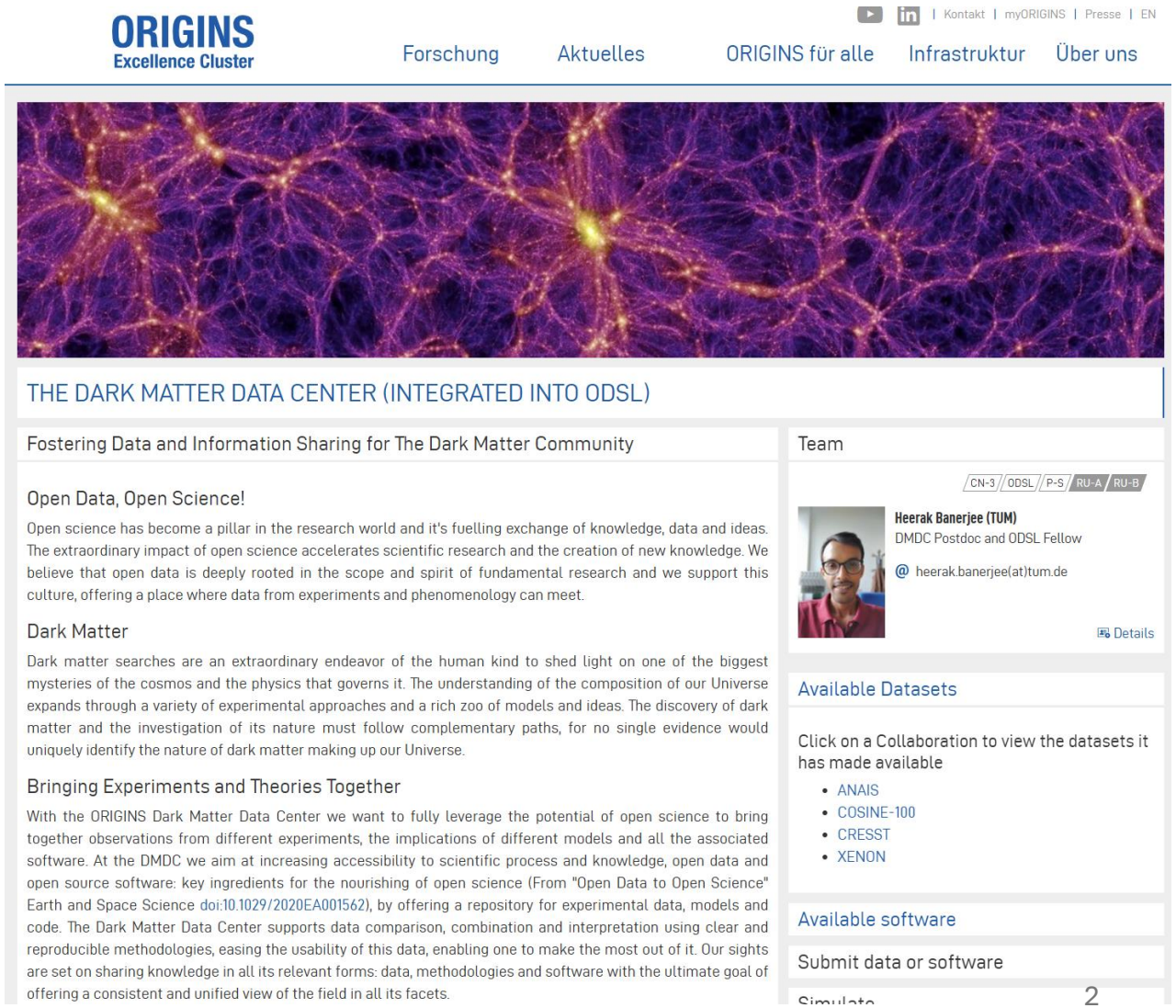
ORIGINS Data Center & Neutrino Oscillation Models

"Democratizing Models" Workshop

Philipp Eller (TUM)

