Atmospheric Neutrino Unfolding

Joakim Sandroos, J. G.U. Mainz

for the IceCube collaboration

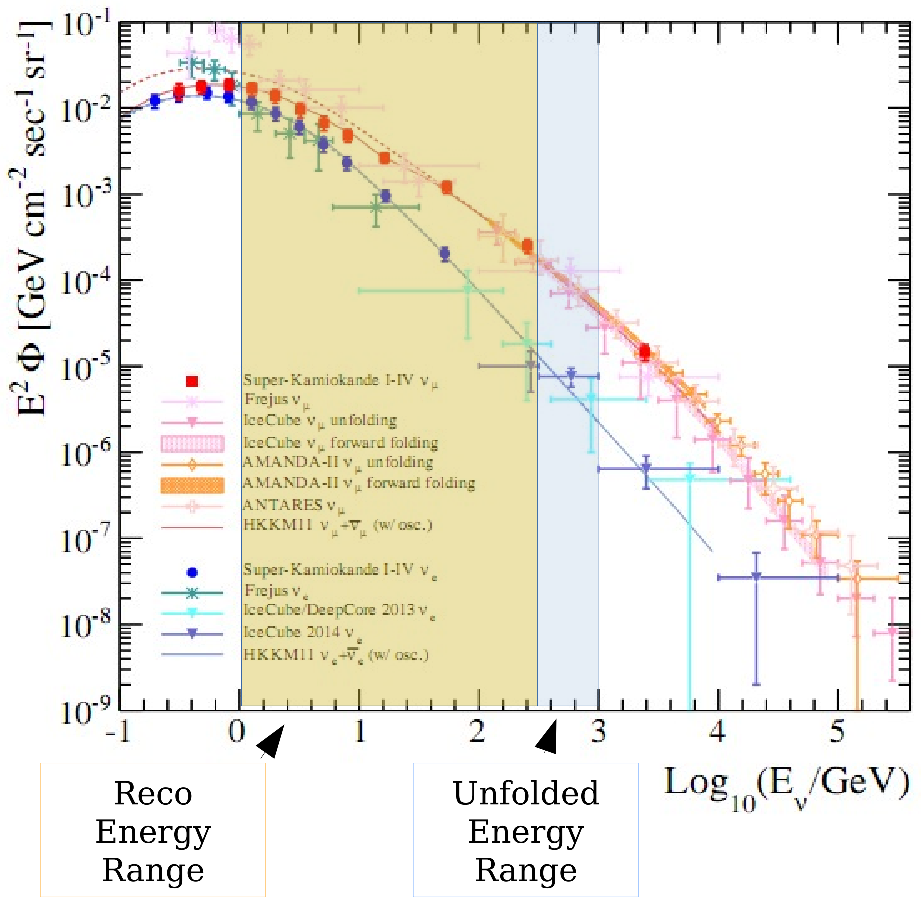
Motivation

- Goal: Provide model agnostic 'true' data sample for model builders to test against

- Goal: Accurate 'beam' measurement for any experiment depending on atmospheric neutrinos

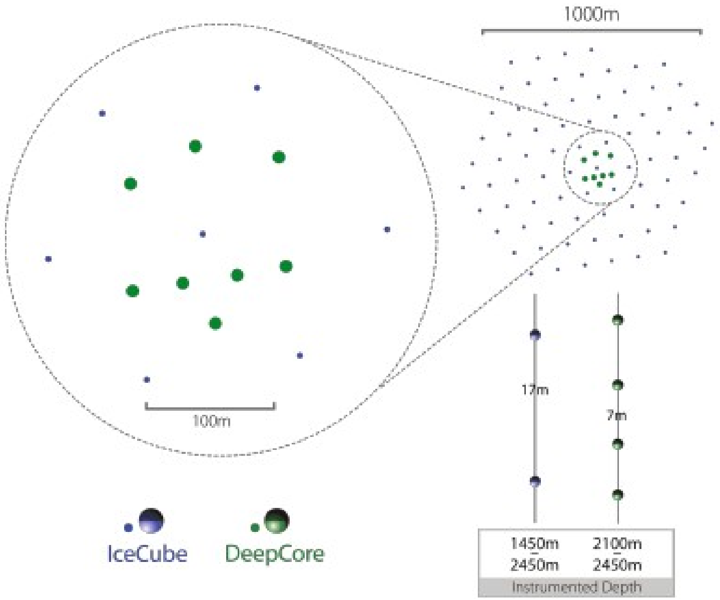
- Means: Model agnostic Unfolding

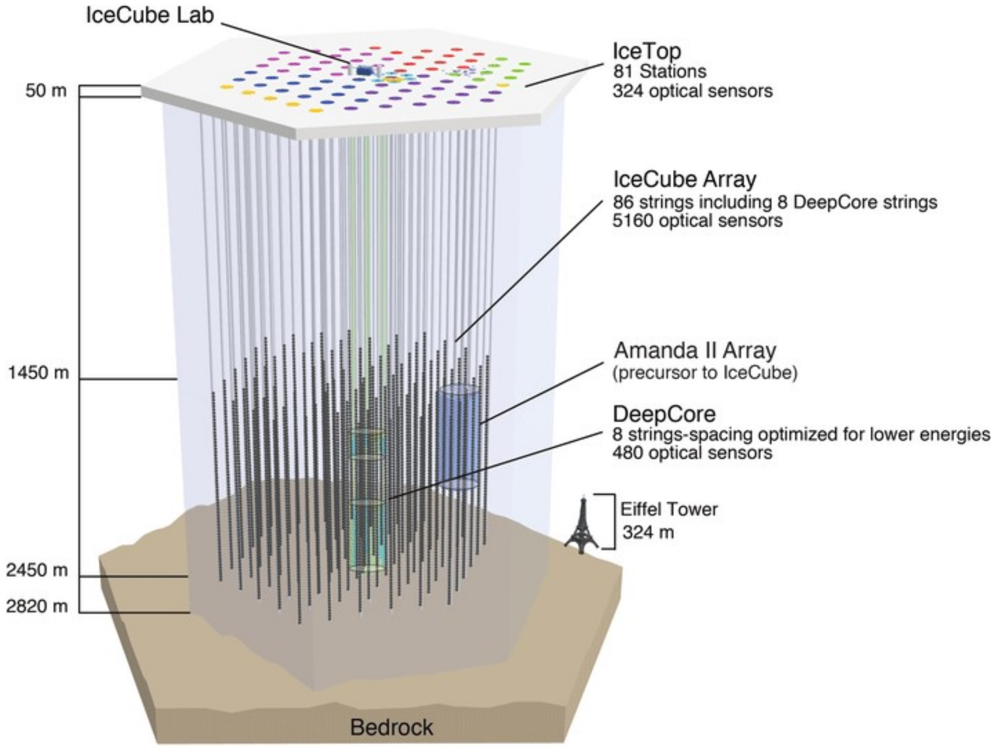
- Blind: Analysis tested on MC and 10% data sample.



