k C_n^k$ ← "shifted" predictions

and $$\chi^2_{pull} = \sum_{n=1}^{N} \left(\frac{R_n^{exp} - \bar{R}_n^{theo}}{u_n} \right)^2 + \sum_{k=1}^{K} \bar{\xi}_k^2$$

$$= \sum_{n=1}^{N} \left(\begin{matrix} \text{pull of} \\ \text{observable} \end{matrix} \right)_n^2 + \sum_{k=1}^{K} \left(\begin{matrix} \text{pull of} \\ \text{systematic} \end{matrix} \right)_k^2$$ ← "diagonal" form

- It turns out that : $\chi^2_{pull} \equiv \chi^2_{covariance}$ (some algebra needed)

  so the methods are numerically equivalent.

- The pull approach may be more convenient for large N. The inversion of a large $N \times N$ covariance matrix may be unstable, especially if systematics dominate. The pull method leads to k equations (linear) in the ξ_k's, which is solved by a $k \times k$ matrix inversion, with $k \ll N$ usually.

- In addition, relatively large ξ_k's may signal systematic "offsets" required to match data and theory.

- Several ν data analyses are now performed in terms of χ^2_{pull}.

See hep-ph/0206162 for details

Comment on low-statistic bins

- The fit to a histogram may become problematic if one (or more) bin contains a low number of events N^{exp} (or even none, $N^{exp}=0$).

  In this case, the Gaussian approximation

  $$\chi^2 \ni \frac{(N^{exp} - N^{theo})^2}{N^{exp}} \qquad \underline{\text{fails}}$$

- In this case, the PDG suggests an alternative form, which embeds more properly the Poisson nature of the fluctuations:

  $$\chi^2 \ni 2\left(N^{theo} - N^{exp} + N^{exp} \ln \frac{N^{exp}}{N^{theo}}\right)$$

  $$\left[\text{or } 2N^{theo} \text{ if } N^{exp}=0\right]$$

- Additional systematics can still be embedded in the pull approach via:

  $$N^{theo} \to N^{theo} + \sum_k \xi_k c_n^k$$

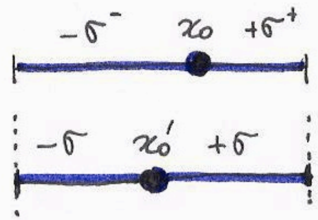
- Gaussian limit recovered at large N.
  Hint: expand logarithm at second order in $(N_{theo} - N_{exp})/N_{exp}$.

Comment on asymmetric errors

- If a variable x is affected by asymmetric errors: $x = x_0{}^{+\sigma^+}_{-\sigma^-}$
  there is no "consensus recipe" to write a χ^2 contribution.

- Sometimes the range is conservatively symmetrized to the largest error: $x = x_0 \pm \sigma_{max}$, $\sigma_{max} = \max(\sigma^+, \sigma^-)$

- A better recipe has been argued in physics/0403086 :
  shift $x_0 \to x_0'$ so that the new $\pm 1\sigma$ range reproduces the old one.



$$x_0' + \sigma \equiv x_0 + \sigma^+$$
$$x_0' - \sigma \equiv x_0 - \sigma^-$$

Then the χ^2 contribution is : $\left(\frac{x - x_0'}{\sigma}\right)^2$.