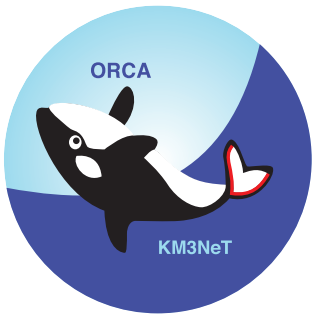
Comparison of ORCA and PINGU Sensitivity Studies

Lukas Schulte[†], Martijn Jongen[⋆]

[†] Physikalisches Institut, Universität Bonn
[⋆] Nikhef Amsterdam

MANTS meeting Geneva
September 2014



Introduction

- Cooperation between experiments
- Hierarchy significances are hard to interpret
- Different inputs and assumptions used
- Need for consistent results

Table of Contents

Overview of Sensitivity Studies

Comparison of Ingredients and Assumptions

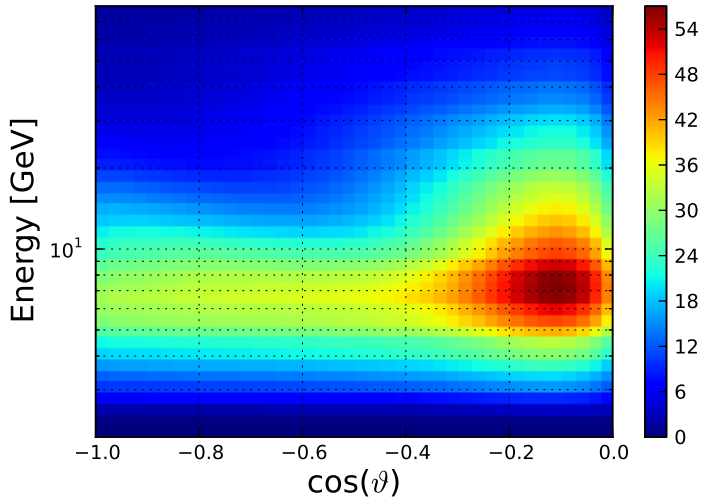
Ongoing Cooperation

Recap and Conclusions

Backup Slides

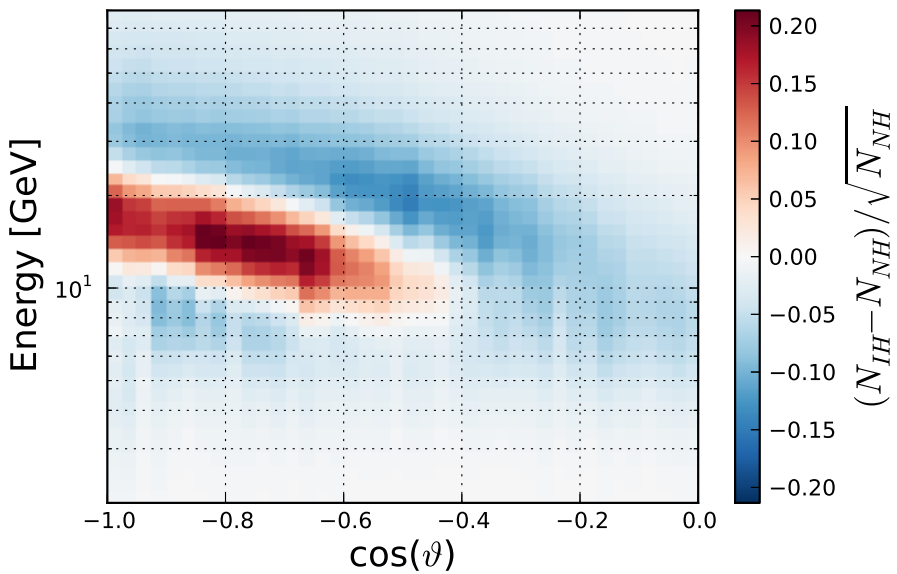
Structure of Sensitivity Studies

- Step 1: calculate expected number of events
  - Physics: atmospheric flux, oscillation, cross-sections etc.
  - Detector-specific: resolution, effective mass, particle ID etc.
  - More details on this later on in this presentation.
- Step 2: extract mass hierarchy significance
  - χ^2-significance
  - Fisher Information Matrix (PINGU LoI)
  - Pseudo-experiments and log likelihood-ratio (ORCA sensitivity study)



Example: expected number of events as a function of energy and zenith angle for ORCA's track channel.

Akhmedov-style χ^2-significance

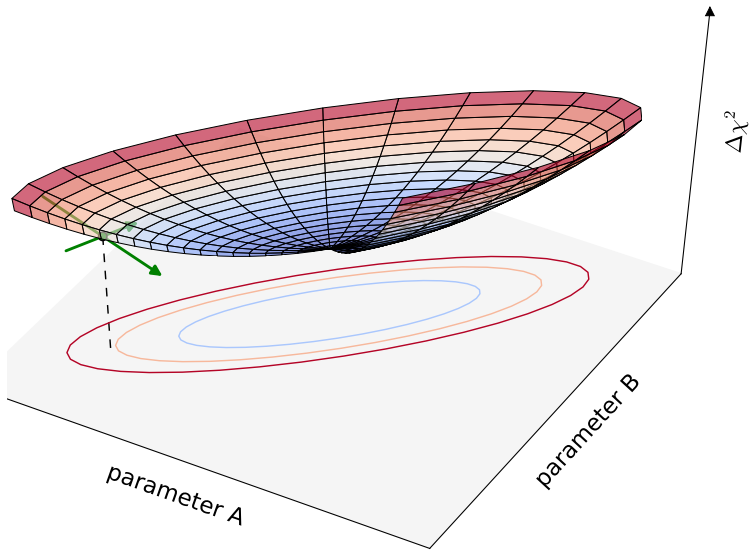


Example: χ^2-significance for PINGU's track channel.

- Bin-by-bin comparison
- Asymmetry $A_i \equiv \frac{N_i^{\mathrm{NH}} - N_i^{\mathrm{IH}}}{\sqrt{(N_i^{\mathrm{NH}})}}$
- $\chi^2 \equiv \sum_i A_i^2$
- Easy to calculate
- Good for debugging
- Shows most important regions
- No parameter uncertainties taken into account

Fisher Information Matrix (FIM)

- Used in PINGU analysis
- Use 'fiducial' values (fixed true values)
- Evaluate bin-by-bin first-order derivatives of expected number of events
  $\Rightarrow$ probe small region around fiducial values
- Covariance matrix from derivatives
- Yields individual and combined uncertainties
- Requires that probed region is sufficiently linear
  - This was checked to be the case
- Quick and easy to add many parameters



Making a linear extrapolation in a multi-parameter space.

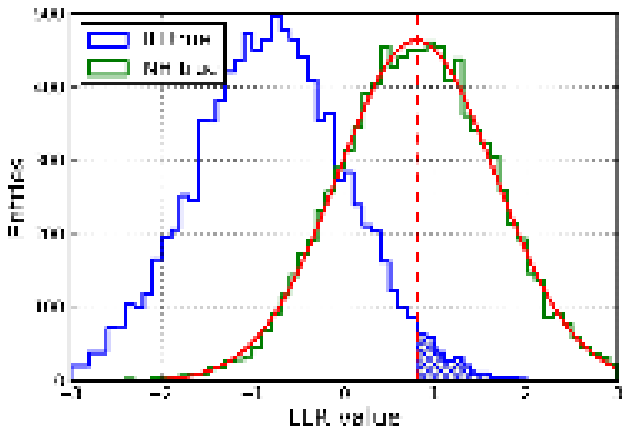
Pseudo-experiments and log likelihood-ratio

- Basis for ORCA analysis
- Used by PINGU to cross-check FIM results
- Method
  - Generate pseudo-experiments
  - Fit hierarchy and parameter values
  - NH/IH log likelihood-ratio (LLR) is test statistic
  - Median significance from LLR distribution
- Takes full parameter correlations into account
- Can handle non-linear behaviour (θ_{23} octant, CP-phase)
- Can work without fiducial values (marginalize), picking true values at random
- Number of free parameters limited by computation time

Are the Fisher Matrix and LLR-Method Equivalent?