IceCube / DeepCore

- Antarctica
- 1km³ instrumented ice
- Cherenkov radiation
- 5160 optical modules
- Outer detector as muon veto

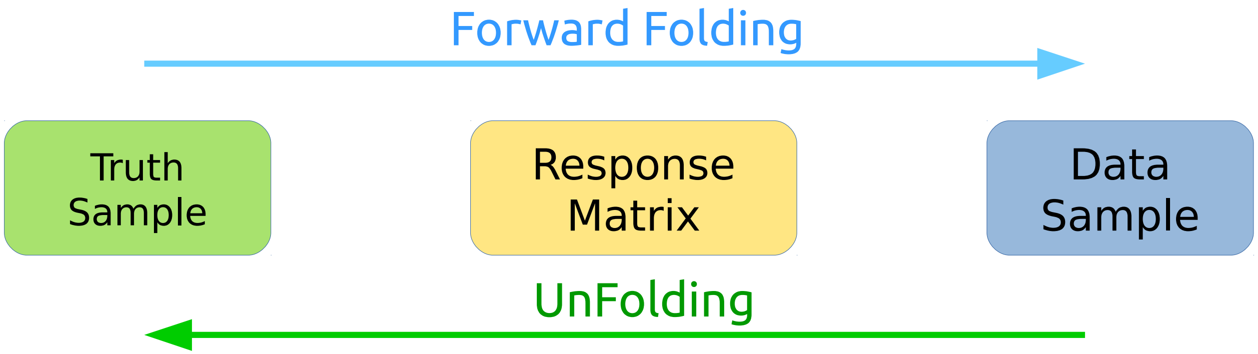




Unfolding Philosophy

JG|U

- Both forward- and unfolding depend on the detector response matrix
- In this analysis we build the response matrix by comparing MC truth to reconstructed MC



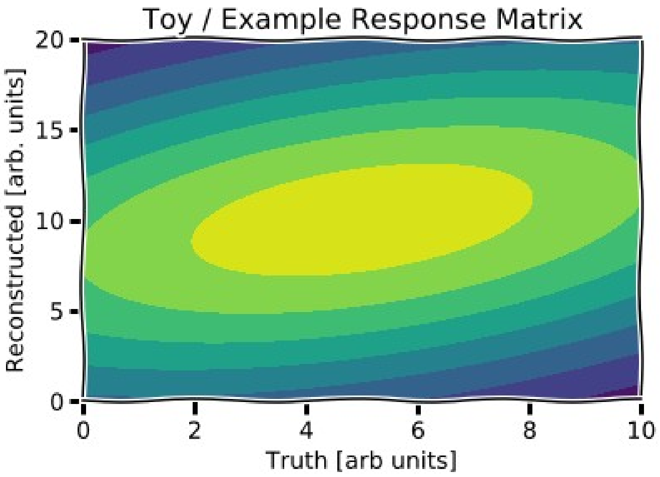
- Unfolding provides not model parameters, but physical quantities
- Unfolding constrains data via parameters

  → Unlike forward folding we cannot fit values of parameters
- Based on our systematic uncertainties, we constrain our unfolded bin content

Unfolding Method

- Bayesian Unfolding (D'Agostini)

$$P_0(i|j)=\frac{P(j|i)\times P_0(i)}{P_0(j)}$$



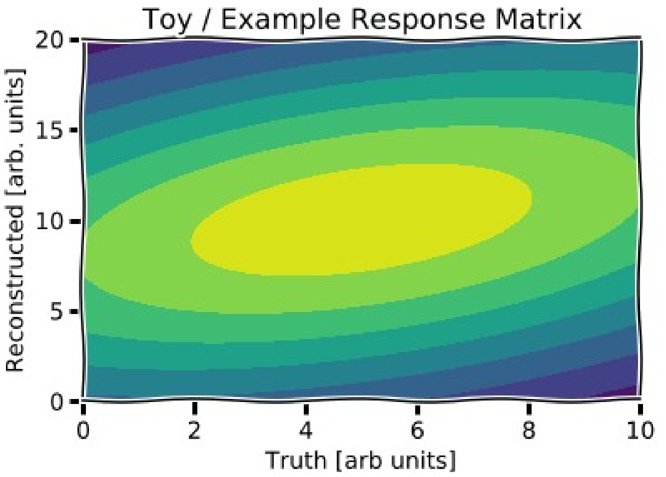
- Prior information about our measurement: If we draw a sample from the MC truth distribution, what is the probability to be draw from bin i?
- Normalization: Probability to observe an event in bin j. Given by the logic of requiring an event in true bin i **and** for that event to contribute to bin j, and summing for all bins. (This is why it's fair to call it a normalization constant)
- All terms known to precision of MC

i = truth bin
j = reco. bin

Unfolding Method

- Bayesian Unfolding (D'Agostini)

$$P_0(i|j) = \frac{\boxed{P(j|i)} \times P_0(i)}{P_0(j)}$$



- Prior information about our measurement: If we draw a sample from the MC truth distribution, what is the probability to be draw from bin i?
- Normalization: Probability to observe an event in bin j. Given by the logic of requiring an event in true bin i **and** for that event to contribute to bin j, and summing for all bins. (This is why it's fair to call it a normalization constant)
- All terms known to precision of MC

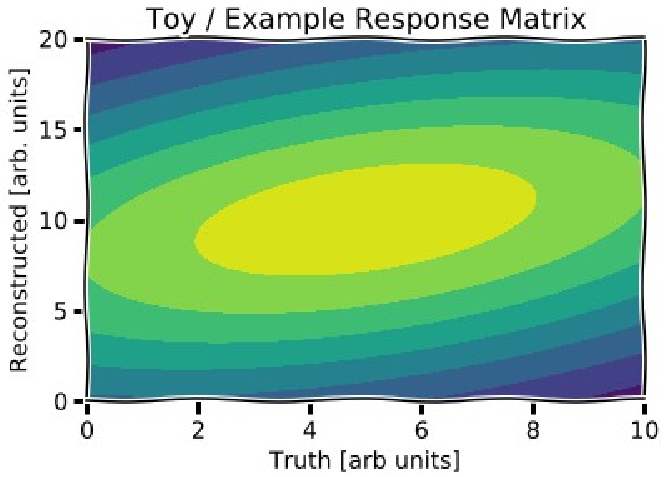
i = truth bin
j = reco. bin

Unfolding Method

- Bayesian Unfolding (D'Agostini)

$$P_0(i|j) = \frac{P(j|i) \times \boxed{P_0(i)}}{P_0(j)}$$

$$\boxed{P_0(i) = \frac{N_i^{MC}}{\sum_k N_k^{MC}}}$$



- Prior information about our measurement: If we draw a sample from the MC truth distribution, what is the probability to be draw from bin i?
- Normalization: Probability to observe an event in bin j. Given by the logic of requiring an event in true bin i **and** for that event to contribute to bin j, and summing for all bins. (This is why it's fair to call it a normalization constant)
- All terms known to precision of MC