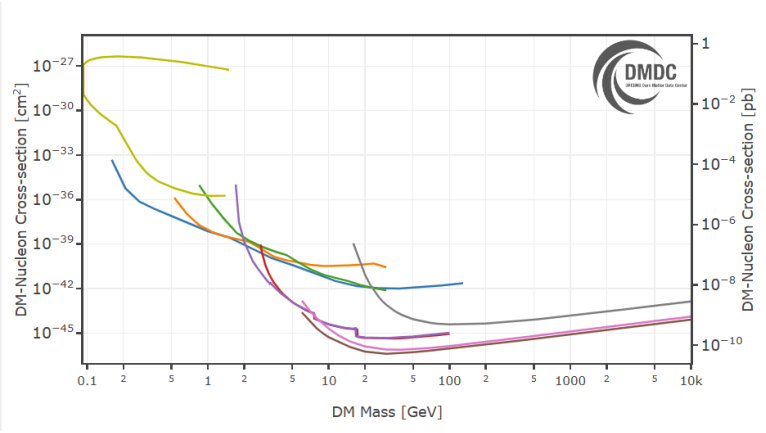
Dark Matter Data Center (DMDC)

- Initiative of the ORIGINS Excellence Cluster in Munich
  - https://www.origins-cluster.de/odsl/dark-matter-data-center
  - Funding Period 2019 – 2025
  - The idea is to provide a central place for (direct) DM detection data
  - Open source / open data
  - facilitate and carry out combined analyses
  - Bring together different DM models (theory) and confront those with different data (experiment)



DMDC

- Currently DMDC has datasets from
  - ANAIS
  - COSINE-100
  - CRESST
  - XENON



Interactive limit plot

- For some experiments, DMDC provides example notebooks on how to run a fit
- At the moment, DMDC is very person-power limited (1 postdoc, Heerak Baerjee basically *is* the DMDC)

→No common model / data format exists yet

ORIGINS 2

- We have submitted the funding proposal for ORIGINS 2
  - Funding Period 2026 – 2032

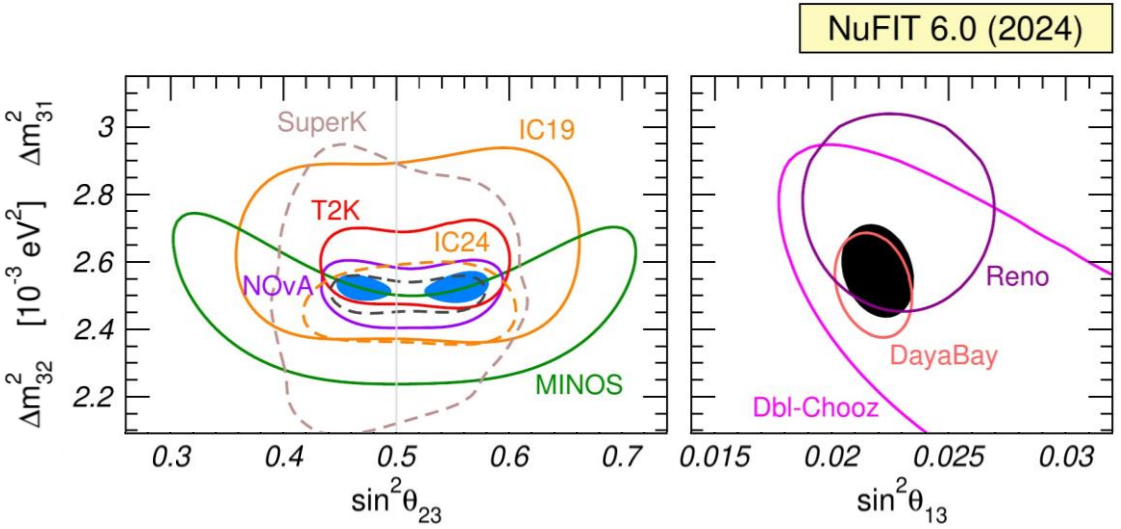
- We have put in an extension of the DMDC to become the larger ORIGINS Data Center (ODC)
  - This will expand beyond direct detection DM, and in particular also include neutrino data
    - **Neutrino oscillations** (my expertiese), Neutrinoless double beta decay, Neutrino Mass, …