Toy study from PINGU

- Templates computed on 2D grid in θ_{23} and Δm^2_{31}.
- Other parameters kept fixed
- Pseudo-experiments drawn from one of the templates
- Minimization on grid for NH and IH hypothesis
- Median significance from Gaussian fit to LLR distribution
- Significance **equal to Fisher Matrix result**.
- Differences could still be possible in more complicated cases.



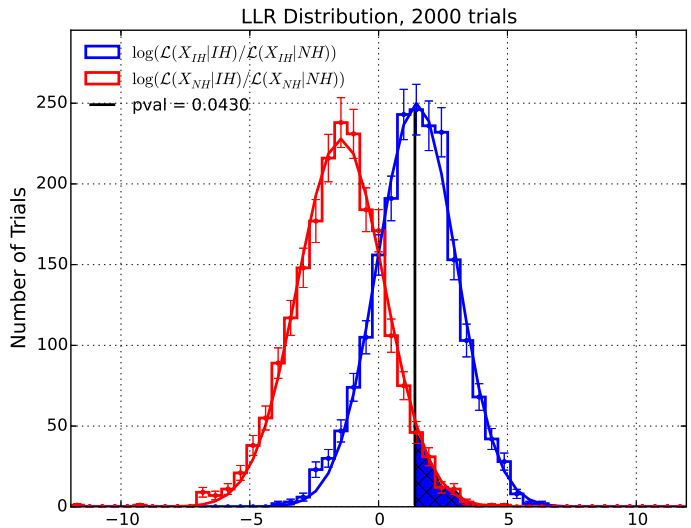
Log likelihood-ratio distributions for true NH and true IH pseudo-experiments.

Are the Fisher Matrix and LLR-Method Equivalent?

Recent more extensive study by Tim Arlen (PINGU)

- Compares Fisher method and LLR-method
- Full minimization
- Five most important systematics: Δm^2_{31}, θ_{23}, θ_{13}, ν and $\bar{\nu}$ cross-section
- Sensitivity
  - 1.717σ (LLR)
  - 1.638σ (FIM)

  Compatible within expected statistical uncertainty (10% @ 2.1k trials)



LLR distributions showing the p-value. Courtesy of Tim Arlen.

Table of Contents

Overview of Sensitivity Studies

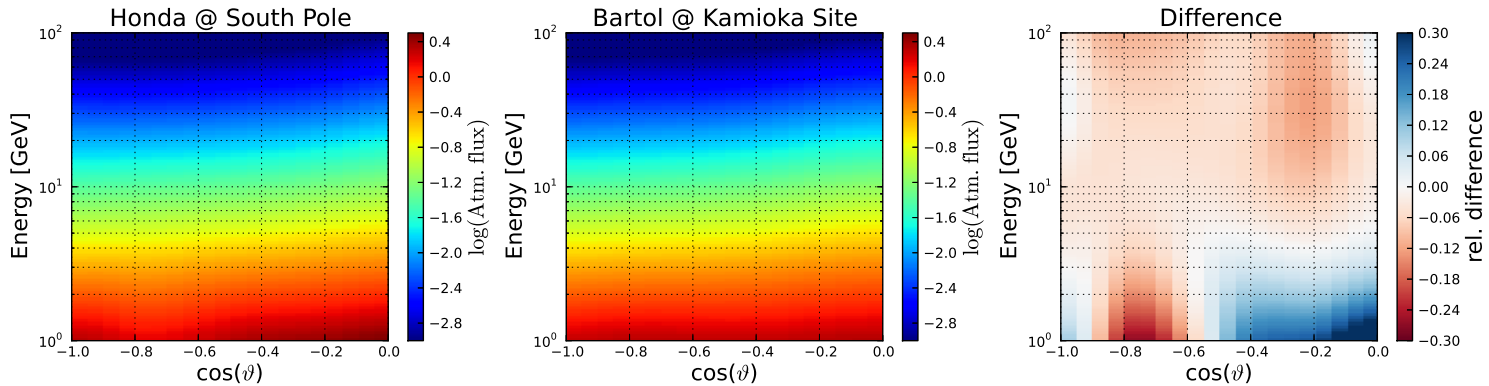
Comparison of Ingredients and Assumptions

Ongoing Cooperation

Recap and Conclusions

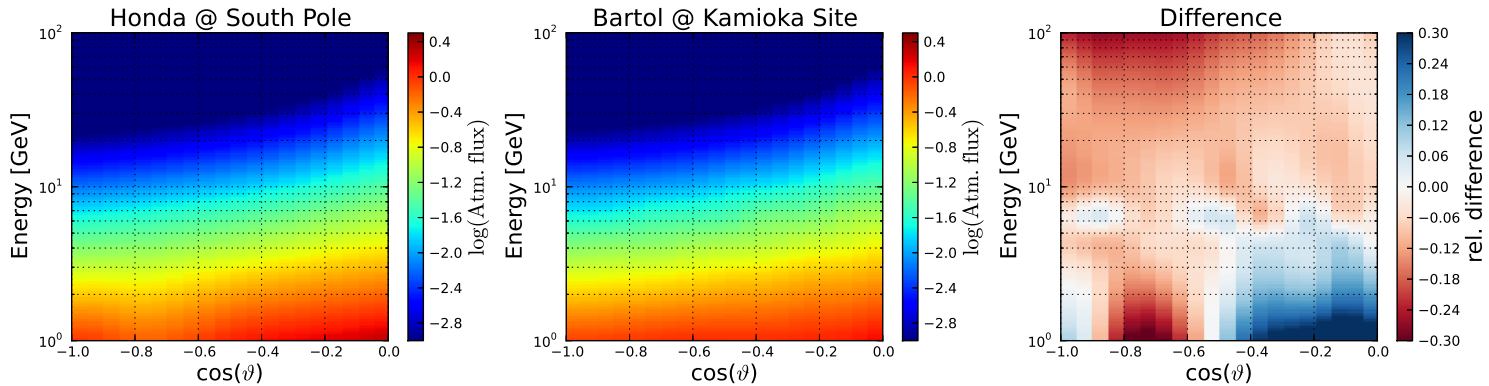
Backup Slides

Atmospheric Fluxes



Comparison of the atmospheric neutrino fluxes used by either analysis. Shown is the **muon neutrino** *flux as a function of neutrino energy and zenith angle.*
Relative differences up to 30%.

Atmospheric Fluxes

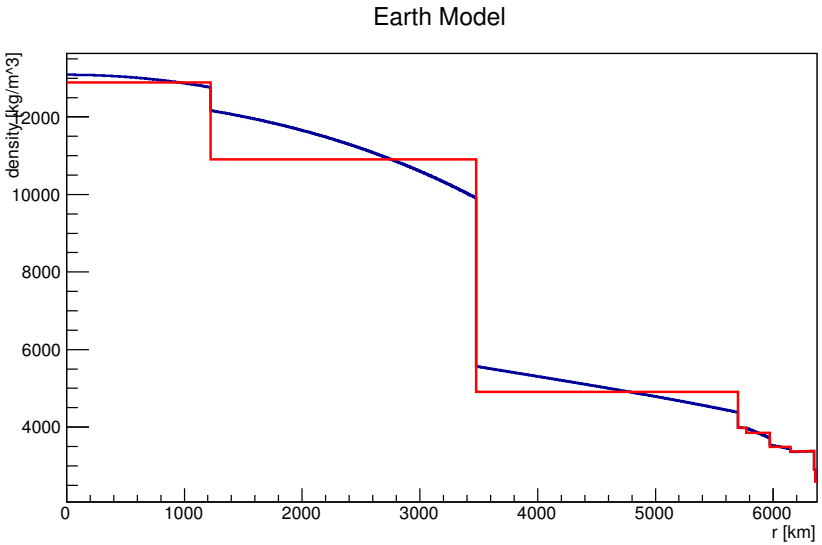


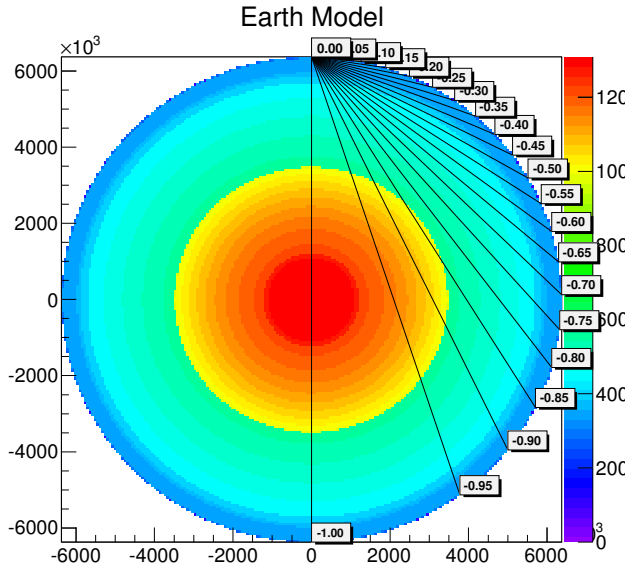
Comparison of the atmospheric neutrino fluxes used by either analysis. Shown is the **electron neutrino** *flux as a function of energy and zenith angle.*

Relative differences up to 30%.