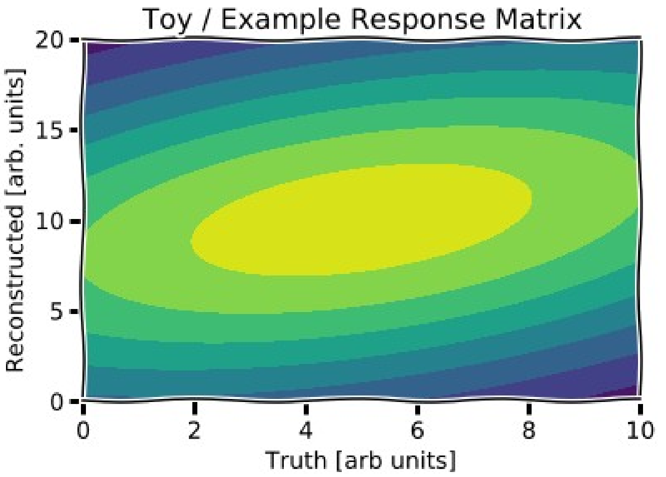
i = truth bin
j = reco. bin

Unfolding Method

- Bayesian Unfolding (D'Agostini)

$$P_0(i|j)=\frac{P(j|i)\times P_0(i)}{\boxed{P_0(j)}}$$

$$\boxed{P_0(j)=\sum_i [P(j|i)\times P(i)]}$$



- Prior information about our measurement: If we draw a sample from the MC truth distribution, what is the probability to be draw from bin i?
- Normalization: Probability to observe an event in bin j. Given by the logic of requiring an event in true bin i **and** for that event to contribute to bin j, and summing for all bins. (This is why it's fair to call it a normalization constant)
- All terms known to precision of MC

i = truth bin
j = reco. bin

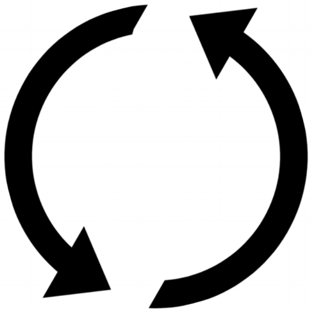
Unfolding Method: Iterative Approach

- Generates unfolding matrix via Bayes' theorem

$$P_0(i|j) = \frac{P(j|i) \times P_0(i)}{P_0(j)}$$

- Estimates unfolded spectrum U, from measurement M:

$$U_n = P_n(i|j) \times M$$

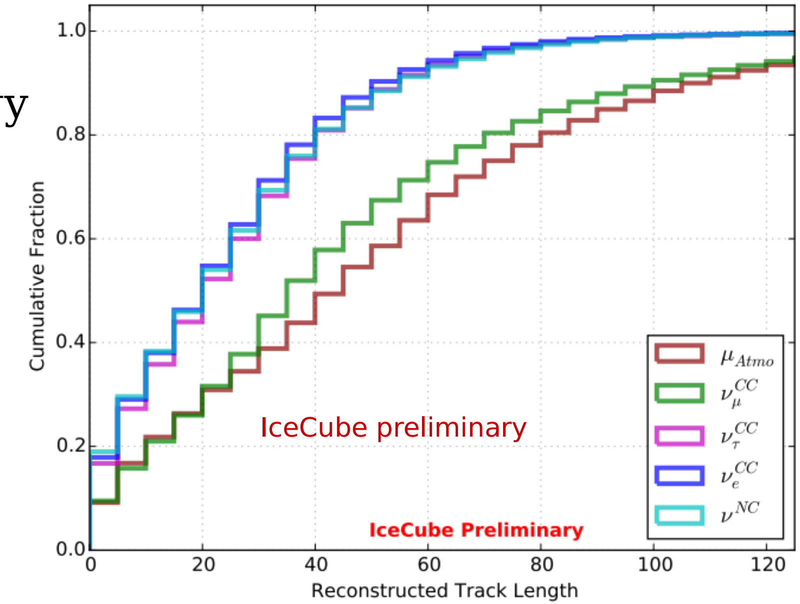


- Output of each step is prior for next step.

- Biased towards MC for low iterations → Bias drops with iterations
- Statistical Uncertainty → Grows with number of iterations
- Final number of iterations must be a trade-off between the above two

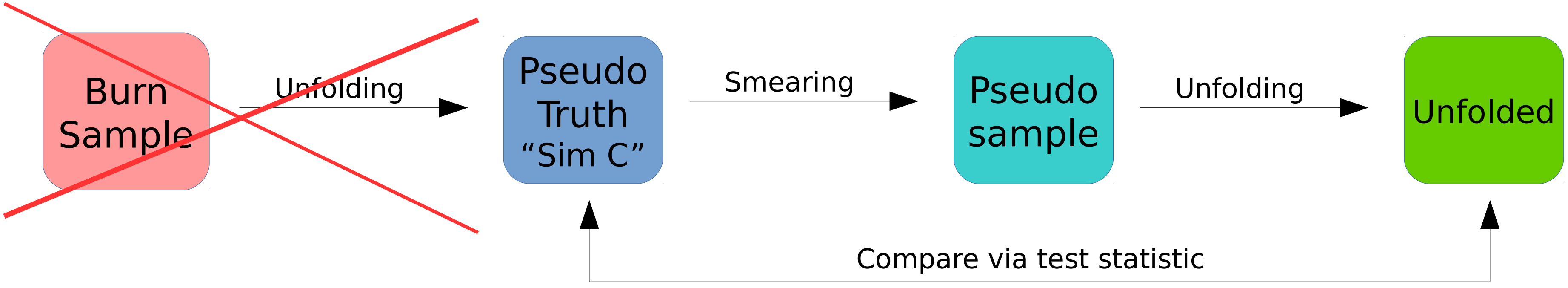
Response Matrix and Channels

- Response matrix constructed from MC
- Reco-side: reconstructed $\cos(\theta_z)$, reconstructed energy and reconstructed track length
- Truth-side: MC PID, MC truth $\cos(\theta_z)$ and MC truth energy
- PID channels: $\boxed{\nu_\mu^{cc} + \bar{\nu}_\mu^{cc} \quad \text{and} \quad \nu_{rest} + \bar{\nu}_{rest}}$
- Weight truth side by: $\left(t_{live} V_{IC}\right)^{-1}$
- Unfolded quantity is “True In-ice interaction rate per volume [/m³ /s]”
- From these distributions the energy and zenith spectra are calculated



