Neutrino Source Searches with Likelihood Landscapes

Neutrino Source Searches

- Hypothesis H0: background only flux
  - Atmospheric neutrino's
  - (Misreconstructed) Atmospheric Muons
- Hypothesis H1: background + signal flux
  - (High energy) Cosmic Neutrinos

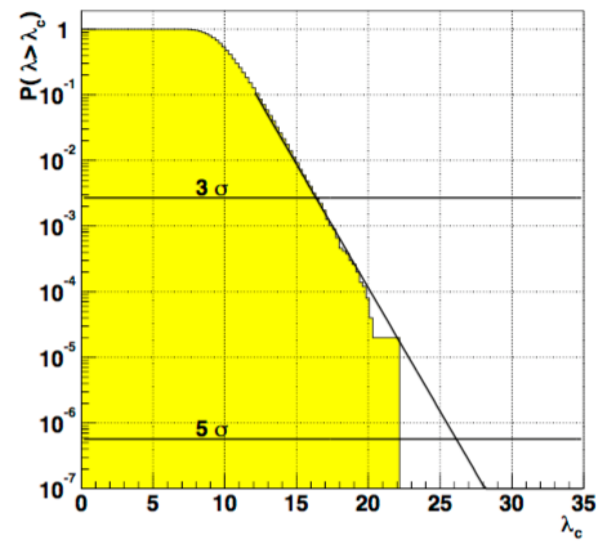
General Procedure

- How compatible is data with H0 or H1?

$$\lambda = \log\left[\frac{P(\text{data}|H_1)}{P(\text{data}|H_0)}\right]$$

- When to claim an observation?
  - Accept H1 if $\lambda > \lambda_c$
  - λ_c such that

    P(accept H1 | H0 = true) < 0.00…1



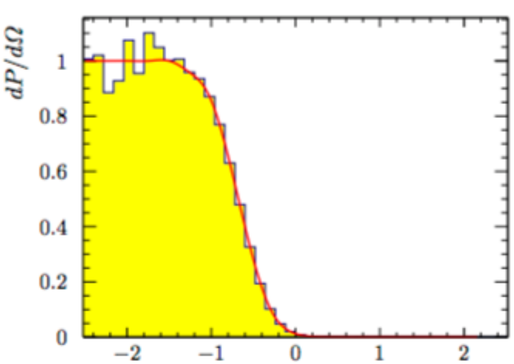
$$\lambda = \log\left[\frac{P(\text{data}|H_1)}{P(\text{data}|H_0)}\right]$$

Test Statistic (Conventional)

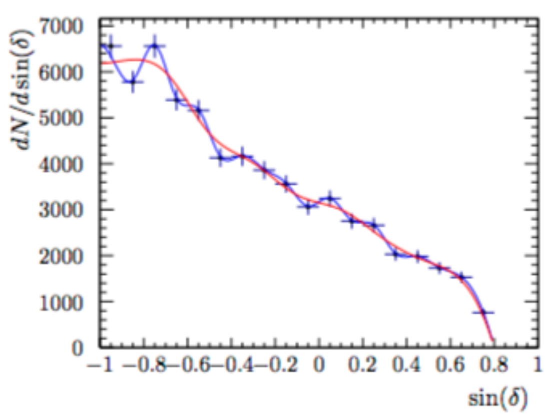
- Given detected (and selected) events $\{ev_i\}$

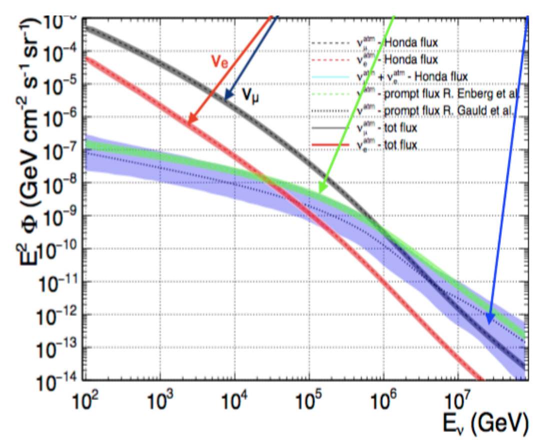
$$P(data|H) = \sum_i \left[\log \int P(x_{reco,i}|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true} \mid H)\, dx_{true}\right] - \mu^{tot}(H)$$

Reconstruction — Detection efficiency — Expected flux



log(angle reco,true)





$$\lambda = \log\left[\frac{P(\text{data}|H_1)}{P(\text{data}|H_0)}\right]$$

Test Statistic

- Given detected (and selected) events $\{ev_i\}$

$$P(data|H) = \sum_i \left[\log \int \underbrace{P(x_{reco,i}|\, x_{true})}_{\text{Reconstruction}} \cdot \underbrace{P^{det}(x_{true})}_{\text{Detection efficiency}} \cdot \underbrace{\mu(x_{true}\,|\,H)}_{\text{Expected flux}}\, dx_{true}\right] - \mu^{tot}(H)$$

- New method:

$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true}\,|\,H)\, dx_{true}\right] - \mu^{tot}(H)$$

- No big deal?