Earth Model

Preliminary Reference Earth Model (PREM)



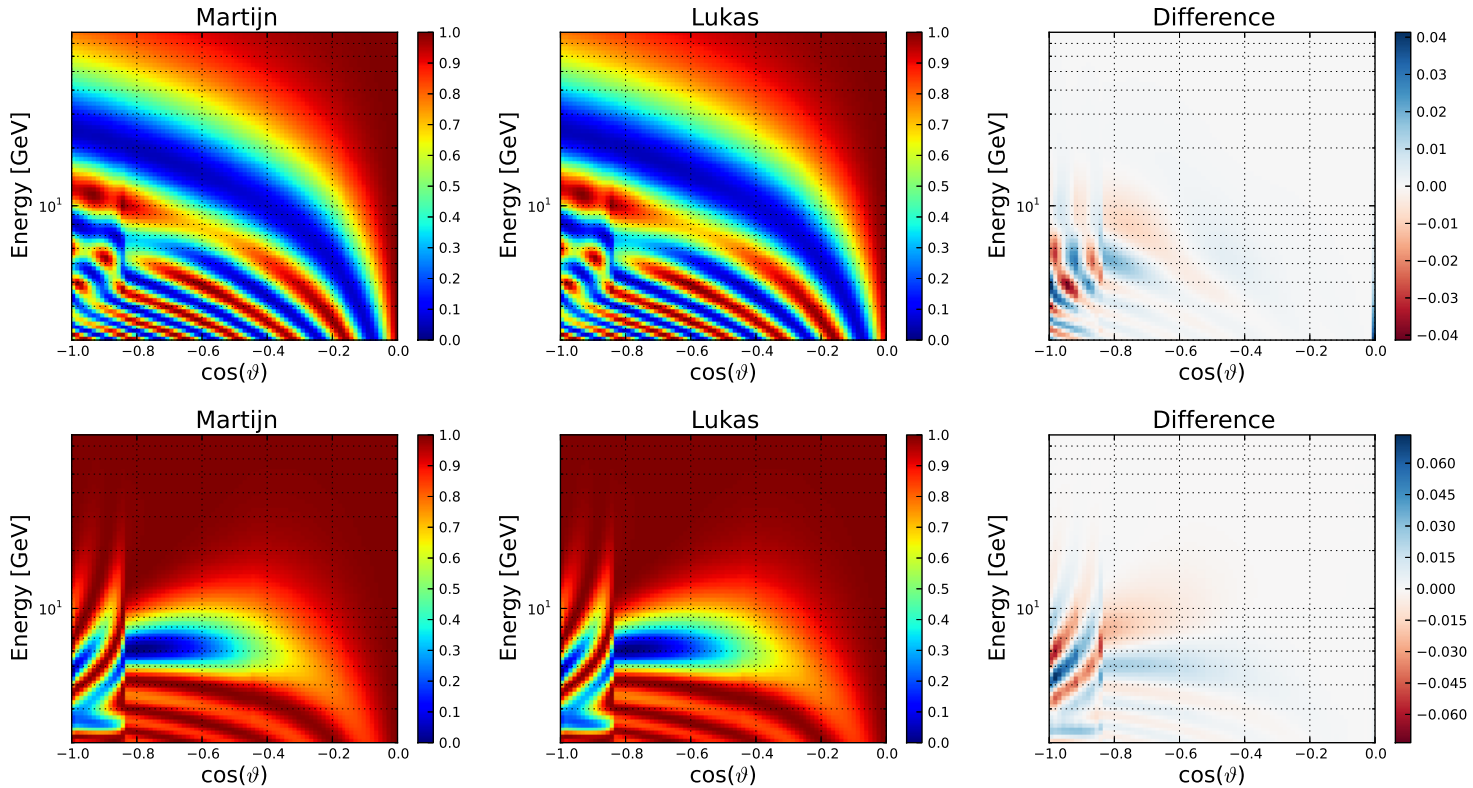


Left: PREM density versus radius. The red line shows the shell model used in the ORCA analysis to speed up the calculations.
Right: 2D representation of the model showing neutrino trajectories for $\cos(\theta)$ values at 0.05 intervals.

Oscillation

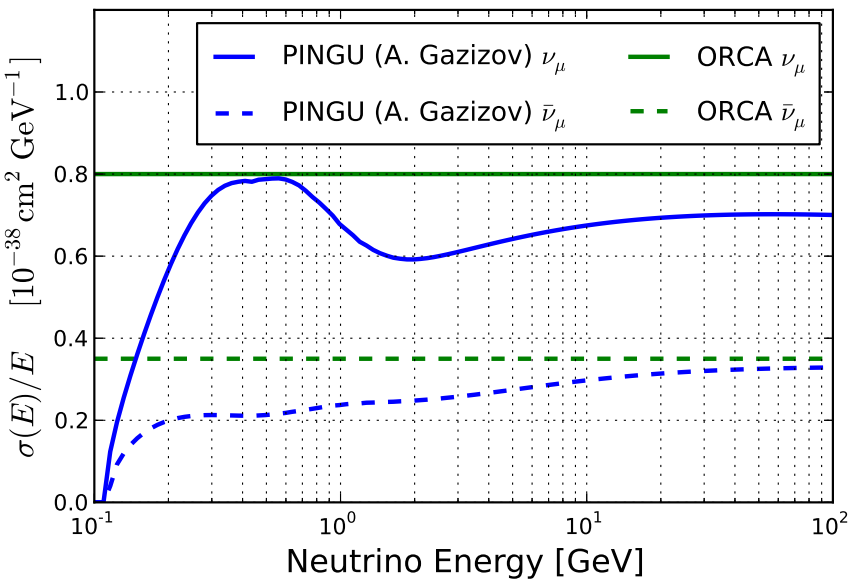
- ORCA: custom oscillation code
  - Based on Cayley-Hamilton formalism (Ohlsson, Snellman)
  - Tested to be in agreement with GLoBES
  - Oscillation calculated on-the-fly (necessary for fit)
- PINGU:
  - AtmoWeights (custom IceCube code)
  - NuCraft (hepforge)
  - Tested to be consistent.
  - Oscillation probabilities are pre-computed
  - Multiple bin sampling to avoid resonances at low energies.
  - Probably switch to Super-K's prob3 in the future. (already used in PINGU LLR analyses)

Oscillation



Oscillation probabilities from ORCA (Martijn) and PINGU (Lukas) code.
The rightmost plots show the relative differences (∼few percent).
The binning is 80x80 and each bin is sampled once.
Top: $P(\nu_\mu \to \nu_\mu)$. **Bottom:** $P(\nu_e \to \nu_e)$

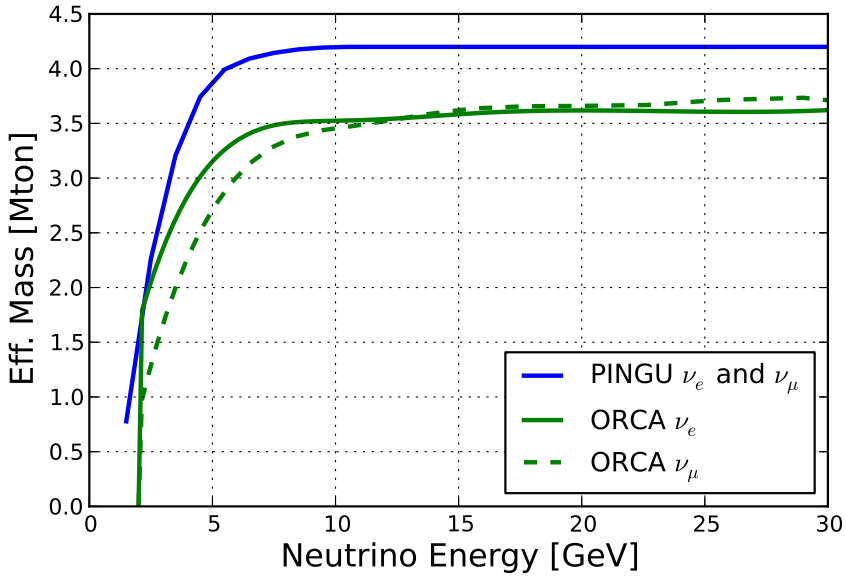
Neutrino-Nucleon Cross Section



Neutrino-nucleon cross-sections used in the sensitivity studies.