→ No common model / data format exists yet

Neutrino Oscillation Data

- Experiments Overview:
  - Atm.:
    - Super-K → No public Data
    - IceCube → CSV files + Notebook
    - KM3NeT
  - Solar:
    - Borexino → Some data files available, but no oscillation specific ones
    - SNO
    - SuperK → No public Data
  - Reactor:
    - KamLAND → Some text files floating around
    - SNO+
    - Double Chooz → Custom text files with data, root file for cov matrix, some tar file with example code
    - Reno
    - DayaBay → Custom text files with data on arXiv
    - JUNO (future)
  - LBL:
    - MINOS → ROOT files from a sterile neutrino analysis on arXiv
    - NOvA → ROOT files with some info, but not detector response
    - T2K → Only chi2 contours available, some actual data release rumored
    - DUNE (future)
    - Hyper-K / T2HK (future)
  - …

Current “Users”



- Global Fitting Group:
  - NuFIT: http://www.nu-fit.org/
  - Valencia: https://globalfit.astroparticles.es/
  - Bari: Lisi, Capozzi, and others, e.g. https://arxiv.org/abs/2003.08511
  - …
- Groups testing specific (BSM) models:
  - E.g. ourselves:
    - https://arxiv.org/abs/2407.20388 (PMNS tests with IC-DeepCore, DayaBay, KamLAND, IC-Upgrade, JUNO)
    - https://arxiv.org/abs/2402.00490 (SM copies with DayaBay, KamLAND, MINOS, NOvA & KATRIN)
- Every group implements their own analysis, that tries to reproduce the official results “reasonably” well. E.g. NuFIT, the perhaps most active global fitting group, they just include norm and shape uncertainties into the analysis until the results kind of match

Why Global Fits?

- In neutrino oscillations particularly important:
  - Usually, an experiment can only access one mixing mode (solar vs. atmospheric vs. reactor) (JUNO can do 2 modes)

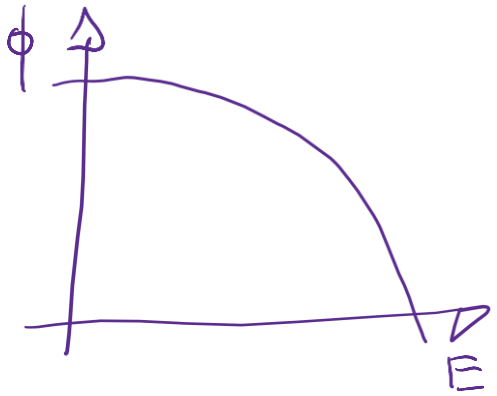
  →Need multiple experiments to constrain full set of mixing parameters
  - Currently hot topics, such as NMO or CP violation, will almost certainly be answered via global analyses:
    - E.g. T2K + NOvA for CP
    - Or Reactor + Atmospheric for NMO

- However, there is no standard on how to share data and models
  - From experience: I put together the 3y IceCube oscillation data release and was completely on my own, so I just did what I thought was best

Neutrino Oscillation Model

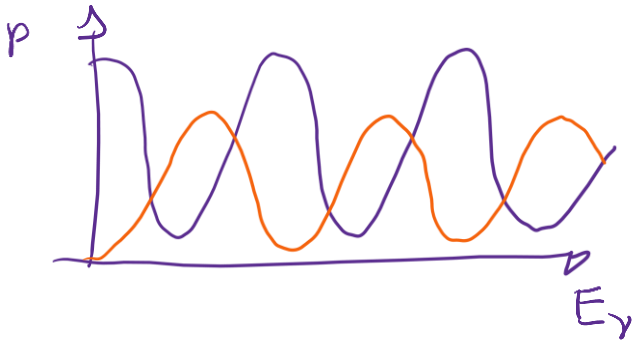
Unoscillated Spectrum

- From Model/Calculation
- From Auxiliary Measurements
- From near detector
- Inclusion of sys. uncertainties



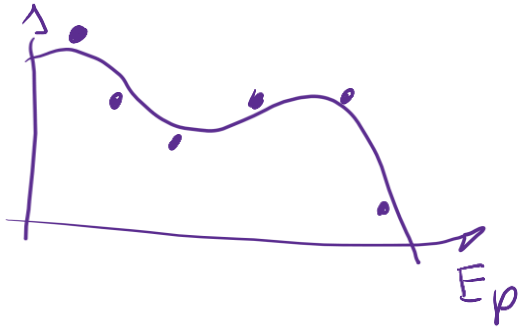
Oscillation Calculation

- Calculation in true neutrino properties (Energy, Baseline, …)
  → Not accessible
- Either 3-flavour Vacuum Oscillations (easy)
- Or BSM physics
- And/or with Matter effects (usually using external code)



