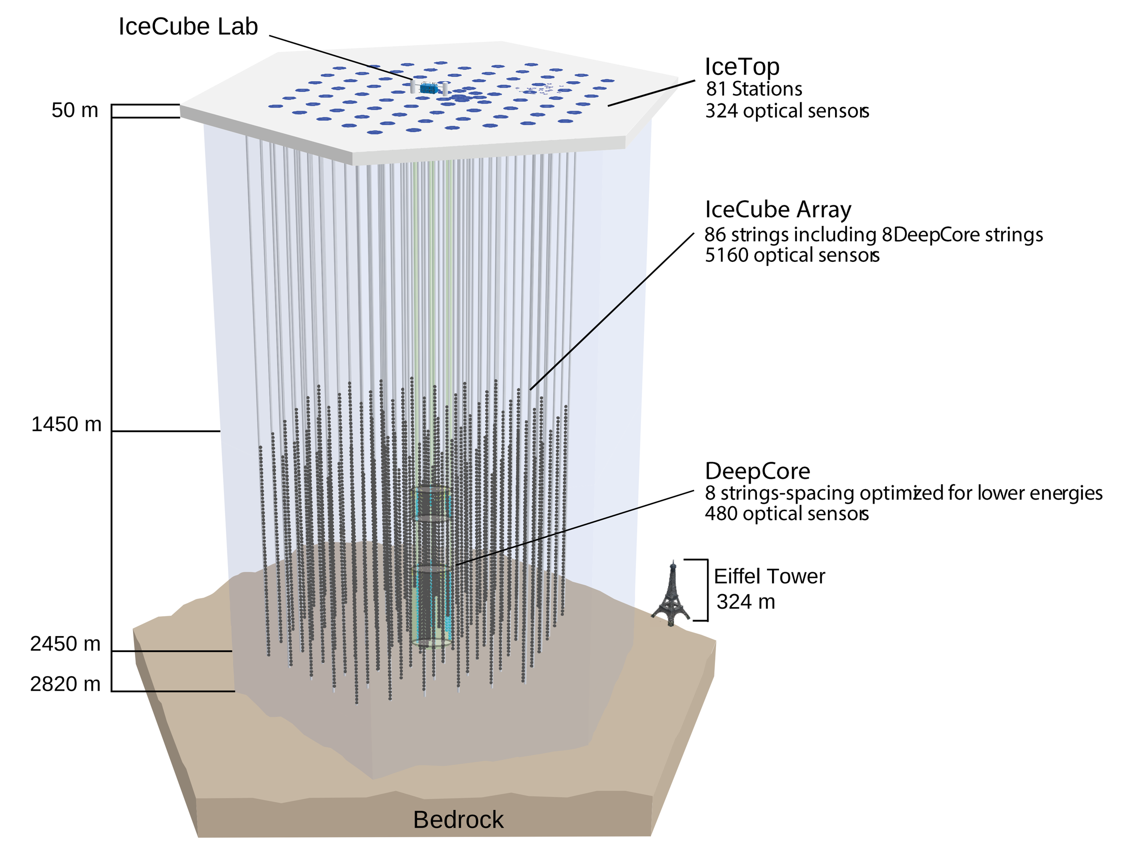
High-statistics and GPU-Accelerated Data Analysis

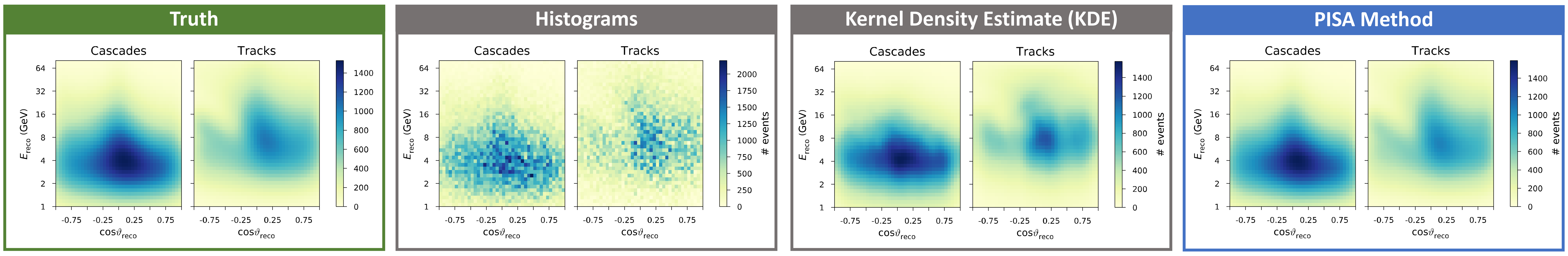
PennState

ICECUBE
SOUTH POLE NEUTRINO OBSERVATORY



IceCube is a gigaton Ice Cherenkov detector located at the South Pole

- Its large volume together + high intensity atmospheric neutrino flux lets us collect unprecedented amounts of data
- Typically one needs ~O(10) more statistics of simulated events to analyze data
- Increasingly difficult for future experiments (PINGU, ORCA, …) anticipated to collect even more data
- Generating distributions from MC needs to be very fast - after every model parameter change distributions have to be regenerated