- ORCA:
  $\sigma_\nu^{CC} = 0.80 \times E \times 10^{-42}$ m^2
  $\sigma_{\bar{\nu}}^{CC} = 0.35 \times E \times 10^{-42}$ m^2
  NC and ν_τ ignored (effects studied to be small)
- PINGU:
  A. Gazizov, M. Kowalski, K. S. Kuzmin, V. A. Naumov, C. Spiering *Neutrino-nucleon cross-sections at energies of Megaton-scale neutrino detectors*, to be published.
  All flavours included.

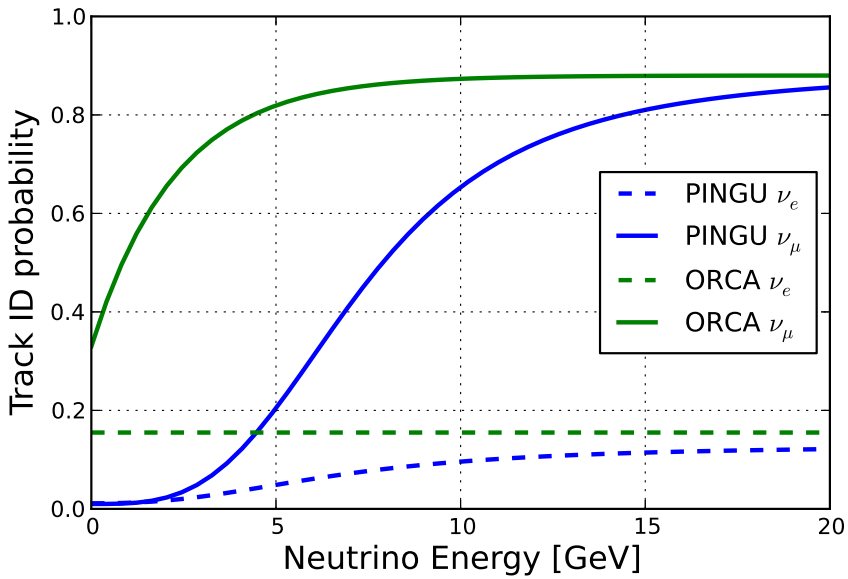
Effective Mass



Effective masses used in the sensitivity studies.

- PINGU: cut on number of detected photons, successful reconstruction, containment of reco vertex, and only upgoing reconstructed events (No MC information).
- ORCA: quality cut on muon tracks. For e-like events: $\geq$3 L1 hits (14 ns coincidence) and true vertex inside instrumented volume.
- Note: muon contamination of signal not yet studied by PINGU or ORCA.

Particle Identification



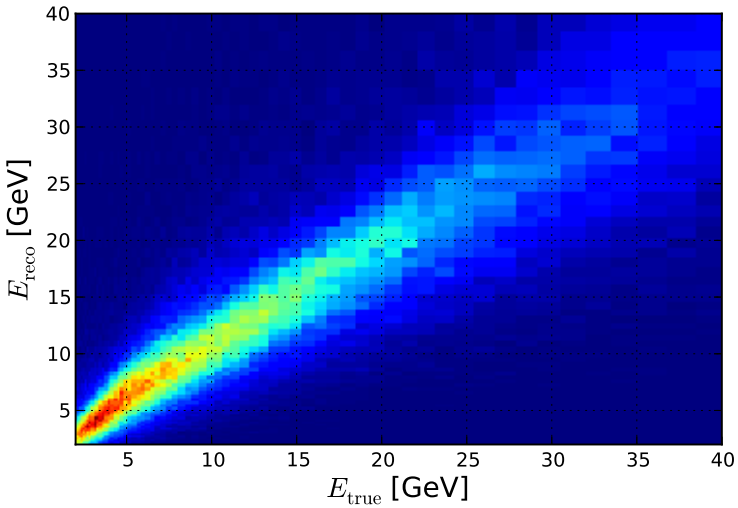
Probability to identify an event as a track.

- PINGU
  - Boosted Decision Tree
- ORCA:
  - Random Decision Forest
  - First study using Premium Events
  - Optimistic
  - To be improved

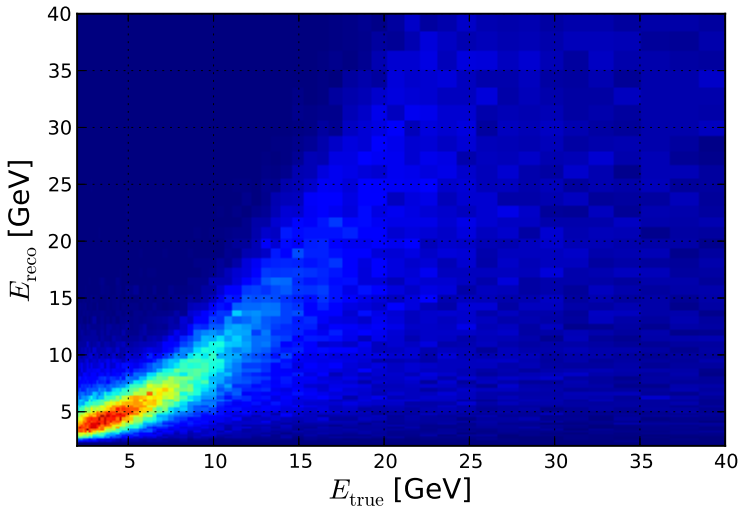
Detector Resolution

- PINGU: 4D smearing histogram
  E_{true}, $\cos(\theta_{\text{true}}) \rightarrow E_{\text{reco}}$, $\cos(\theta_{\text{reco}})$.
- ORCA: 2D energy + 3D angle smearing histogram
  - $E_{\text{true}} \rightarrow E_{\text{reco}}$
  - E_{true}, $\cos(\theta_{\text{true}}) \rightarrow \cos(\theta_{\text{reco}})$
- NOTE: misidentified events should have different resolutions.
  Not yet implemented for either of the two experiments.