New vs. Conventional

Conventional

- Only best solution kept from reconstruction
- Selection criteria needed to select well-reconstructed events -> events are lost

- Different reconstruction algorithms (showers/tracks/tau double bang) patched together
- Event identification by BDT's and other black magic algorithms

- Parameterizations of MC events
- Fast

New Method

- Detailed knowledge of event likelihood landscape
- All events can be used

- Single 'reconstruction' algorithm for all events

- Neutrino flavour identification automatically taken into account

- Event-by-event
- Probably slow

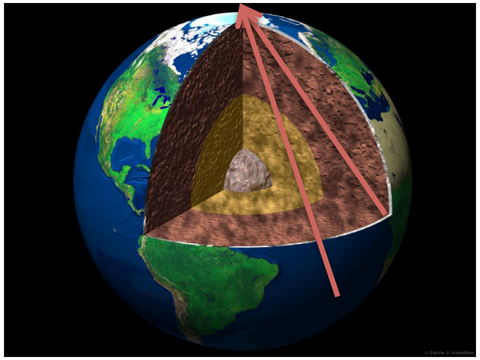
Likelihood Ingredients

$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true}\ |\, H)\, dx_{true}\right] - \mu^{tot}(H)$$

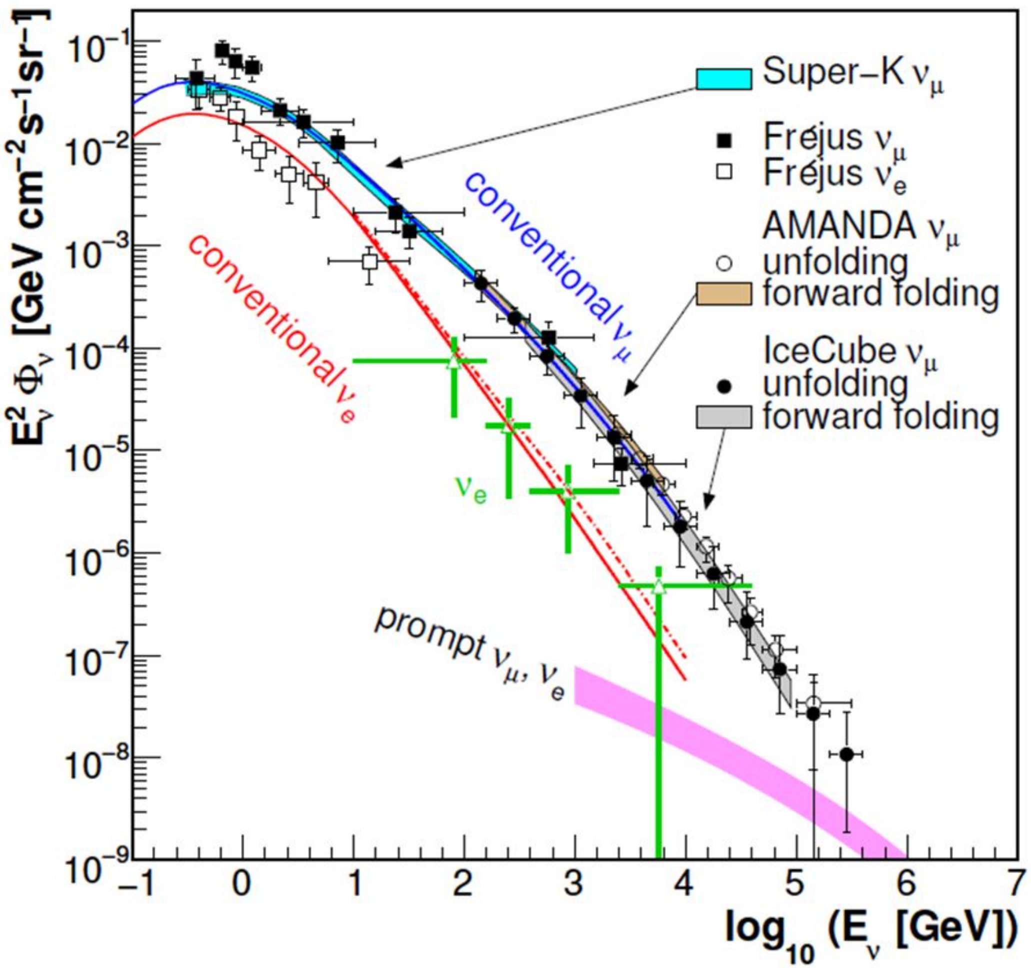
$\mu(x_{true}\ |\, H)$ Number of expected background or signal events in our detector (can)