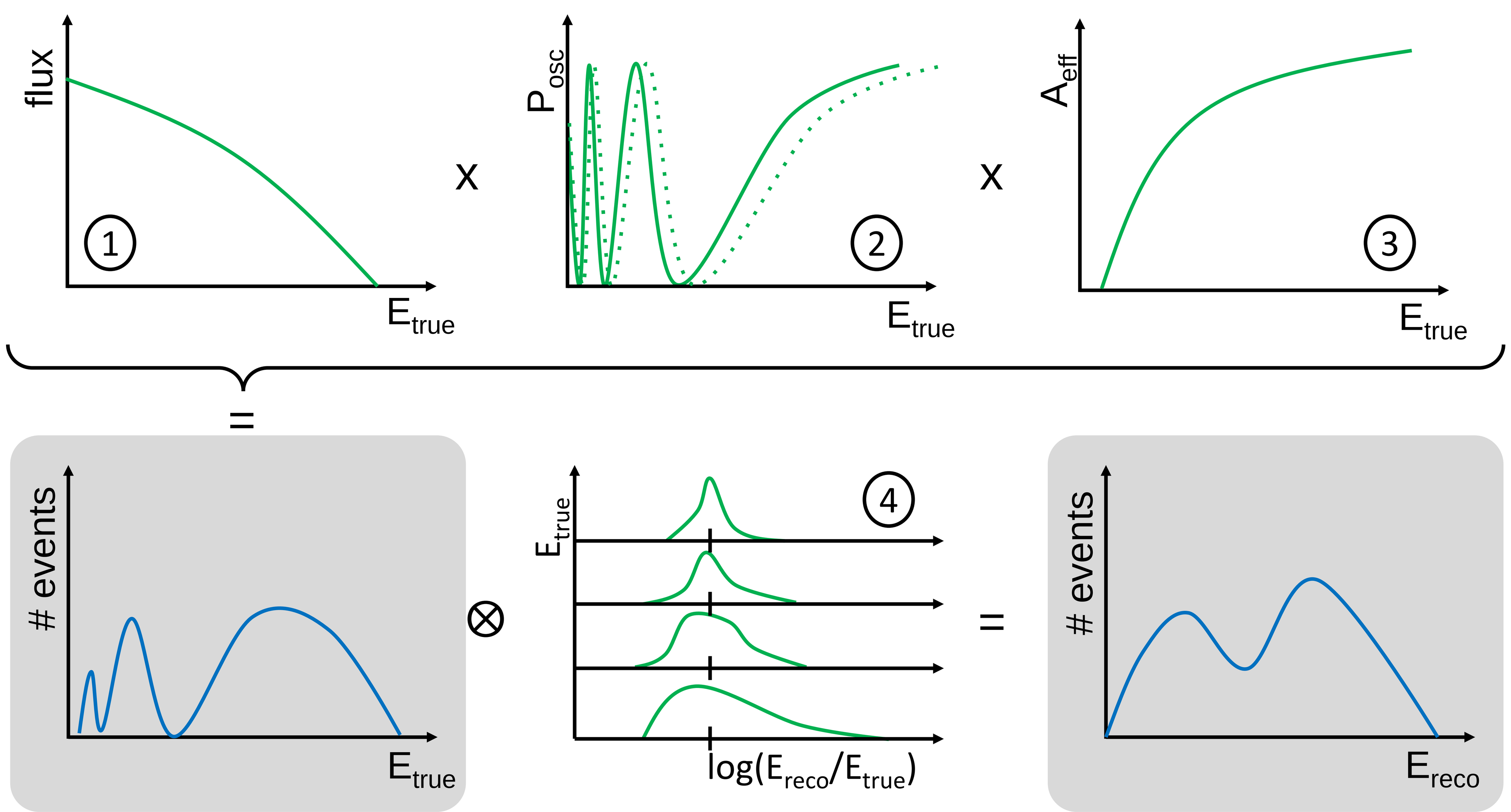
PISA Method

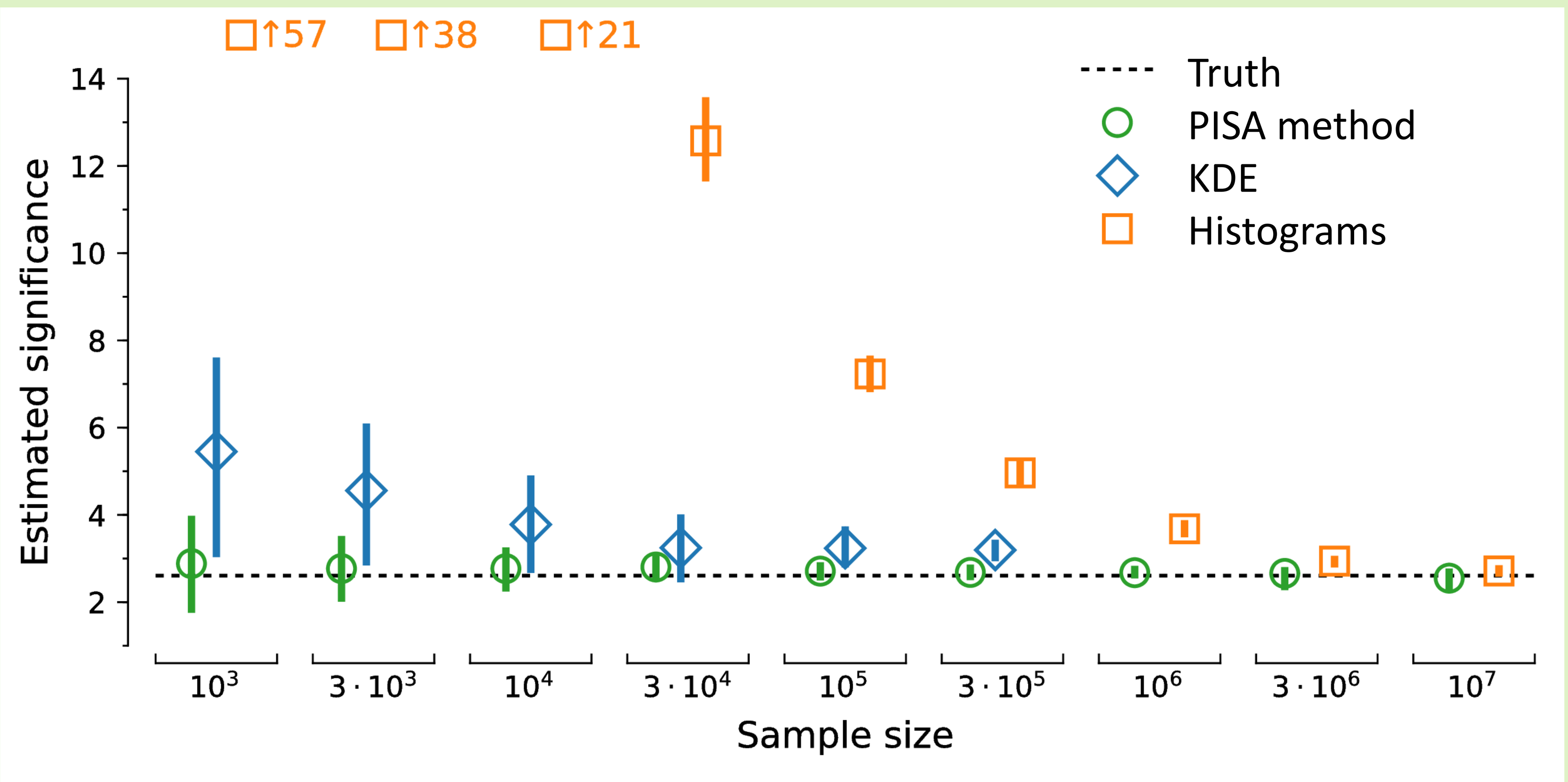
1. Atmospheric Flux interpolation
Creating a continuous function interpolating discrete values from external inputs

2. Neutrino Flavor Oscillation Probabilities
The probability of flavor oscillations, depending on energy and matter profile

3. Neutrino Interaction
Probability that an event interacts in the detector and is selected, obtained from MC simulation
+ smoothing

4. Event Reconstruction
Smears out the events according to detector resolutions, obtained from MC simulation
+ smoothing (adaptive KDE)





Example Neutrino Oscillation Analysis

- Illustrating the effect of low MC statistics
- Comparing models to data
- True significance (result) in this example lies at 2.7σ
- Estimated significances using the different methods with various MC set sizes
- Other methods (grossly) overestimate for small sample sizes
- PISA method yields reliable results virtually independent of sample size

Philipp Eller (pde3@psu.edu)
for the IceCube-GenII Collaboration

ACAT 2017
Seattle

GPU Acceleration

- All of the above methods can be accelerated by parallelizing certain calculations on GPUs
- Implemented in NVIDIA CUDA
- We get speedups of usually > O(10), depending on the GPU/CPU configuration
- Especially event-by-event calculations are suitable for simple parallelization
- Copying MC events to device at startup, and then just retrieving final histograms for futher processing (everything event specific handled on the GPU)