Detector Resolution (ν_μ energy smearing)

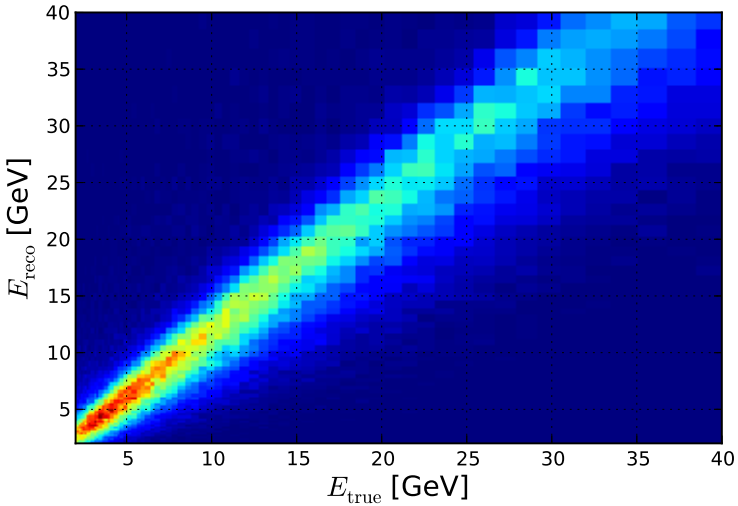


Projection of PINGU's resolution in the log(E_{true})-log(E_{reco}) *plane. See talk by J.P. de André.*

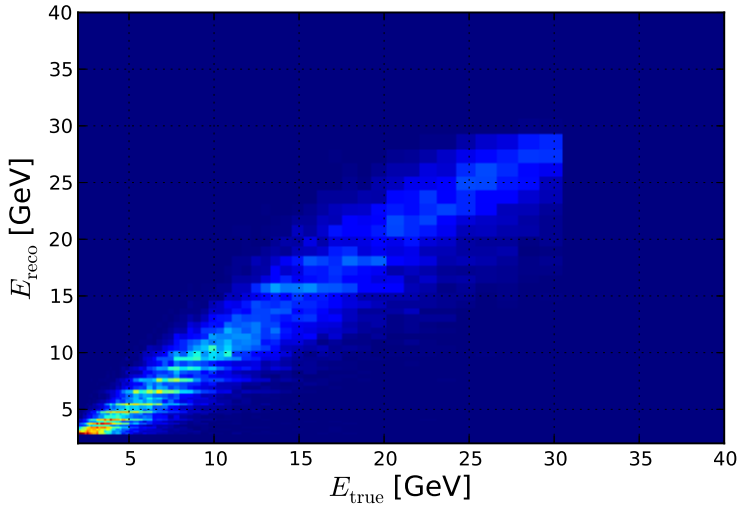


Energy smearing used for ORCA analysis. Based on track-fitting algorithm. See talk by Jannik Hofestädt.

Detector Resolution (ν_e energy smearing)

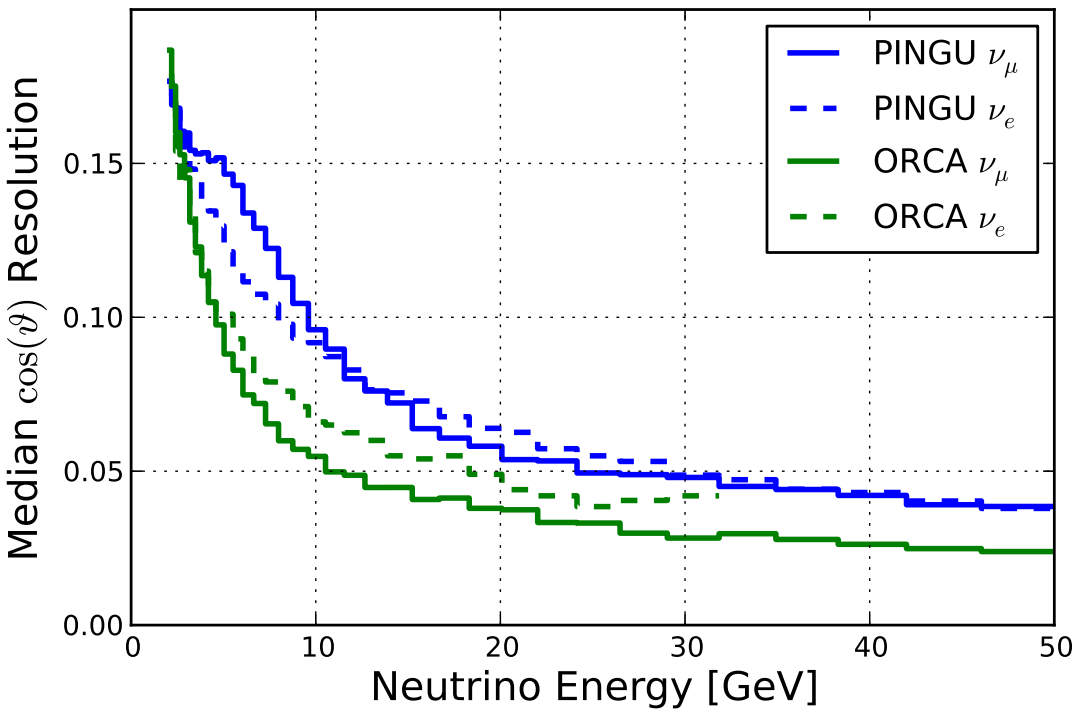


Projection of PINGU's resolution in the log(E_{true})-log(E_{reco}) *plane. See talk by J.P. de André.*



Energy smearing used for ORCA analysis.
Originally based on Premium Events but similar resolution has been achieved in a full study - see talk by Jannik Hofestädt.

Detector Resolution (zenith angle)

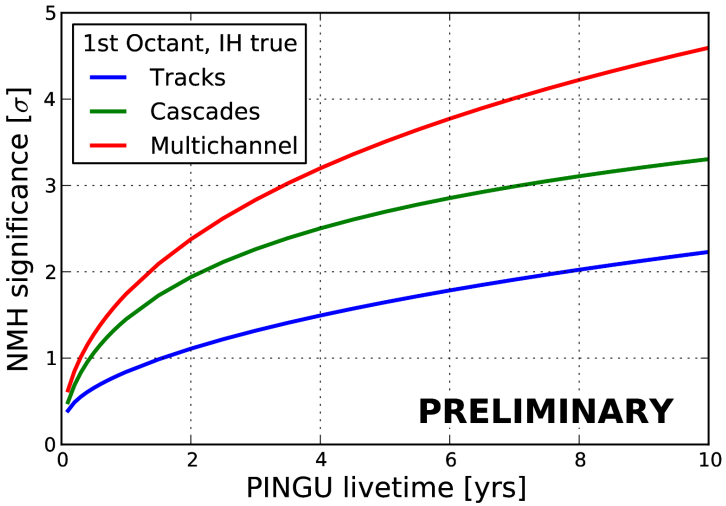


- Same note as before for ORCA ν_e
- Original study based on Premium Events but new results are similar.
- See talk by Jannik Hofestädt for more info.

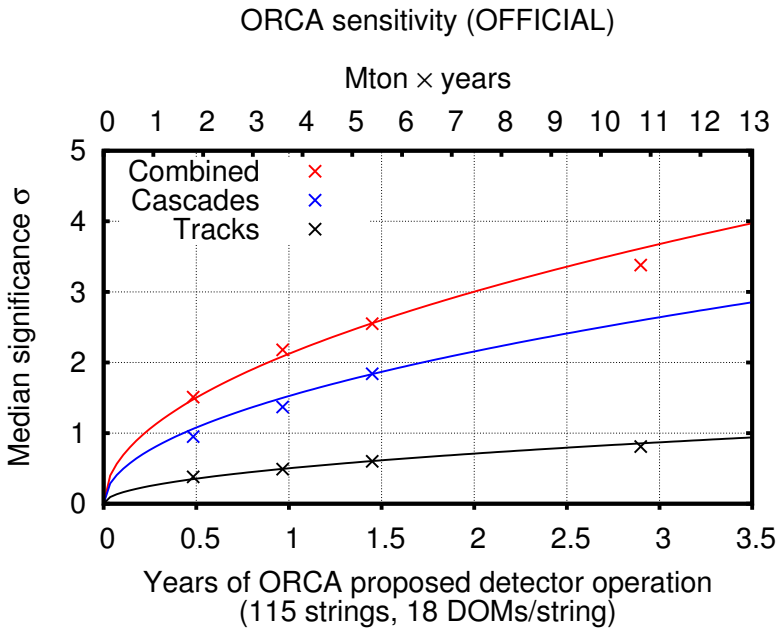
Systematics

- PINGU: many systematics studied
  - Overall flux factor
  - $\nu/\bar{\nu}$ ratio
  - ν_e/ν_μ ratio
  - Energy scale
  - Oscillation parameter uncertainties
    All except δ_{CP}
    Ignoring θ_{12} and δm^2_{21} (shown to be negligible)
  - ...and more
- ORCA:
  - Full correlations of the oscillation parameters.
  - Most influential parameters (θ_{23}, ΔM^2 and δ_{CP}) fitted from data
  - Other parameters as nuisance
  - Currently studying effect of overall flux factor
  - More systematics will be added