TSU: A Blind Burn Sample Closoure Test

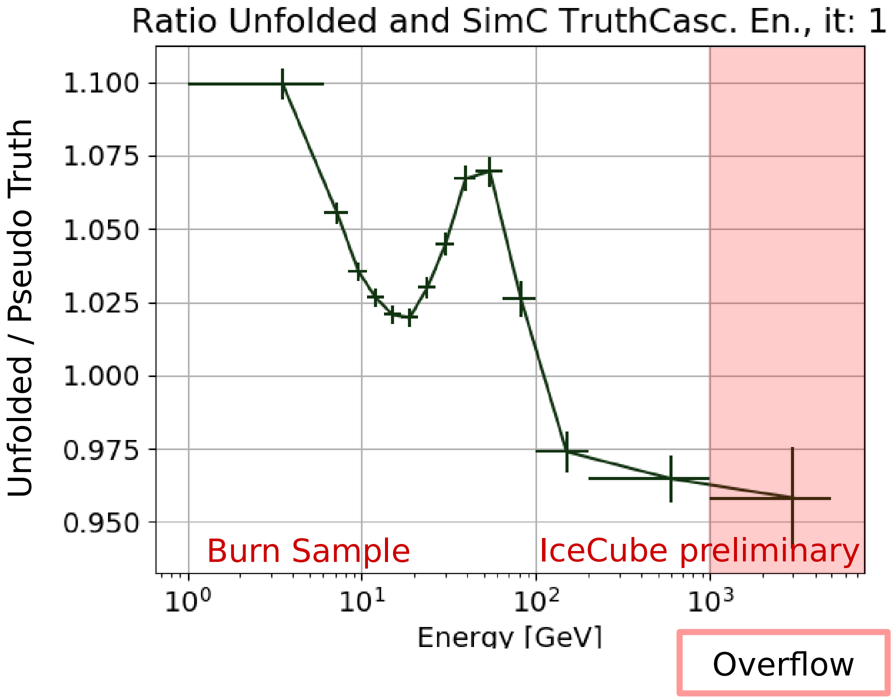
- Problem: When unfolding a real data sample we do not have access to truth information like in the MC case
- Aim: Show stability of unfolding method across smearing and unfolding
- Closure test: Truth-Smeared-Unfolded (TSU) test

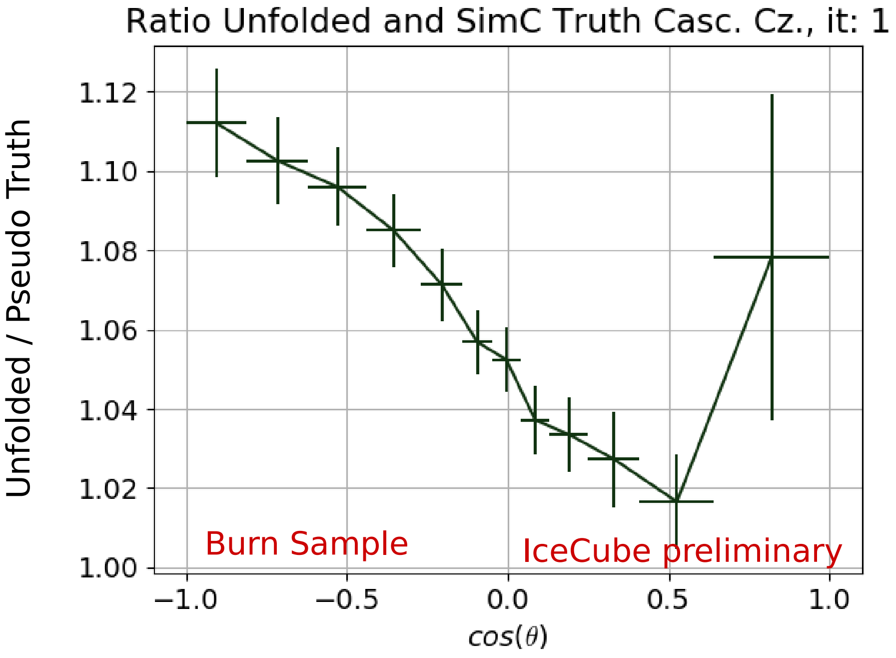


- Unfold to 25 iterations
- Result: Converges on pseudo-truth to well within statistical uncertainty
- Careful consideration of stopping condition is necessary

TSU: Burn Sample Consistency

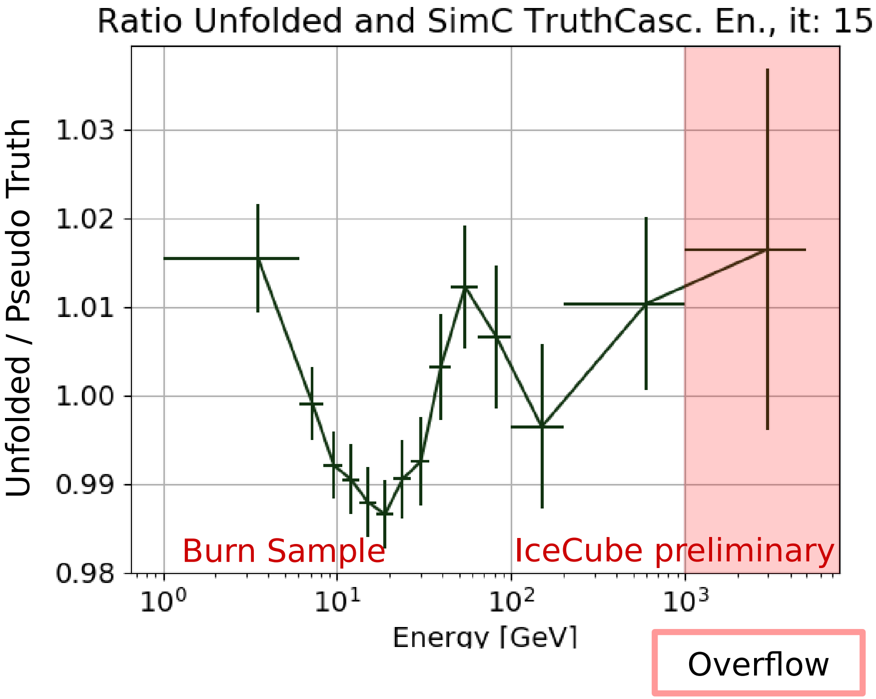
- Blind Check: TSU-Ratios, 1 iteration,
- Checks consistency in unfolding