Oscillated Spectrum

- Including Detector effects / acceptance, etc
- Typically realized via MC
- Inclusion of sys. uncertainties



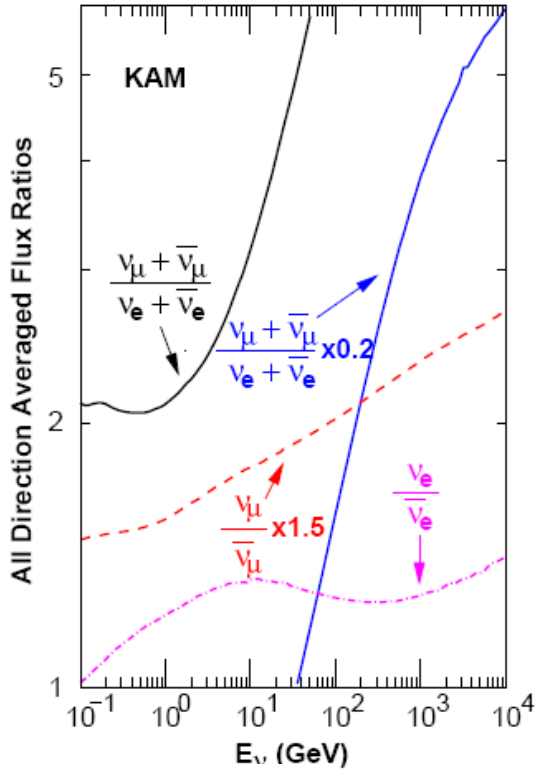
Measured Data

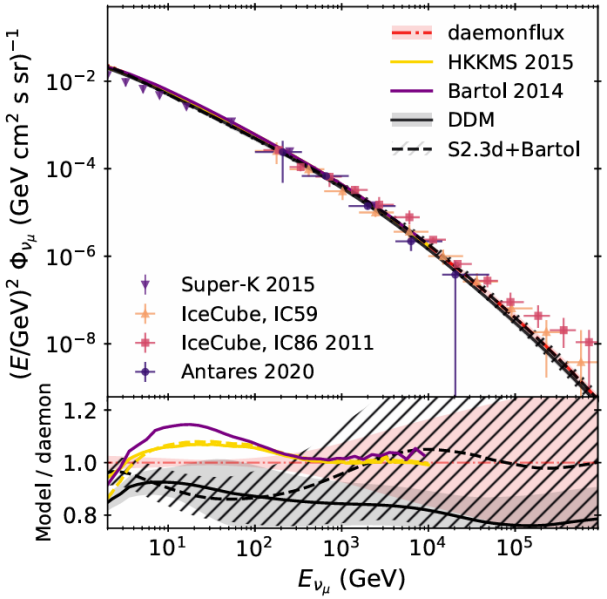
"standard" likelihood

E.g. binned Poisson

Example: Atm. Flux

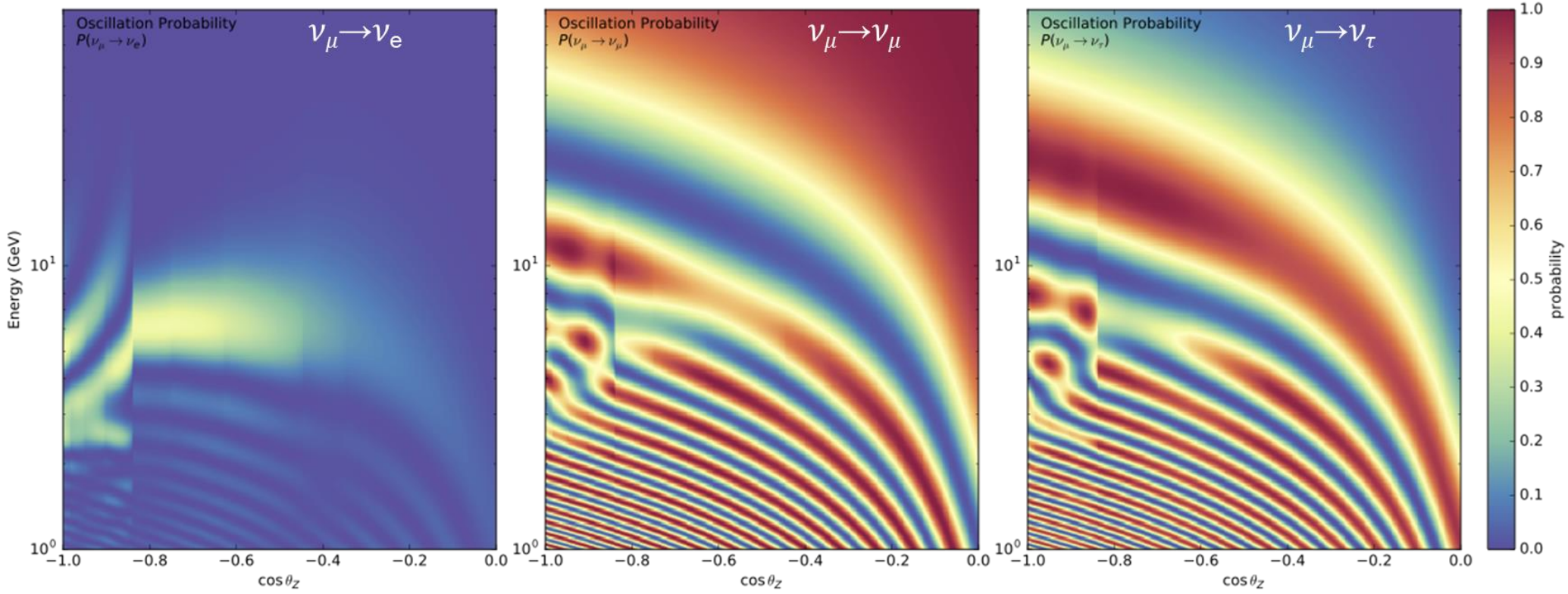
- Based on external calculations
  - For example HKKM
  - Tabulated values in 2-3 dimensions
- Or, based on on-the-fly calculations
  - For example MCEq
- Or, data-driven unfolding model
  - E.g. DEAMON flux
- All of those are in truth variables





Oscillations

- Example prob3 function + PREM earth profile, ca. 1000 lines of highly optimized code, highly non-linear