Current Official Results

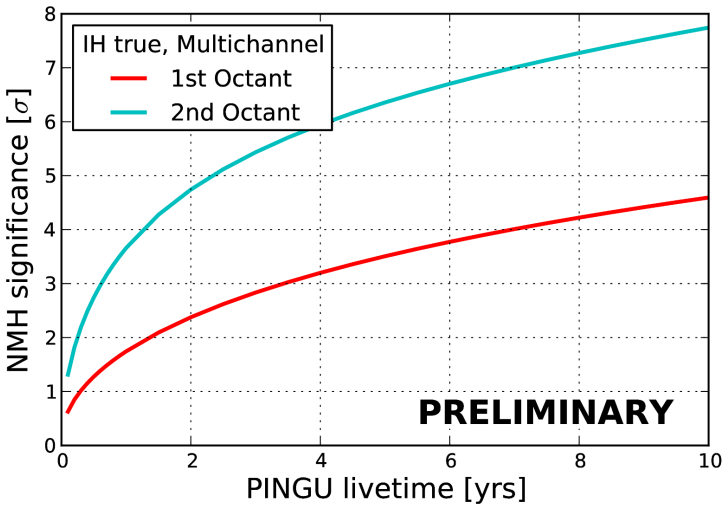


PINGU official hierarchy significance plot (from LoI)

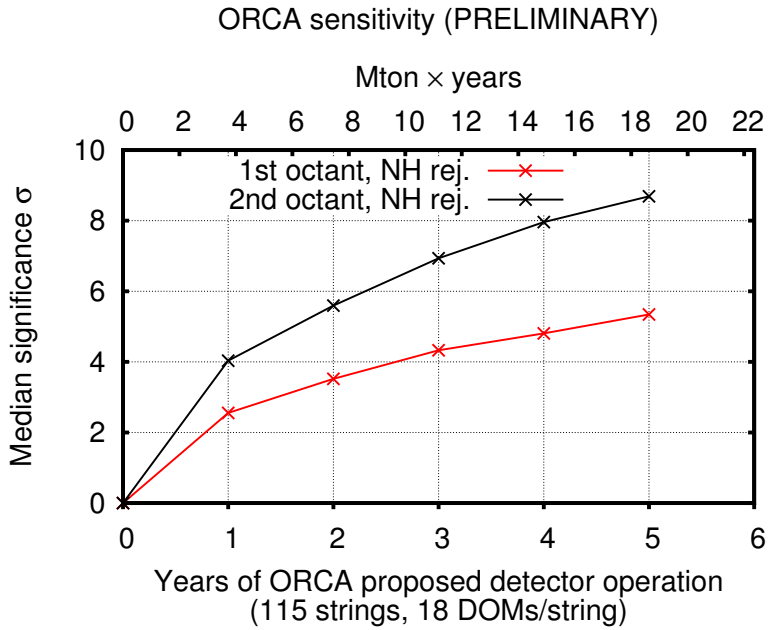


ORCA official hierarchy significance plot, using the ingredients described throughout this presentation.

Current Official Results



PINGU official hierarchy significance plot for first and second octant (from LoI)

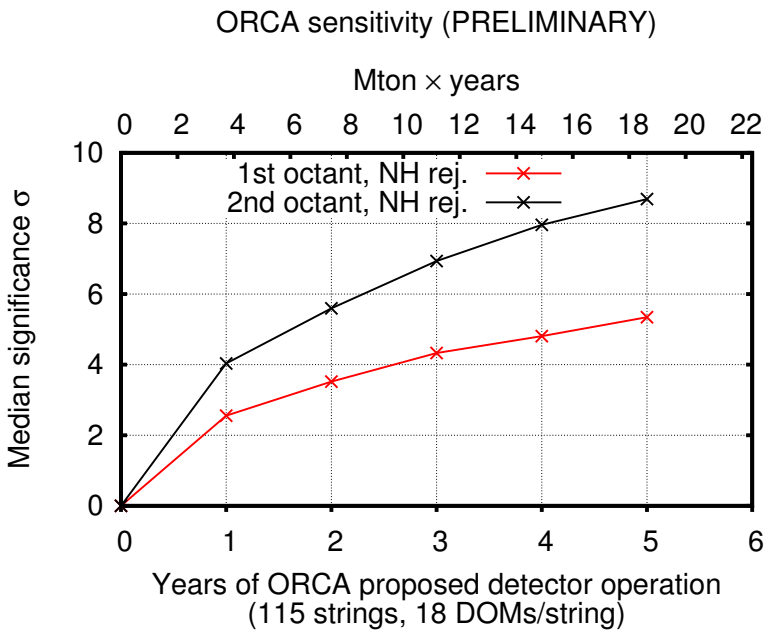


ORCA **preliminary** *plot for illustration purposes only.*

Current Official Results

Disclaimer for plot on the right

- Generated differently than official plot!
- Using PINGU fiducial values instead of distributions
- Gaussian fits made to the LLR distributions
- Parameters δ, θ_{23} and ΔM^2 are fitted
- Other uncertainties are **not** taken into account.
- Currently not understood feature: 2nd octant IH rejection (not shown) is very high.



ORCA **preliminary** *plot for illustration purposes only.*

Current Official Results - Disclaimer

Comparison should not be made carelessly. Many differences between analyses.

- Parameter values (PINGU: fiducial, ORCA: distribution)
- Systematics
  - PINGU: uncertainties on oscillation parameters and many other systematics.
    For full list see PINGU LoI.
  - ORCA: θ_{23}, ΔM^2, δ fitted from data alone.
    Uncertainty on other oscillation parameters taken into account.
    No other systematics.
- ORCA: no NC/tau and optimistic particle ID
- Systematic uncertainty on implementation details (flux, oscillation code, binning, range etc.)
- No definitive comparison can be made at this point!

Table of Contents

Overview of Sensitivity Studies

Comparison of Ingredients and Assumptions

Ongoing Cooperation

Recap and Conclusions

Backup Slides

Comparing Sensitivity Studies

Initial Idea

- Exchange input and settings
- Both run analysis in 'PINGU mode' and 'ORCA mode'
- See that significances are identical

Results are not the same!

- Comparison not that easy
- Many details go into implementation
- Differences must be identified and dealt with

Comparing Sensitivity Studies

Intermediate step

- Trying to get identical results for simple toy model
- Have converged up to few percent differences
- Getting similar values for hierarchy significance and measurement of δM^2 and θ_{23}.

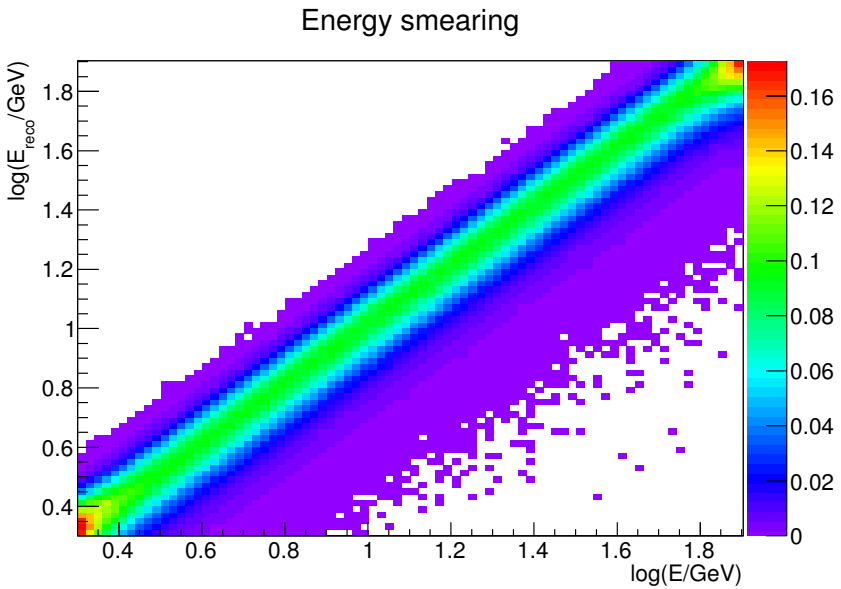
Final goal

- Full comparison
- Hope to present final results at next MANTS meeting.