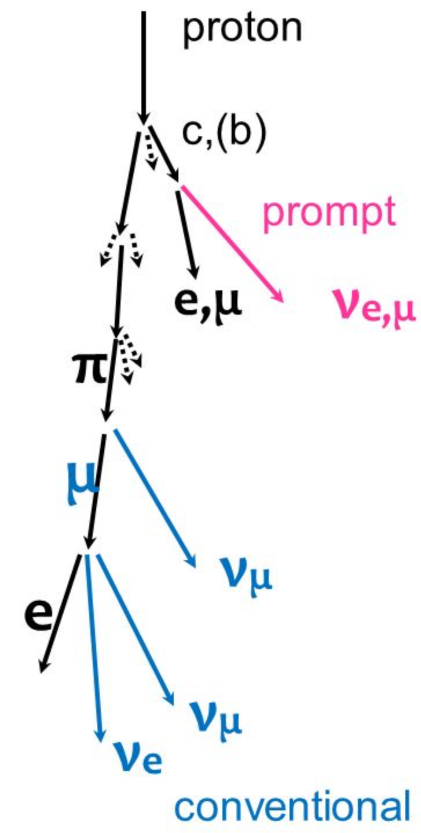
$P^{det}(x_{true})$

$P(ev_i|\, x_{true})$



Atmospheric Neutrinos

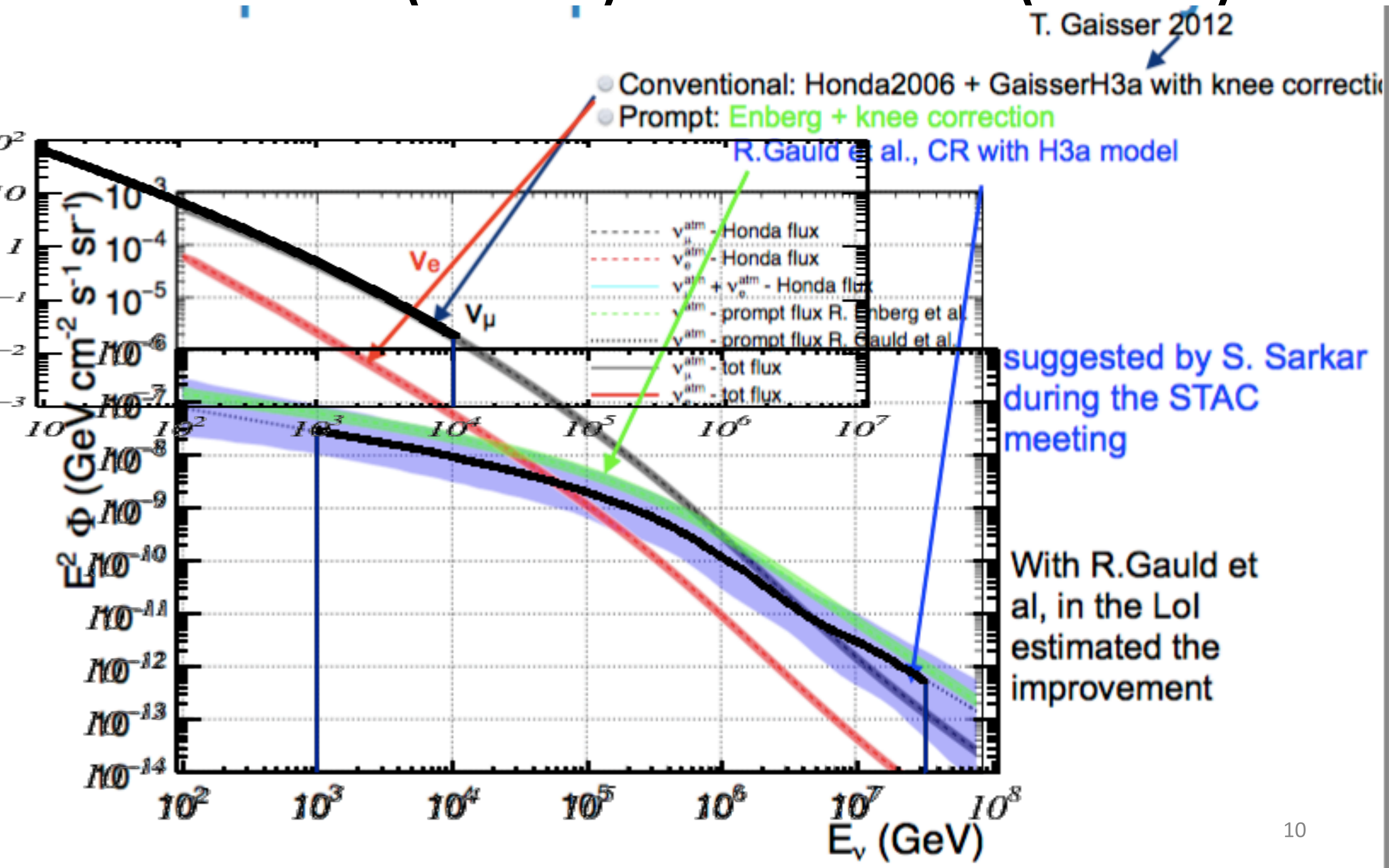




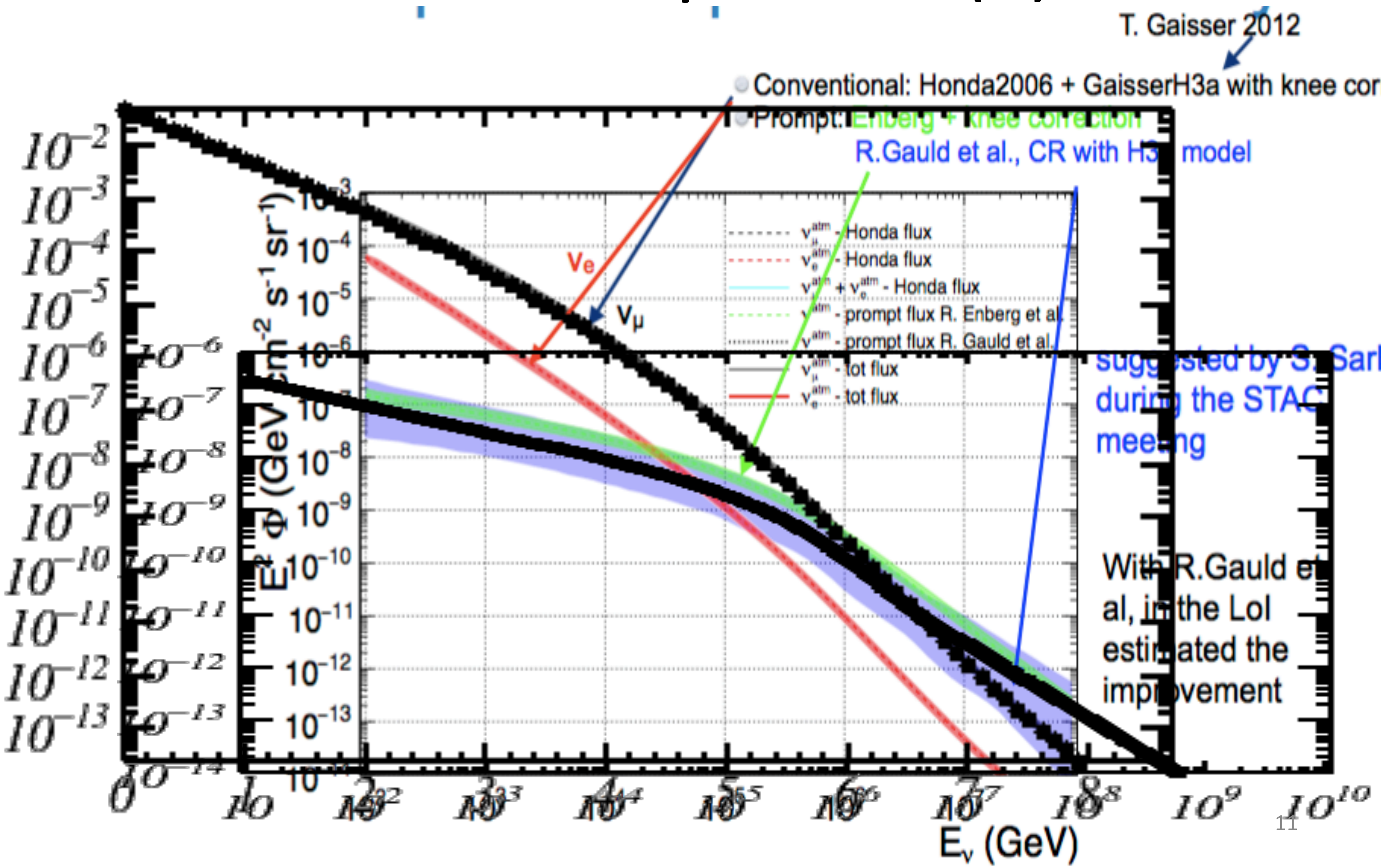
Current Parameterization

- KM3NeT Letter of Intent
- Based on Seatray