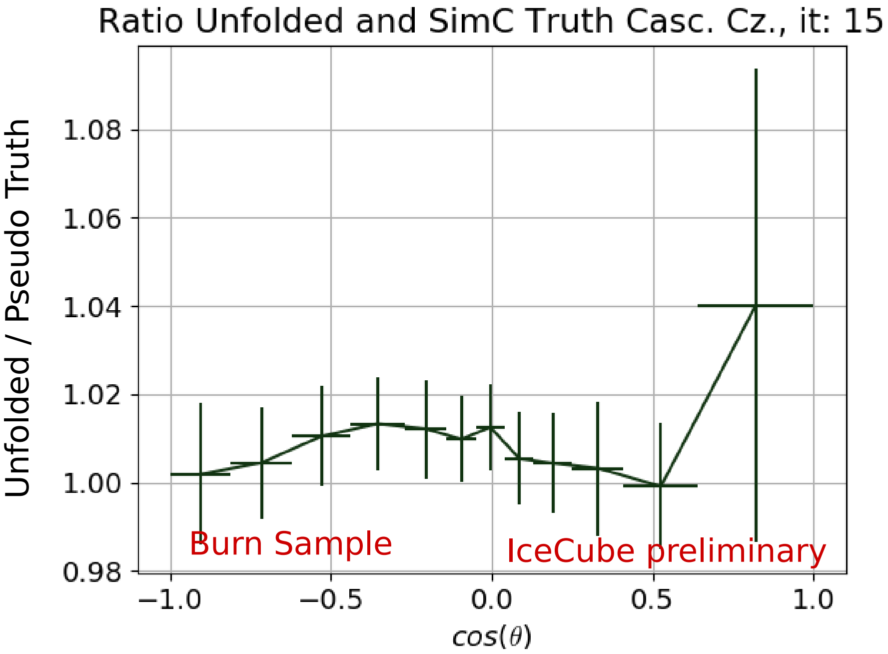




TSU: Burn Sample Consistency

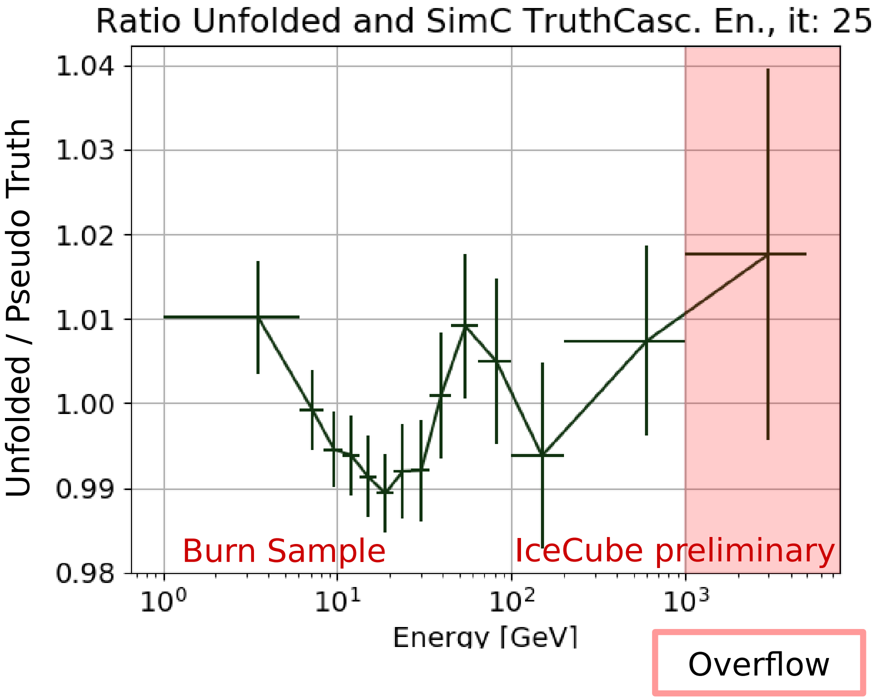
- Blind Check: TSU-Ratios, 15 iterations,
- Reasonable consistency

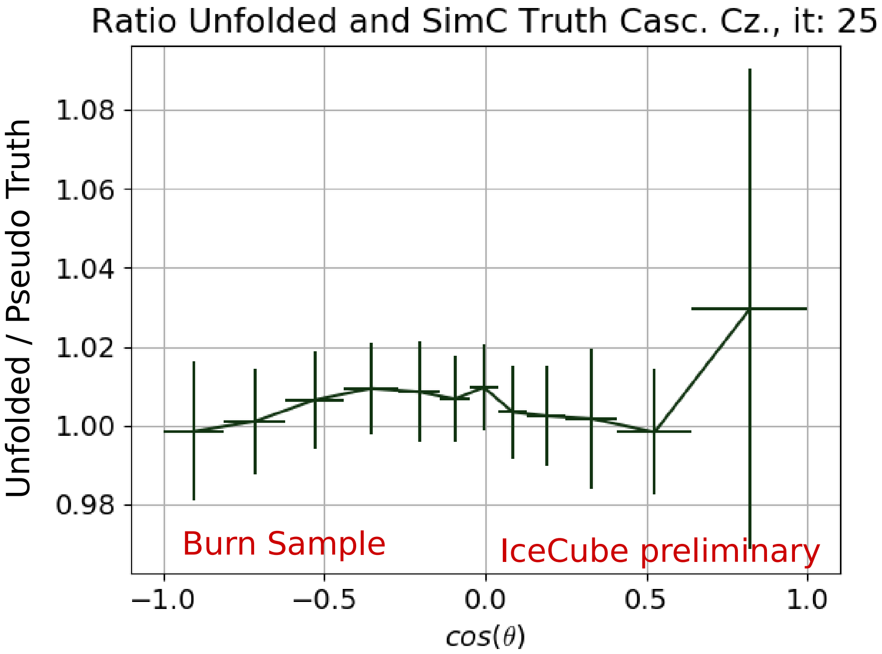




TSU: Burn Sample Consistency

- Blind Check: TSU-Ratios, 25 iterations,
- Reasonable Consistency





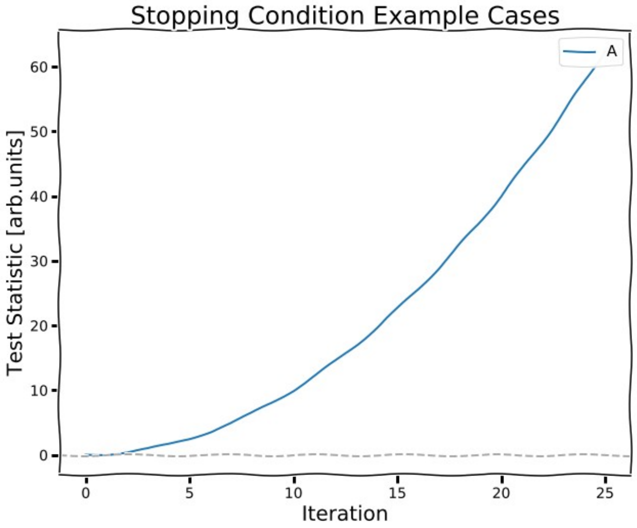
Setting the Stopping condition

- Two test statistics:
- Statistics only:

$$TS = \frac{1}{UM} \sum_i \frac{\left(m_i U - u_i M\right)^2}{m_i + u_i}$$

- Uncertainty based:

$$TS_\sigma = \frac{1}{N} \sum_i^N \left(\frac{\sigma_i^{max} - \sigma_i^{min}}{n_i} \right)^2$$



- The stopping condition plots consist of calculating a test statistic for every iteration between unfolded and pseudo truth – can take many different shapes.