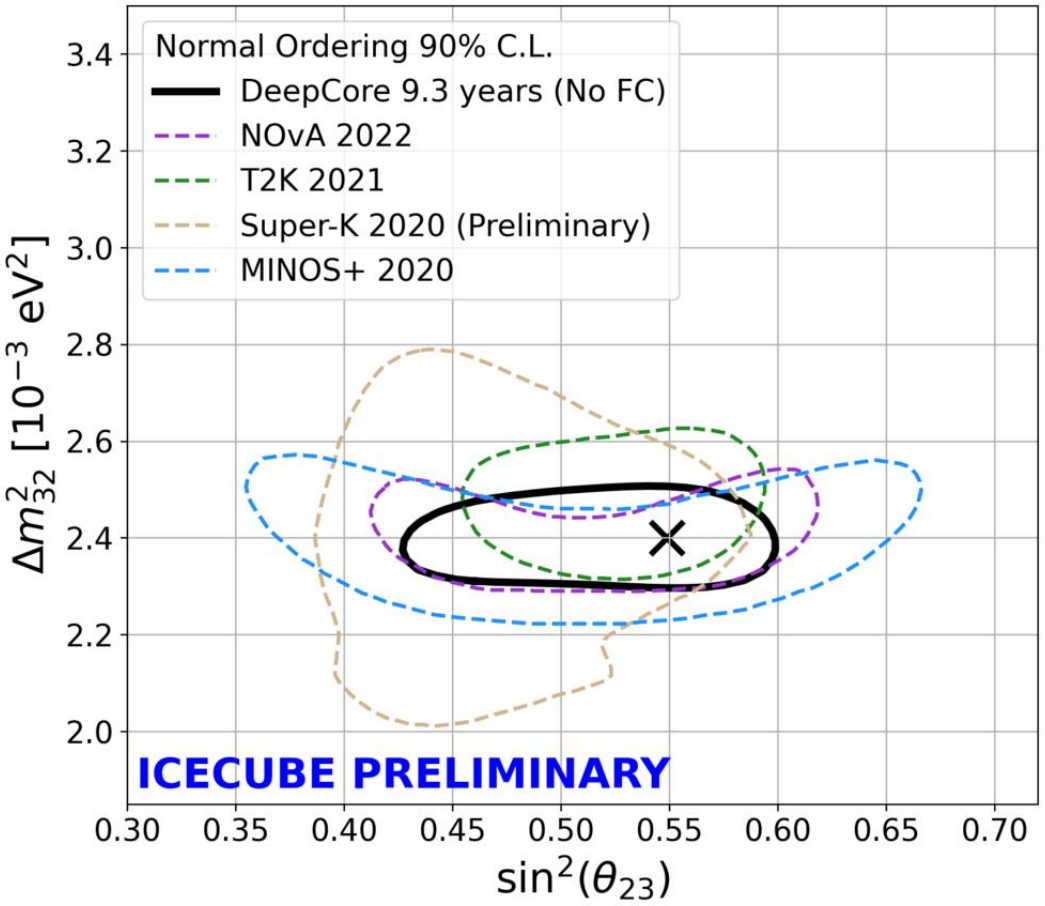


Observed Spectrum

- Transfer from truth to observed variables:
  - Either achieved via smearing functions (convolution), for example works rel. well for reactor neutrino experiments
  - Or via transfer kernels obtained via MC: e.g. LBL can be done via fine 2d histograms

- For atmospherics, this would become a 5-6d histogram, which is becoming infeasible
- Alternative is to do an event-by-event reweighting of the MC after each change of a (nuisance) parameter