- Polynomial fit of Honda tables
  - Extrapolation to higher energy ranges
  - Outdated? Honda 2006 used.
  - Gaisser H3a knee correction
- Polynomial fit of Gauld tables 2015
  - From PromptNuFlux, L. Rottoli

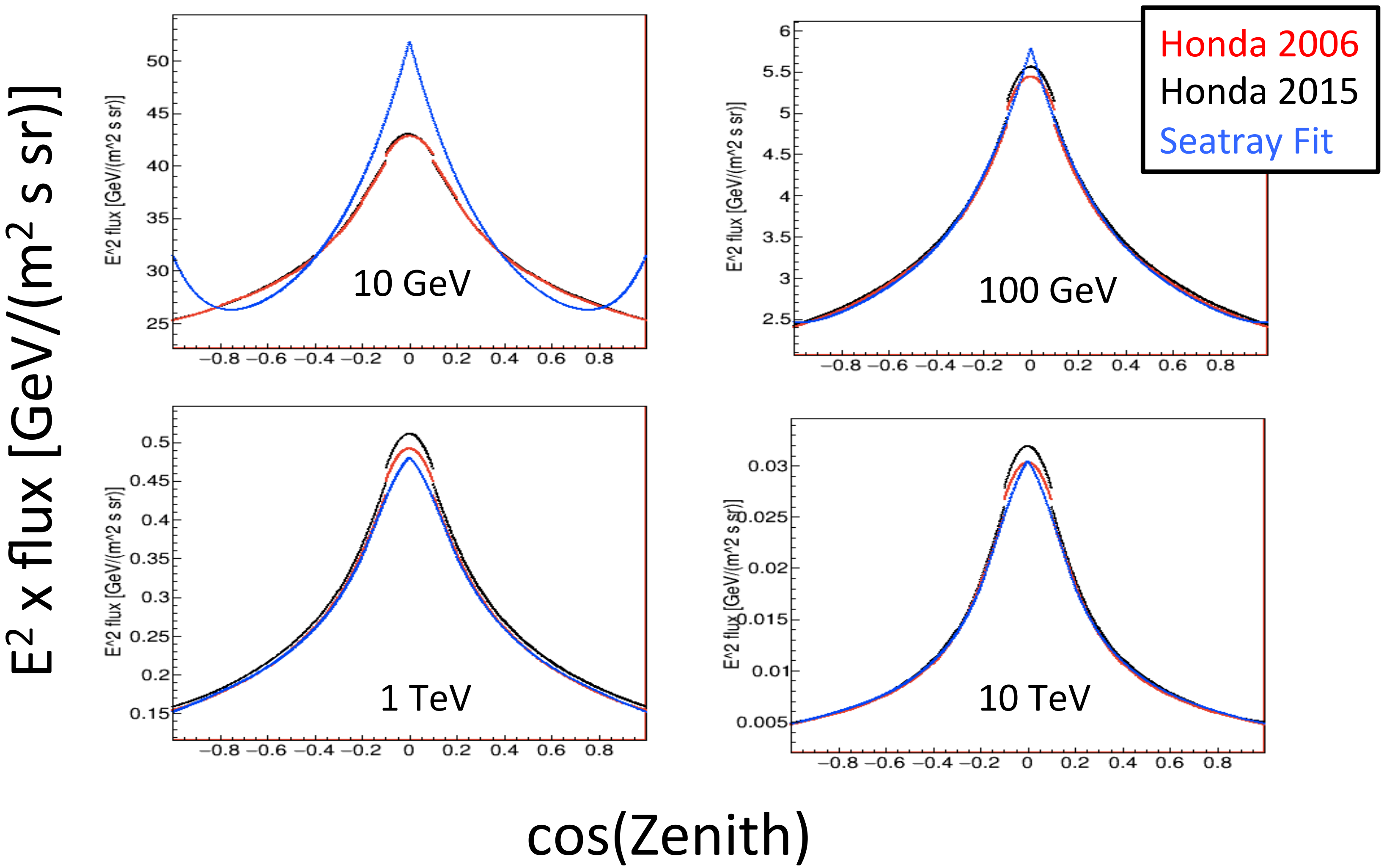
Honda (2006) and Gauld (2016)