Hierarchy of procedure:

- In case of divergence: Stop after 4 iterations, as advised by D'Agostini (A)

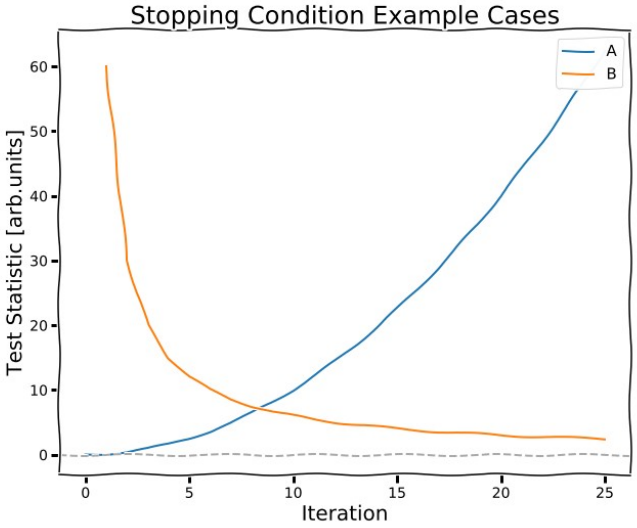
Setting the Stopping condition

- Two test statistics:
- Statistics only:

$$TS = \frac{1}{UM} \sum_i \frac{\left(m_i U - u_i M\right)^2}{m_i + u_i}$$

- Uncertainty based:

$$TS_\sigma = \frac{1}{N} \sum_i^N \left(\frac{\sigma_i^{max} - \sigma_i^{min}}{n_i} \right)^2$$



- The stopping condition plots consist of calculating a test statistic for every iteration between unfolded and pseudo truth – can take many different shapes.

Hierarchy of procedure:

- In case of divergence: Stop after 4 iterations, as advised by D'Agostini (A)
- In case of convergence: Use systematics based stopping condition. (B)

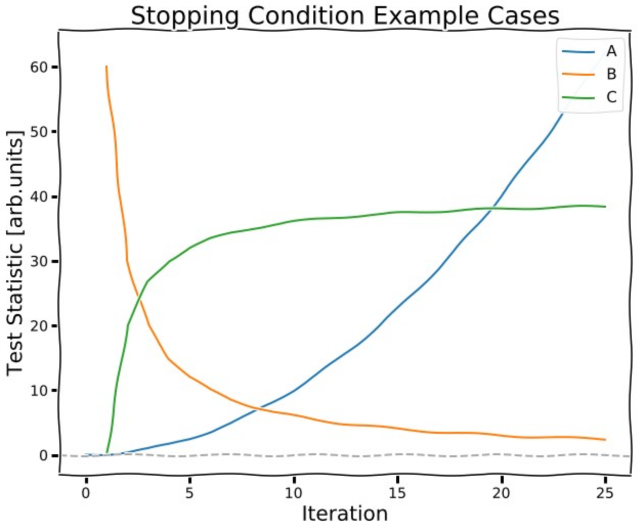
Setting the Stopping condition

- Two test statistics:
- Statistics only:

$$TS = \frac{1}{UM} \sum_i \frac{\left(m_i U - u_i M\right)^2}{m_i + u_i}$$

- Uncertainty based:

$$TS_\sigma = \frac{1}{N} \sum_i^N \left(\frac{\sigma_i^{max} - \sigma_i^{min}}{n_i} \right)^2$$



- The stopping condition plots consist of calculating a test statistic for every iteration between unfolded and pseudo truth – can take many different shapes.

Hierarchy of procedure:

- In case of divergence: Stop after 4 iterations, as advised by D'Agostini (A)
- In case of convergence: Use systematics based stopping condition. (B)
- In case of systematic dominance after 1 iteration: Use statistics only stopping condition. (C)

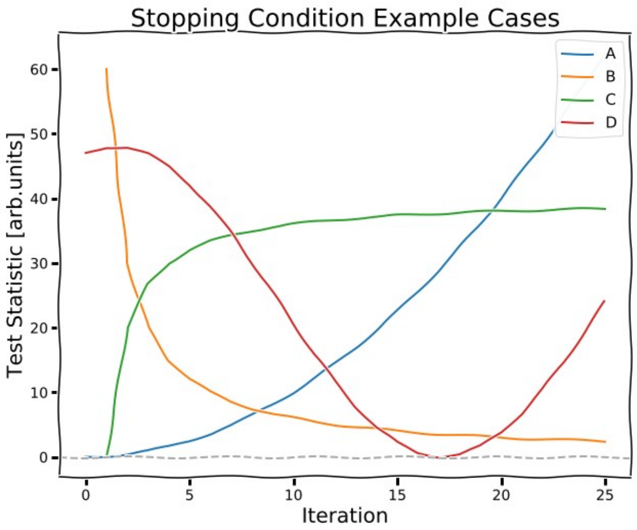
Setting the Stopping condition

- Two test statistics:
- Statistics only:

$$TS = \frac{1}{UM} \sum_i \frac{(m_i U - u_i M)^2}{m_i + u_i}$$

- Uncertainty based:

$$TS_\sigma = \frac{1}{N} \sum_i^N \left(\frac{\sigma_i^{max} - \sigma_i^{min}}{n_i} \right)^2$$



- The stopping condition plots consist of calculating a test statistic for every iteration between unfolded and pseudo truth – can take many different shapes.

Hierarchy of procedure:

- In case of divergence: Stop after 4 iterations, as advised by D'Agostini (A)
- In case of convergence: Use systematics based stopping condition. (Minimize distribution) (B)
- In case of systematic dominance after 1 iteration: Use statistics only stopping condition (C)
- Otherwise: Minimize the distribution (D)

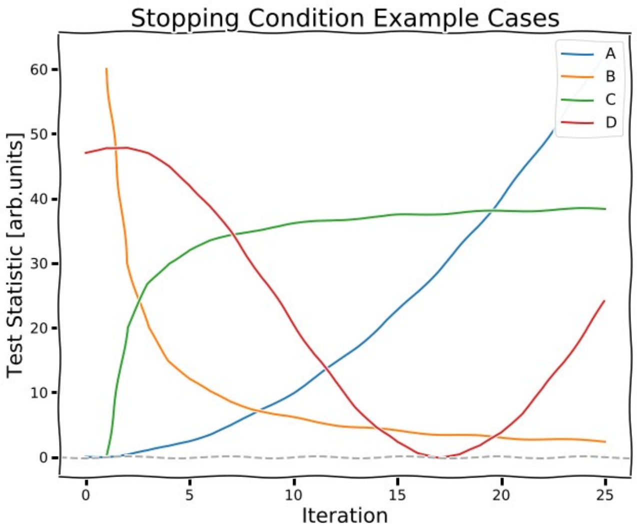
Setting the Stopping condition

- Two test statistics:
- Statistics only:

$$TS = \frac{1}{UM} \sum_i \frac{\left(m_i U - u_i M\right)^2}{m_i + u_i}$$

- Uncertainty based:

$$TS_\sigma = \frac{1}{N} \sum_i^N \left(\frac{\sigma_i^{max} - \sigma_i^{min}}{n_i} \right)^2$$