Simple Toy Model

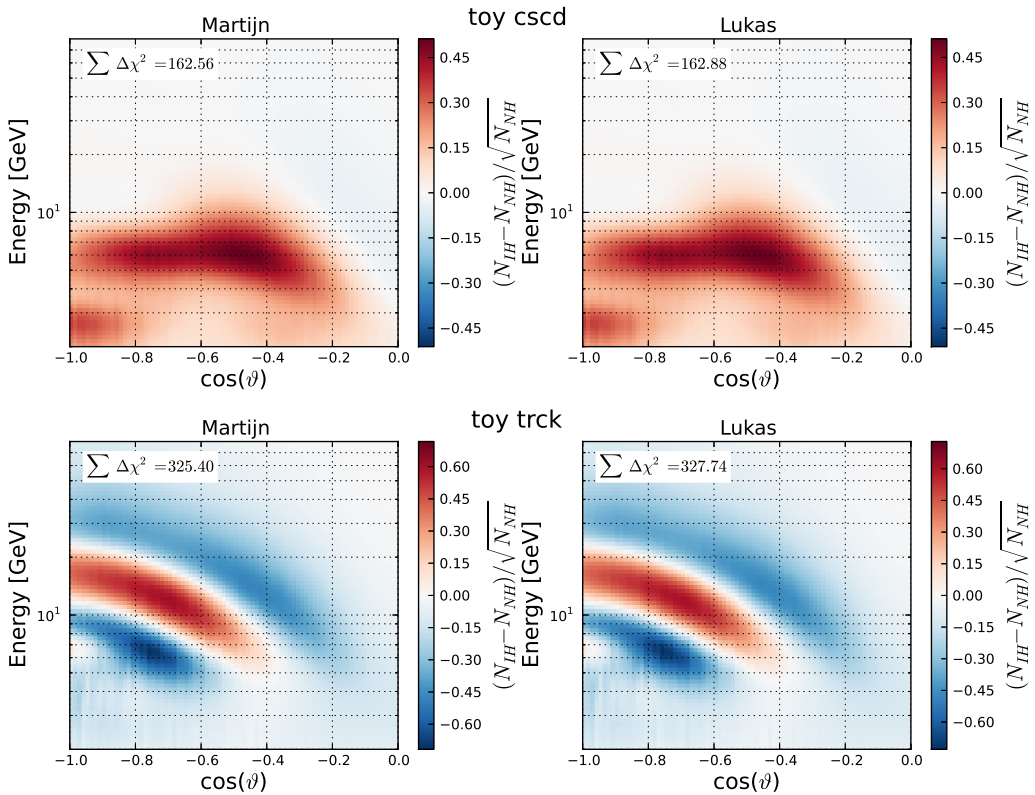
Ingredients

- Parametrized atmospheric flux
- Constant 4 Mton effective mass
- Resolution same for all channels
- Perfect particle ID
- Ignoring ν_τ and NC events
- Using PINGU default fiducial parameter values
- Only θ_{23} and ΔM^2 as free parameters



Energy smearing used for the toy model. It is based on Gaussian-generated "MC events".

Simple Toy Model - significance



Mass hierarchy asymmetry in the toy model using ORCA (**left**) and PINGU (**right**) code for the cascade channel (**top**) and the track channel (**bottom**).

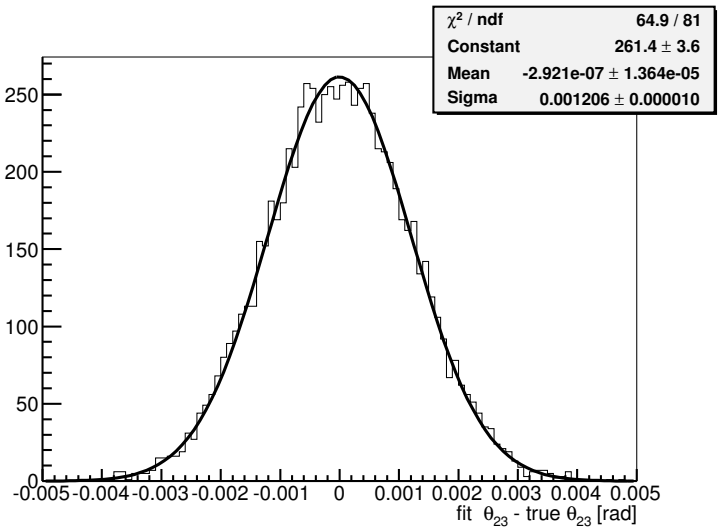
Simple Toy Model - significance

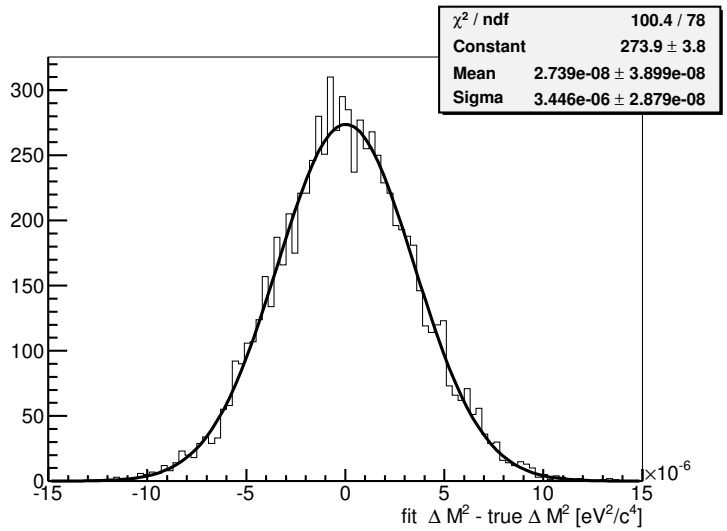
χ^2-significance from ORCA/PINGU code:
22.09/22.15 $\Rightarrow$ **0.27% difference**

	No free pars.	θ_{23} free	ΔM^2 free	both free
LLR	22.50	19.11	17.76	13.971
FIM	22.15	19.61	17.44	14.21
difference	**1.6%**	**2.6%**	**1.8%**	**-1.7%**

Hierarchy significance in σ for the toy model. With current statistics, the estimated error on the LLR method is ∼2%.

Simple Toy Model - parameter uncertainties

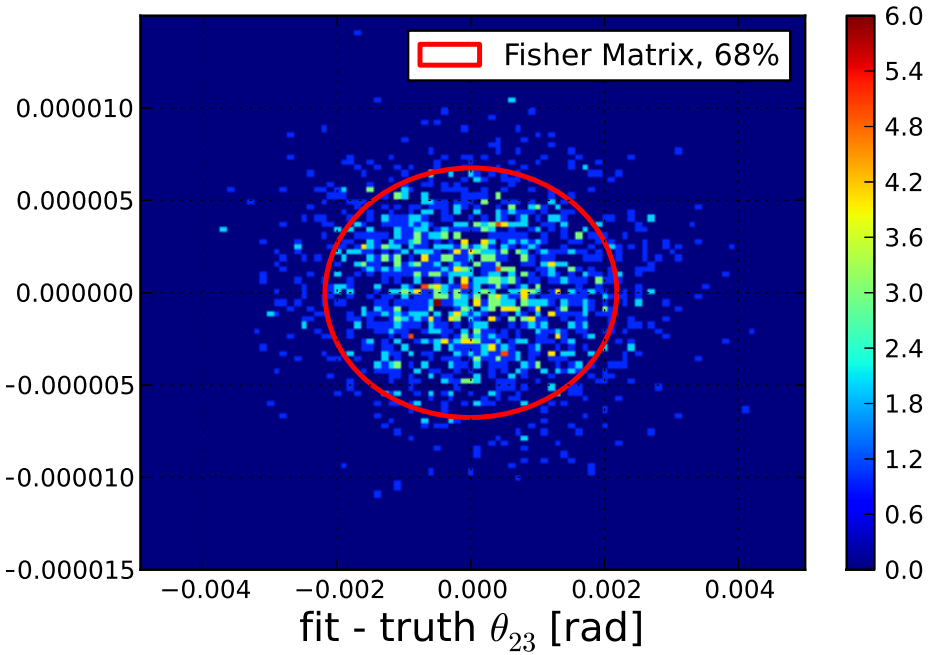




	fitting both	fitting only one parameter
$\sigma(\theta_{23})$ 10^{-2} deg	6.910/8.284 (**-17%**)	6.807/6.732 (**1.1%**)
$\sigma(\Delta M^2)$ 10^{-6} eV^2	3.446/4.765 (**-28%**)	3.403/3.208 (**6.1%**)

Uncertainty on parameters as fitted from data. Format:
*Martijn (LLR)/Lukas (FIM) (**relative difference**)*
The estimated error on the fitted values of the LLR method is ∼2%
The discrepancies are still to be resolved.