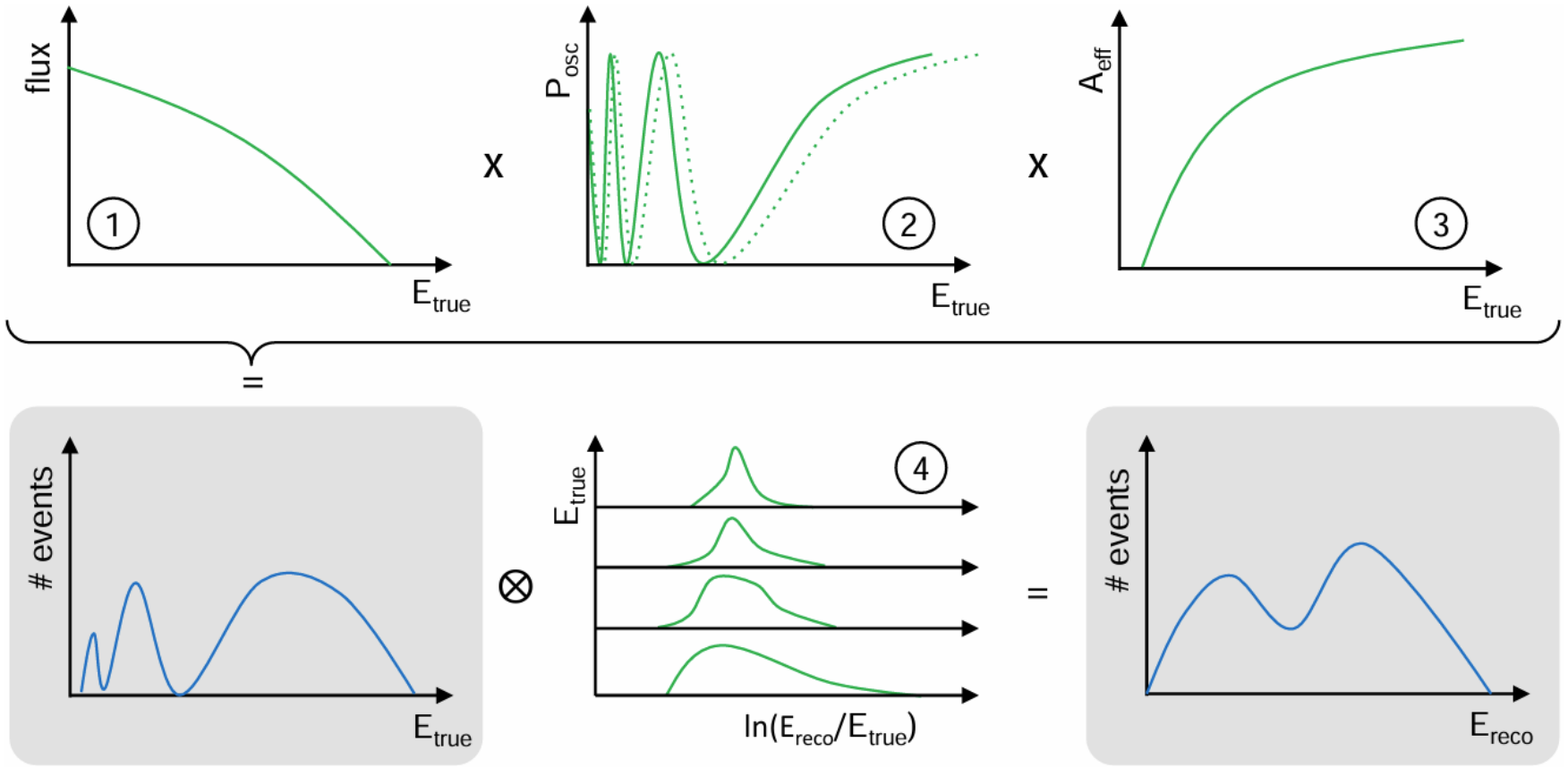
Example Result



- 2 measurement parameters
  - Here 2d C.L. contour
- 20-30 nuisance parameters
- Some parameters may be correlated in global fits with other experiments
  - E.g. Atm. Flux model, Reactor Model, …

Neutrino Oscillation Model

- Computation of likelihood involves:
  1. Computation of un-oscillated spectrum / flux
     - Should allow for external calculations and correlations across experiments
  2. Evaluation of oscillation probabilities
     - Should allow for external calculations
  3. Transfer from true neutrino properties to observables
     - May require costly MC reweighting / convolution
  4. Vanilla binned Poisson Likelihood
     - Easy