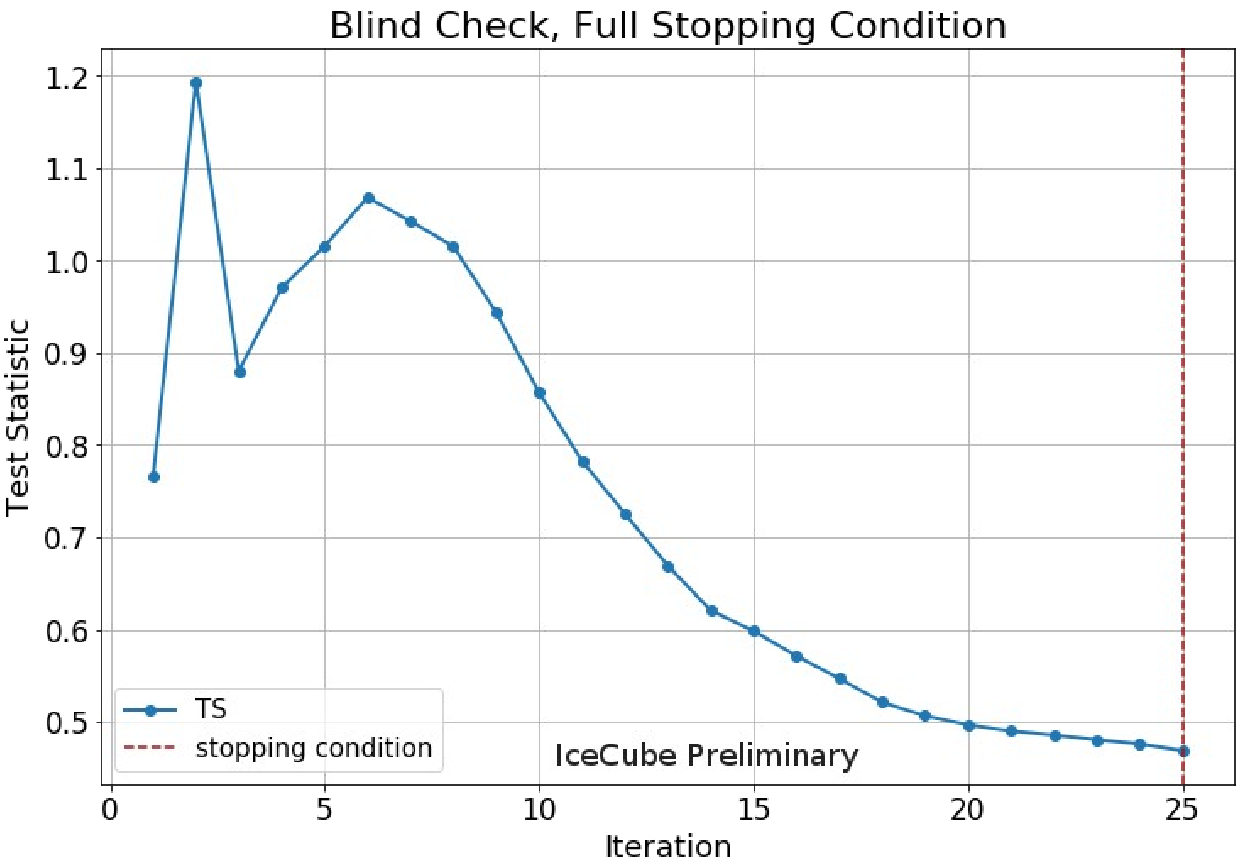
- The stopping condition plots consist of calculating a test statistic for every iteration between unfolded and pseudo truth – can take many different shapes.

Hierarchy of procedure:

- In case of divergence: Stop after 4 iterations, as advised by D'Agostini (A)
- In case of convergence: Use systematics based stopping condition. (Minimize distribution) (B)
- In case of systematic dominance after 1 iteration: Use statistics only stopping condition (C)
- Otherwise: Minimize the distribution (D)
- Burn sample test case: D, minimum at 20 iterations