Both Extrapolated (2)



Honda: Zenith Dependence



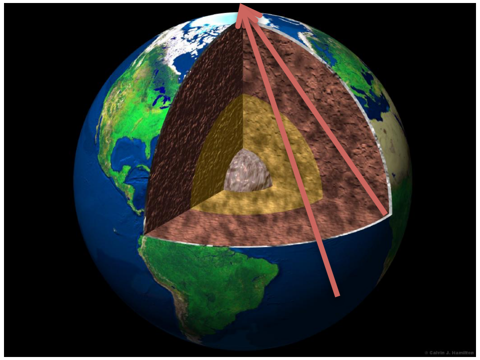
Likelihood Ingredients

$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true} \mid H)\, dx_{true}\right] - \mu^{tot}(H)$$

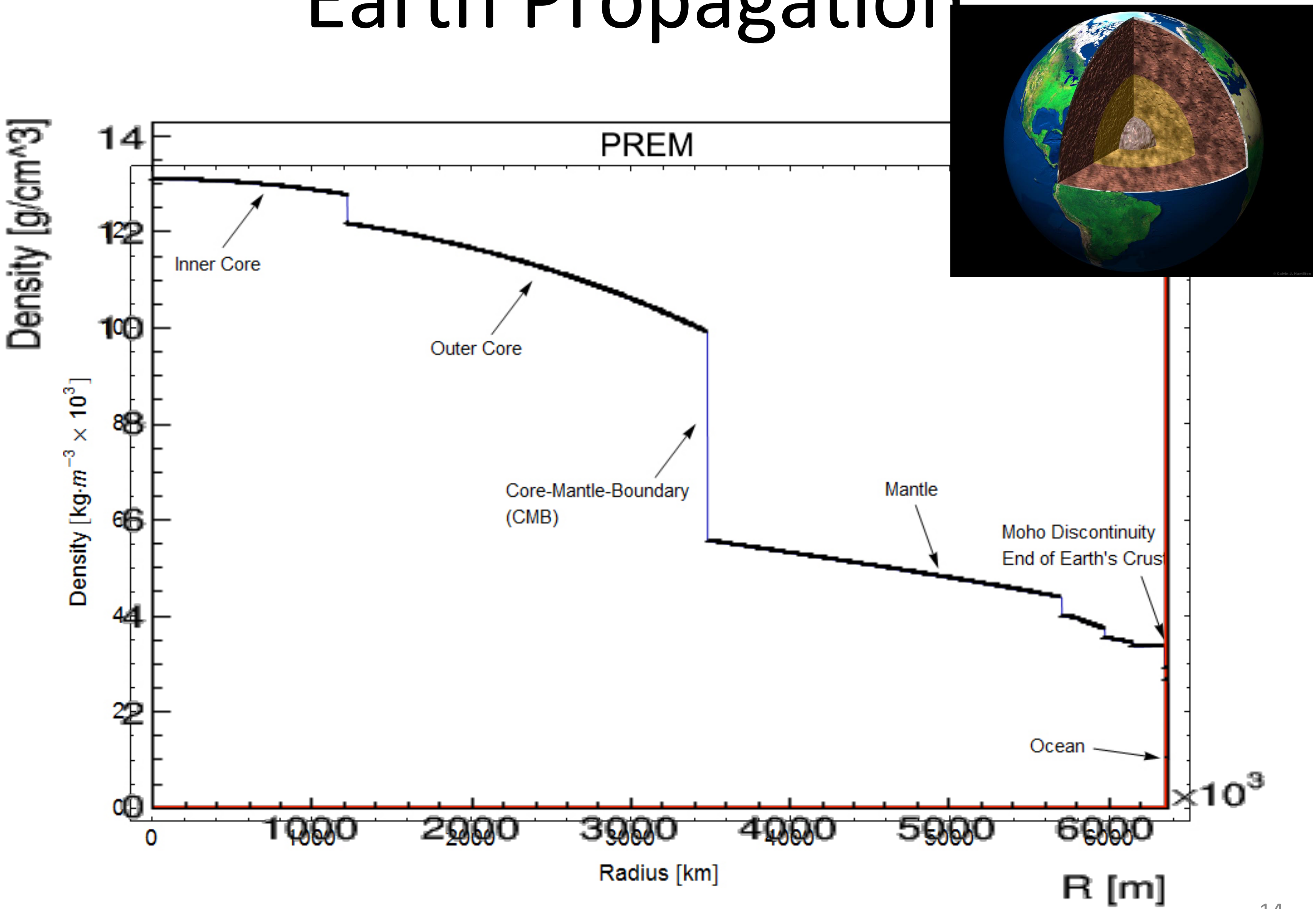
$\mu(x_{true} \mid H)$ Number of expected background or signal events in our detector (can)