https://arxiv.org/pdf/1803.05390 "Computational Techniques for the Analysis of Small Signals in High-Statistics Neutrino Oscillation Experiments"

HS^3

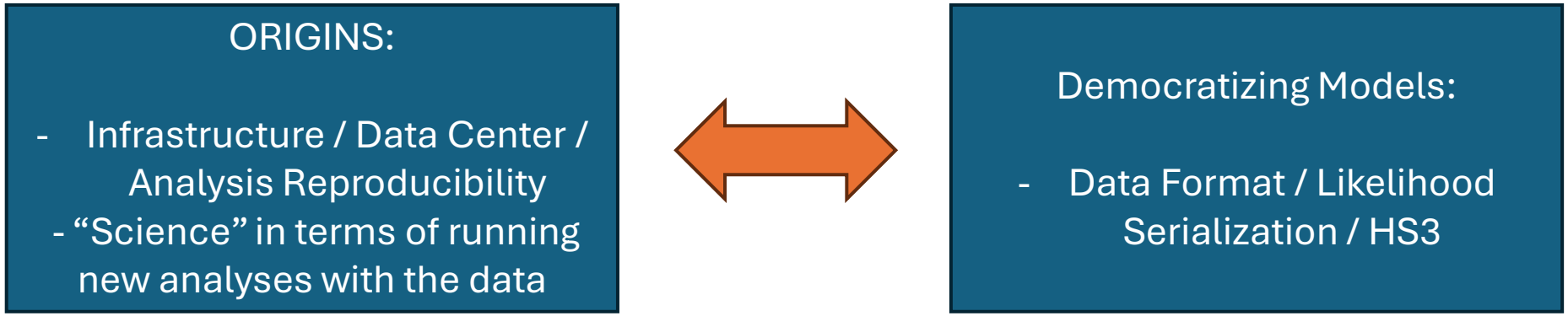
"how well the HS3-like json representation work for your type of the models?"

→ This is a question I don't have a good answer for, but maybe you have…?

My hunch is, that of course it could be used as a data exchange format, but to provide an actual likelihood it would require quite some work.

Summary

- ORIGINS and ORIGINS2
  - Will provide research data infrastructure for DM and neutrino data
  - But neither, DM nor Neutrino, have ANY common way of sharing data or models
  - This is where I would see some potential connection to “Democratizing Models”



- Hosting data / providing platforms could likely be viewed as double funding