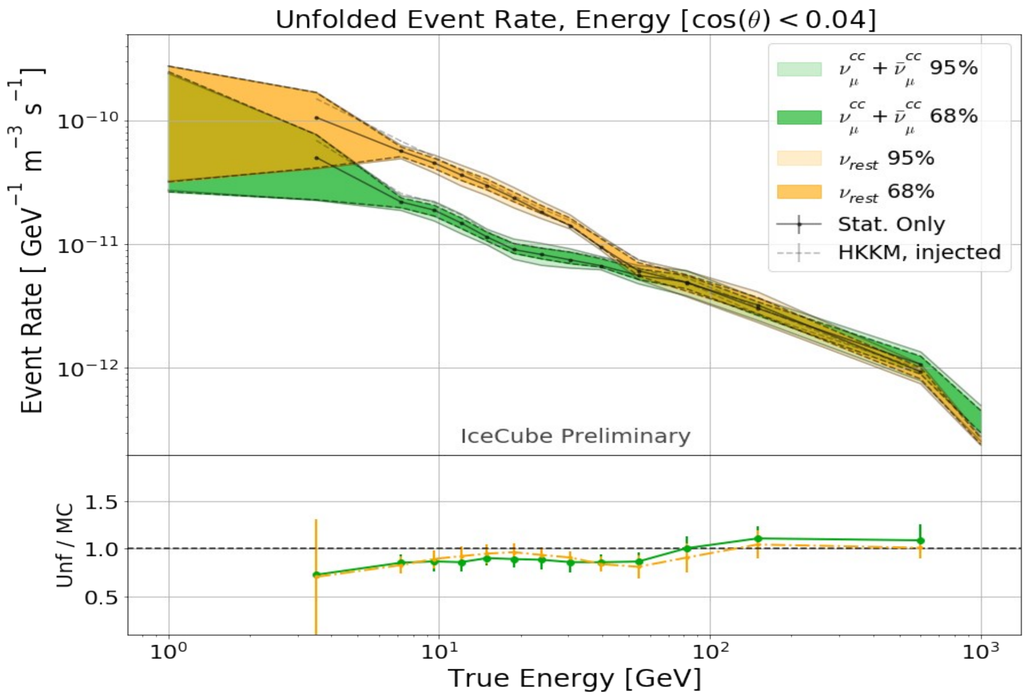
Stopping condition

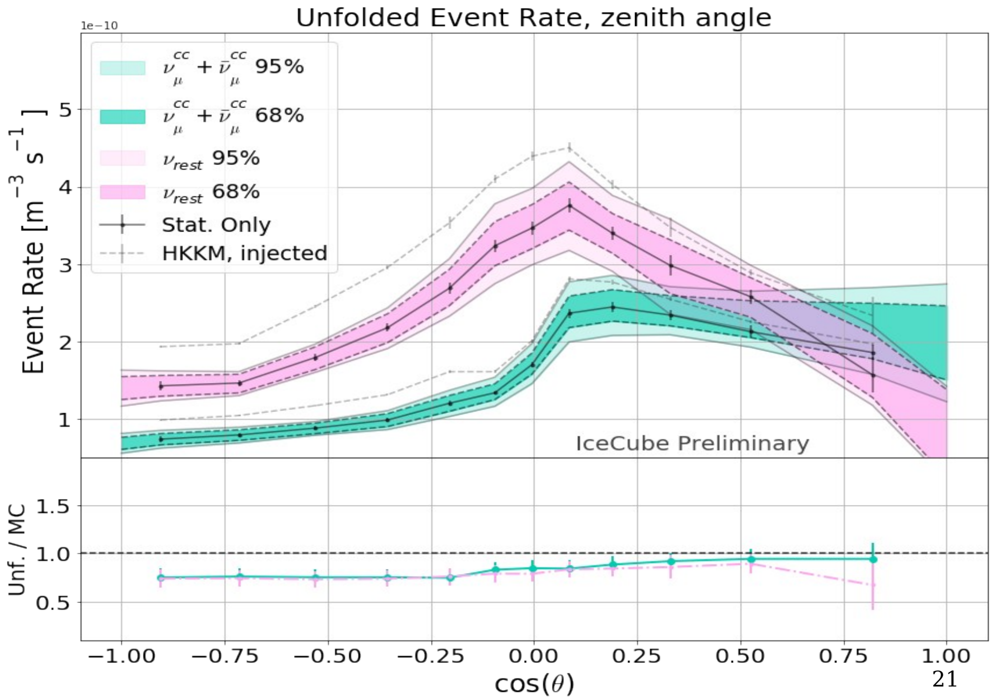
- For full data sample
- Only small variation above ~20 iterations.



Unfolded Event Rate

- 2 channels based on idealized event signature in detector
- $\nu_\mu^{cc} + \bar{\nu}_\mu^{cc}$
- Everything else: ν_{rest}
- ~1400 trials





Closing Remarks

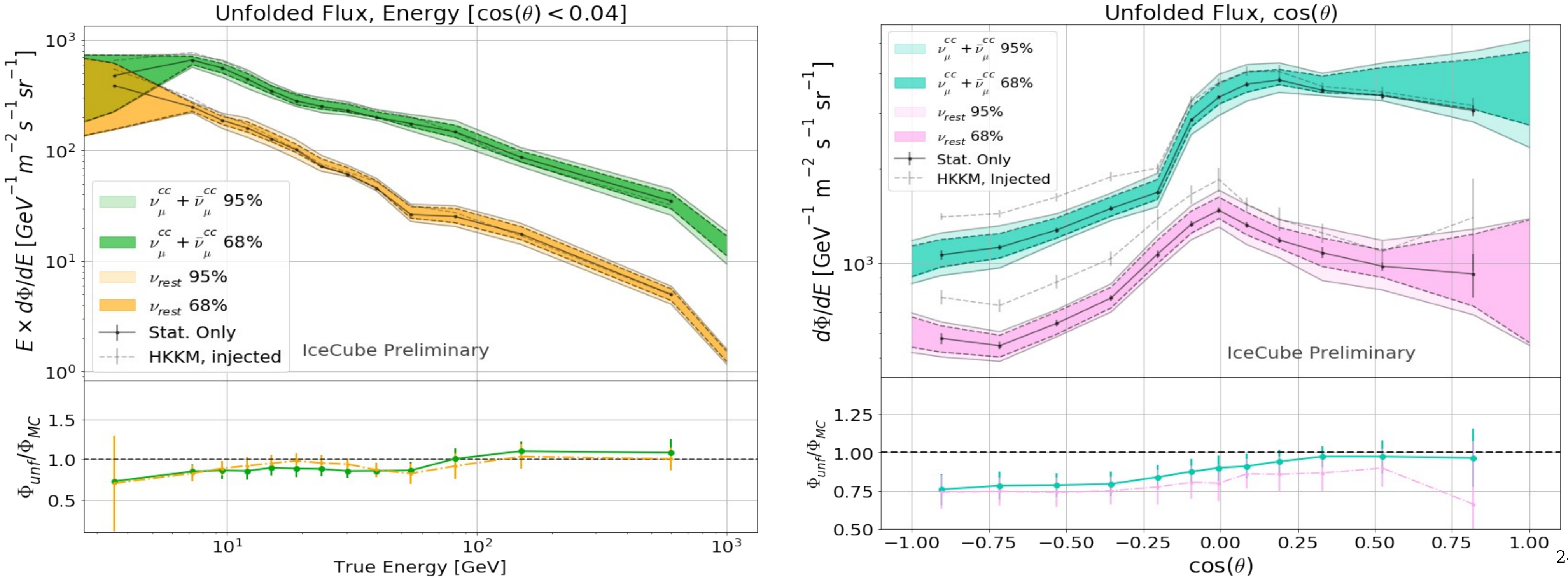
- Unfolded measurement of atmospheric neutrino flux at south pole
- Allows model builders to test predictions on many parameters
- Some tension with expectation below 10 GeV and in up going region
- Data release and publication in preparation

Backup

- From here

Addendum: Unfolded Flux

- Same 2 channels, now also compensated for cross sections
- $\nu_\mu^{cc} + \bar{\nu}_\mu^{cc}$
- Everything else: ν_{rest}
- ~ 75 trials



Systematics

- Random sampling of systematics due to non-linear effects
- ~1400 trials

Oscillation and Weighting Systematics		
Systematic	Value	Prior
θ_{12}	34.5^{o}	$\pm 1.1^{o}$ (1)
θ_{23}	41^{o}	$\pm 0.11^{o}$ (1)
θ_{13}	8.41^{o}	$\pm 0.17^{o}$ (1)
Δm_{21}	$7.56\ e^{-5}\ eV^{2}$	$\pm 0.19\ e^{-5}\ eV^{2}$ (1)
Δm_{31}	$2.55\ e^{-3}\ eV^{2}$	$\pm 0.04\ e^{-3}\ eV$ (1)
δ_{cp}	252^{o}	$\pm 24^{o}$ (1)
Livetime	4.8 [yr]	1%
Muon Scale	1.0	5%
Noise	1.0	10%

1: Salas et al, arXiv: 1708.01186

Discrete Systematics		
Systematic	Value	Prior
Dom eff	1.0	10%
Hole ice	25	±5
Bulk ice scattering	1.0	10%
Bulk ice absorbtion	1.0	10%

3: IceCube Standard

Analysis Chain:

Sample

Osc. re-weighting

Discrete sys.

Unfolding

Efficiency

- Definiton:

$$\epsilon = \frac{N_{sel}}{N_{gen}} = \frac{R_{sel} V_{fidu}}{R_{gen} V_{gen}}$$

- With the number of efficiency corrected unfolded events:

$$N_{unf}^{\epsilon} = \frac{N_{unf}}{\epsilon} = \frac{R_{unf} V_{fidu}}{\left(\frac{R_{sel} V_{fidu}}{R_{gen} V_{gen}}\right)} = \frac{R_{unf} R_{gen} V_{gen}}{R_{sel}}$$

- To get a rate independent from the volume, it is prudent to modify the efficiency; we divide out the generator volume to arrive at:

$$\epsilon' = \frac{N_{sel}}{N_{gen} / V_{gen}} = \epsilon V_{gen} \quad \longrightarrow \quad R_{unf}^{\epsilon'} = \frac{N_{unf}}{\epsilon V_{gen}} = \frac{R_{unf} R_{gen}}{R_{sel}}$$