$P^{det}(x_{true})$

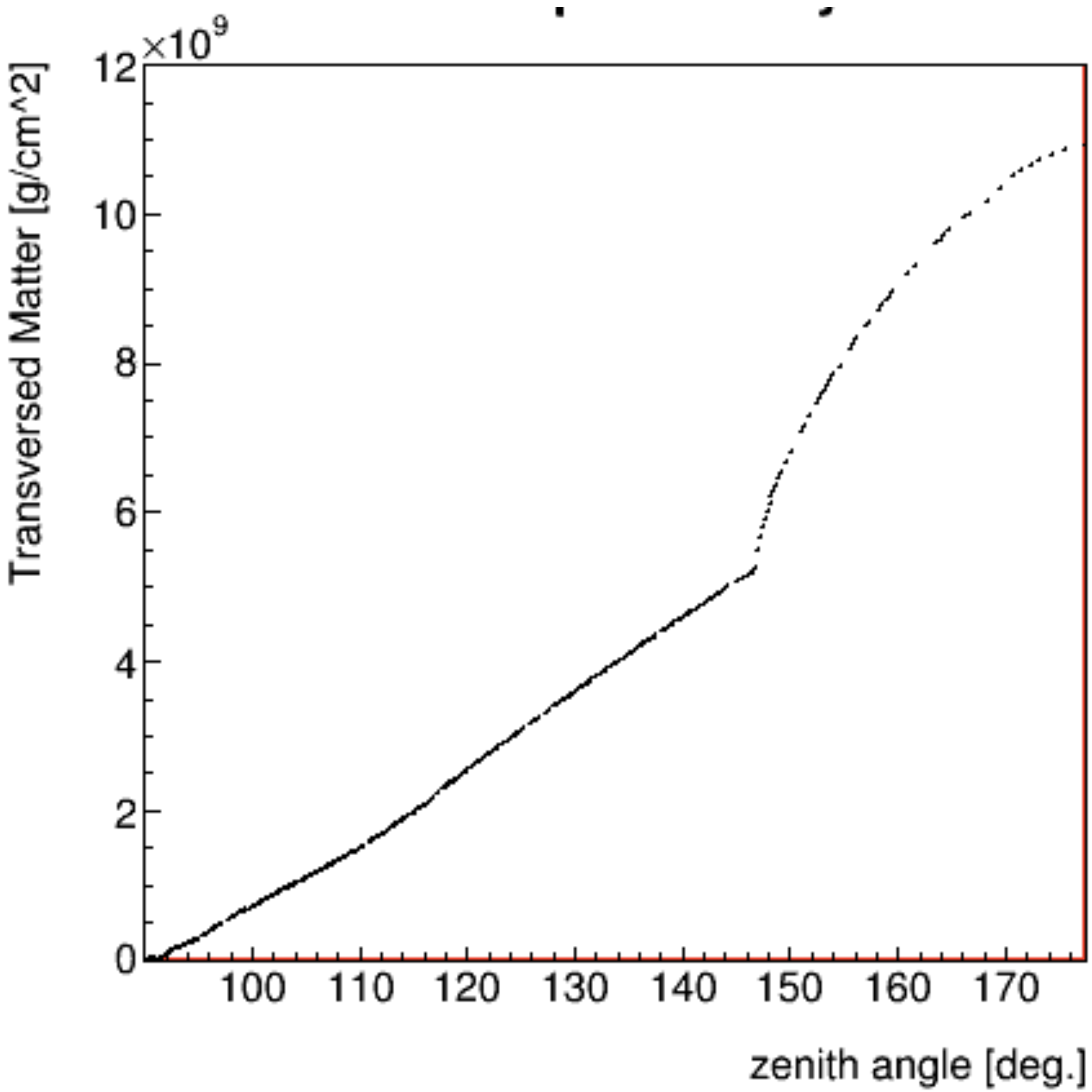
$P(ev_i|\, x_{true})$

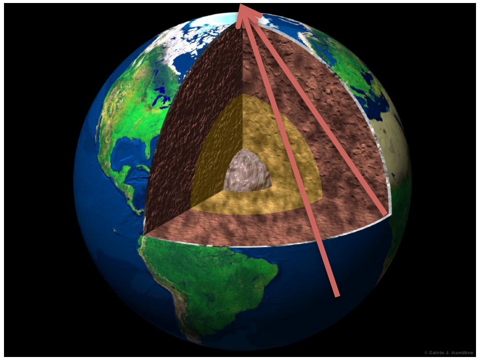


Earth Propagation



Transversed Matter Density