Simple Toy Model - parameter correlation



Correlation between the error on the fitted θ_{23} and ΔM^2 from the LLR method. Overlaid is an ellipse showing the result of the FIM method. The correlation coefficients we get are not consistent: -8.9%/34.3% (LLR/FIM). To be investigated. Note that these results are very recent.

Table of Contents

Overview of Sensitivity Studies

Comparison of Ingredients and Assumptions

Ongoing Cooperation

Recap and Conclusions

Backup Slides

Recap and Conclusions

Recap

- Discussed ORCA and PINGU sensitivity studies
- Compared ingredients
- Discussed ongoing efforts to get consistent results

Conclusions

- Comparing ORCA and PINGU is far from trivial
- Effort must be put into honest comparison
- Will continue cooperating on this

Goals for Near Future

- Eliminate remaining differences from simulation chain.
- Gradually build up more complex model and compare
  - Mass hierarchy significance
  - Parameter uncertainties
  - Parameter correlations
  - Other systematics

Table of Contents

Overview of Sensitivity Studies

Comparison of Ingredients and Assumptions

Ongoing Cooperation

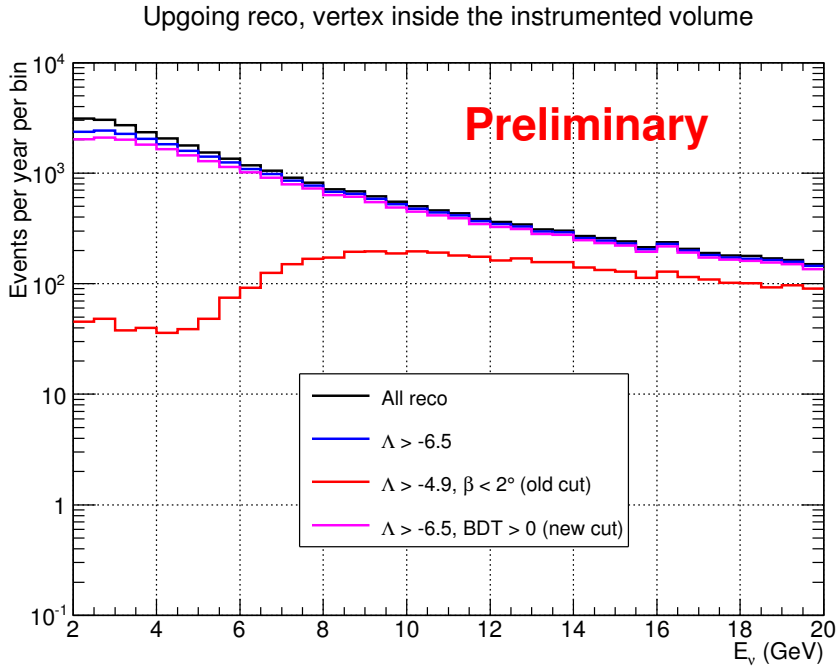
Recap and Conclusions

Backup Slides

PINGU fiducial parameter values

- Used for PINGU mass hierarchy sensitivity
- Also used for toy study
- Mixing angles:
  $\theta_{23} = 38.6°$
  $\theta_{23} = 8.93°$
  $\theta_{23} = 33.6°$
- Mass-squared differences:
  $\Delta m^2_{31} = 2.46 \times 10^{-3}$ eV^2
  $\Delta m^2_{21} = 7.54 \times 10^{-5}$ eV^2
- Definition of "hierarchy flip" is that
  $\Delta M^2 \equiv \Delta m^2_{31} - 0.5 \times \Delta m^2_{21}$
  switches sign

Atmospheric Muon Rejection in ORCA



Result of different atmospheric muon rejection cuts on the effective mass. The blue line shows the cut used for the sensitivity study, the violet one ('new cut') the most recent result for which only 10% muon contamination remains in the energy region of interest.
See presentation by Luigi Fusco.

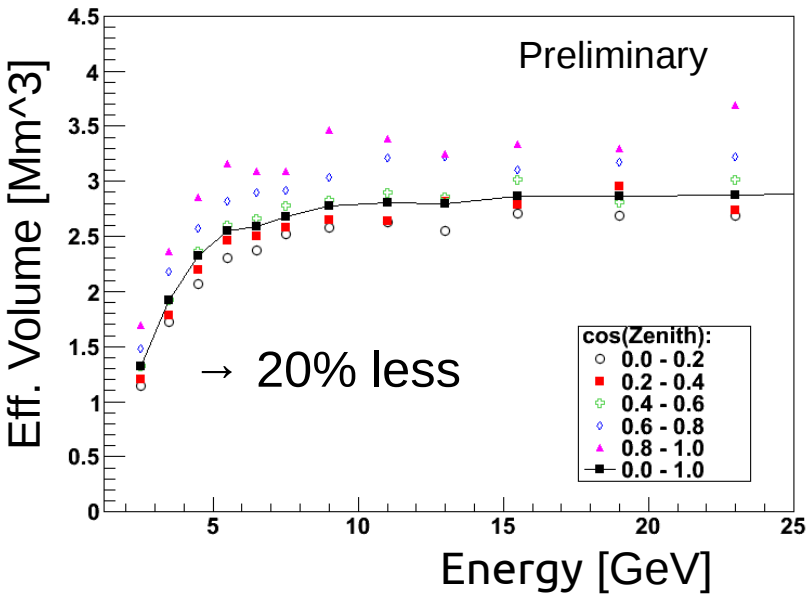
Shower Reconstruction in ORCA

- The following two slides are courtesy of Jannik Hofestädt
- Comparison of cascade energy/angle resolution used in sensitivity study and newest results
- Conclusion: the new results from full simulation are comparable to the initial results from Premium Events.

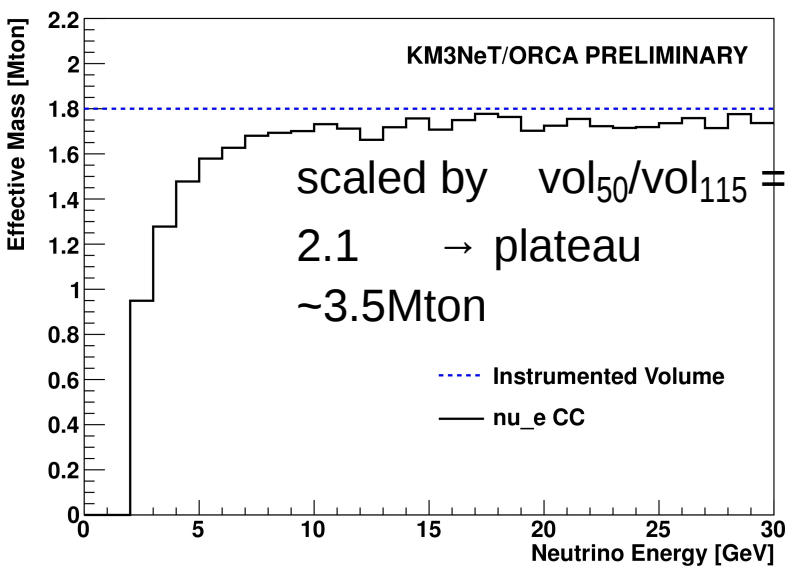
Comparison with the Past (I)

- Official sensitivity study used:
  - reference detector
  - resolutions from 'premium events' (→ optimistic assumptions)
  - effective volume of 50 string detector scaled to 115 string detector

Now



Sensitivity study input



Comparison with the Past (II)