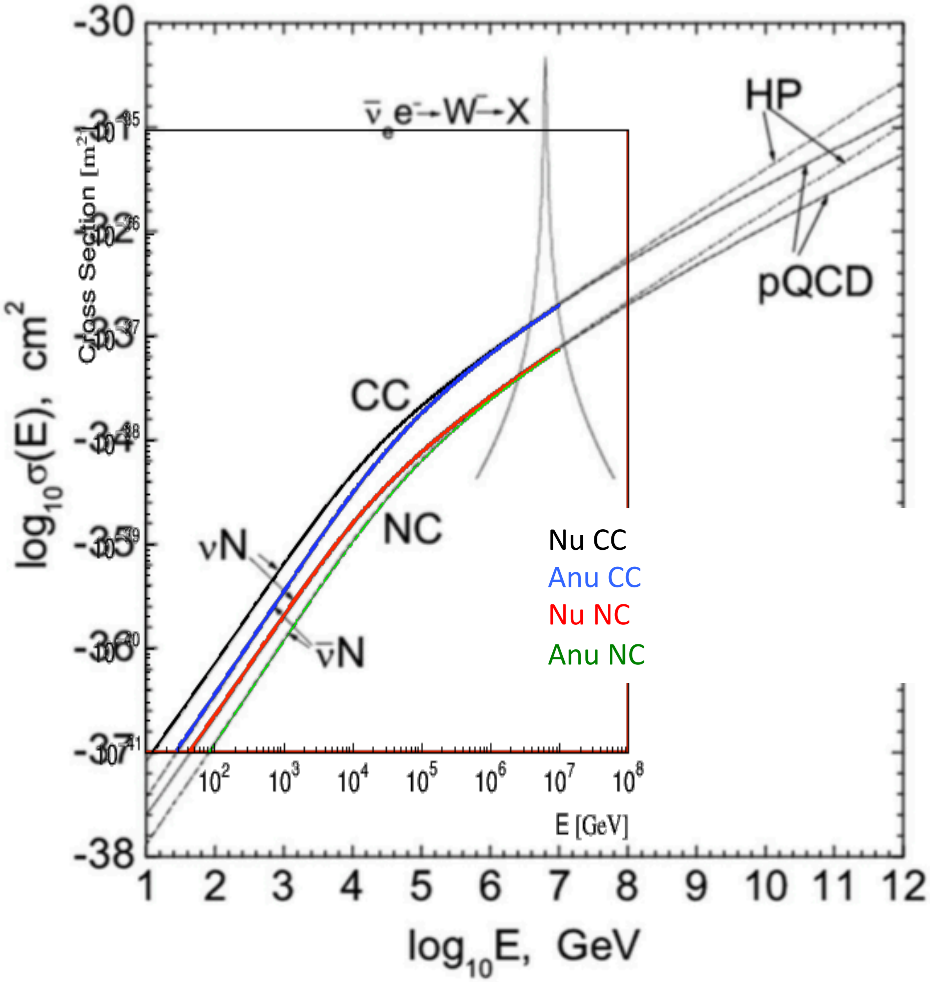




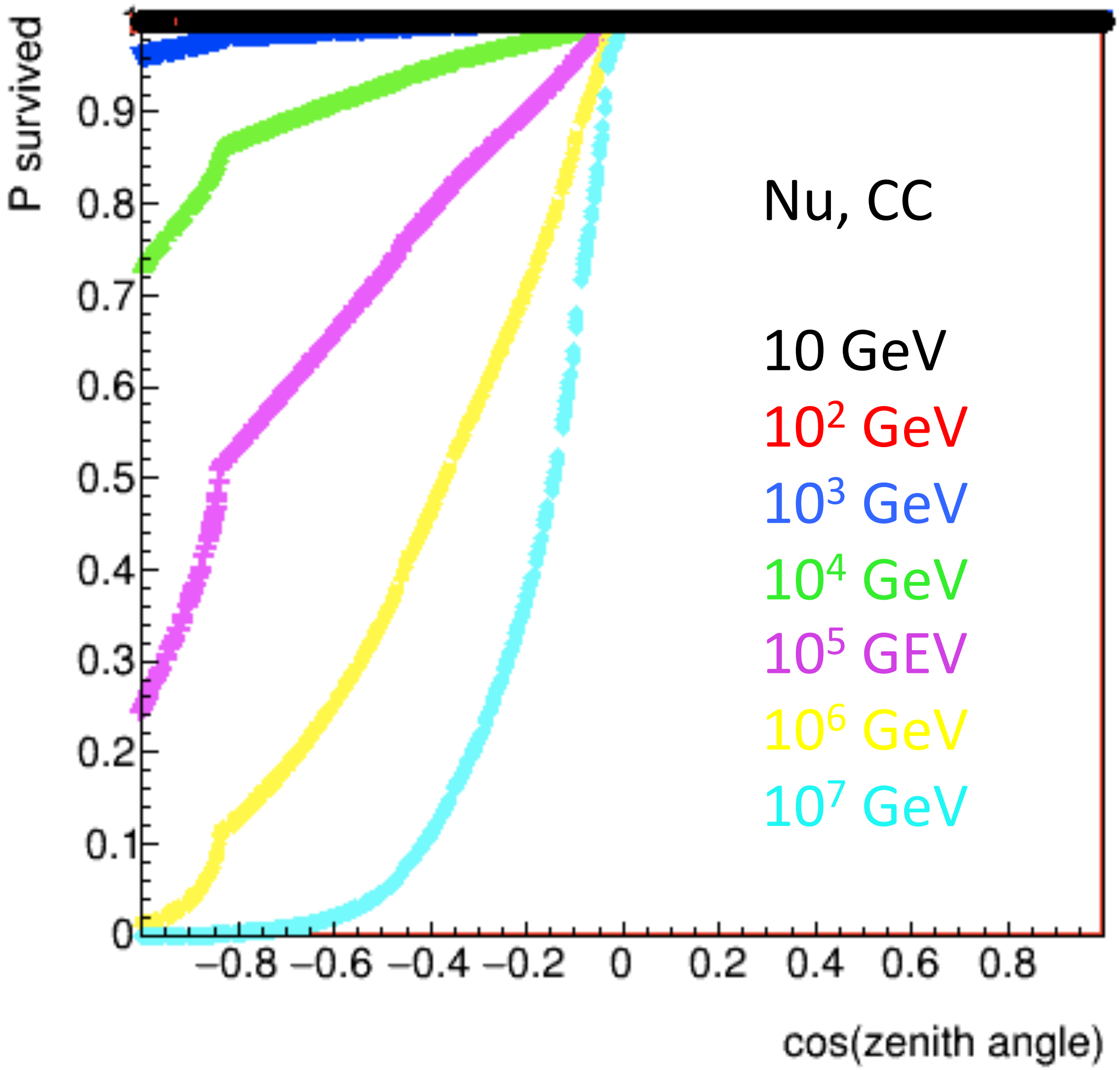
Analytically derived -> very fast

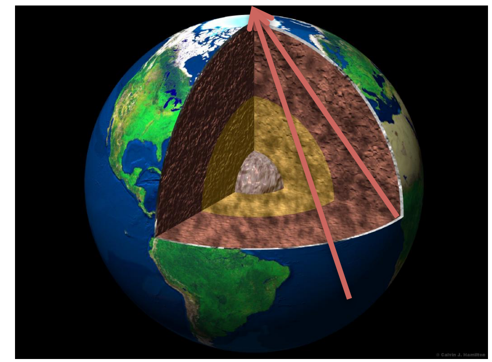
ANIS, Kowalski 2003
Figure from Colnard 2009

Neutrino Cross Sections

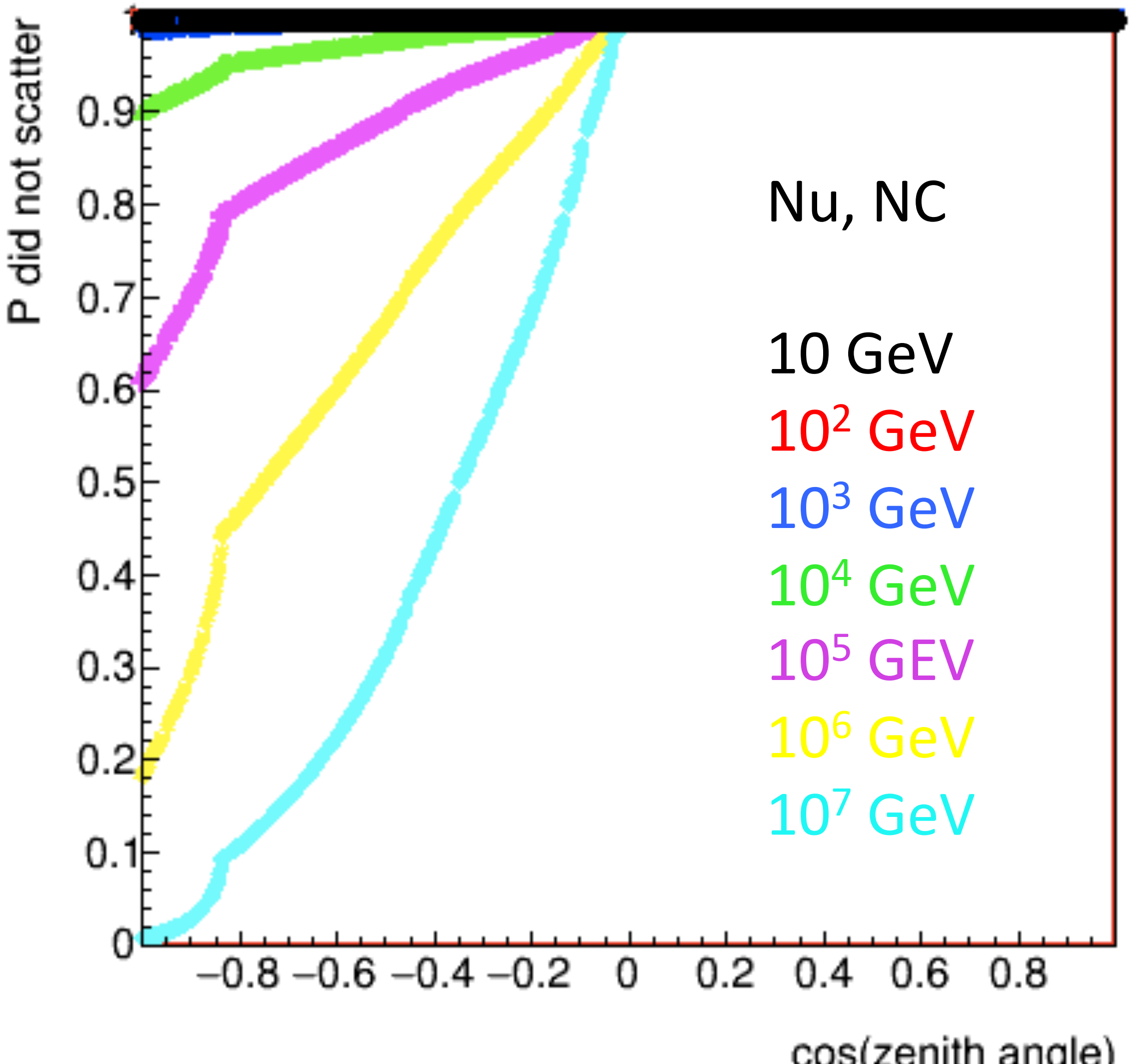


Neutrino Absorption



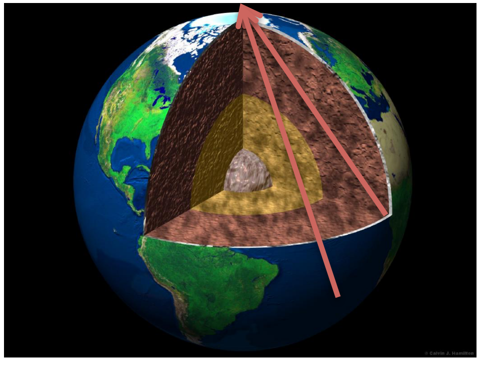


Neutrino NC Scattering (1)

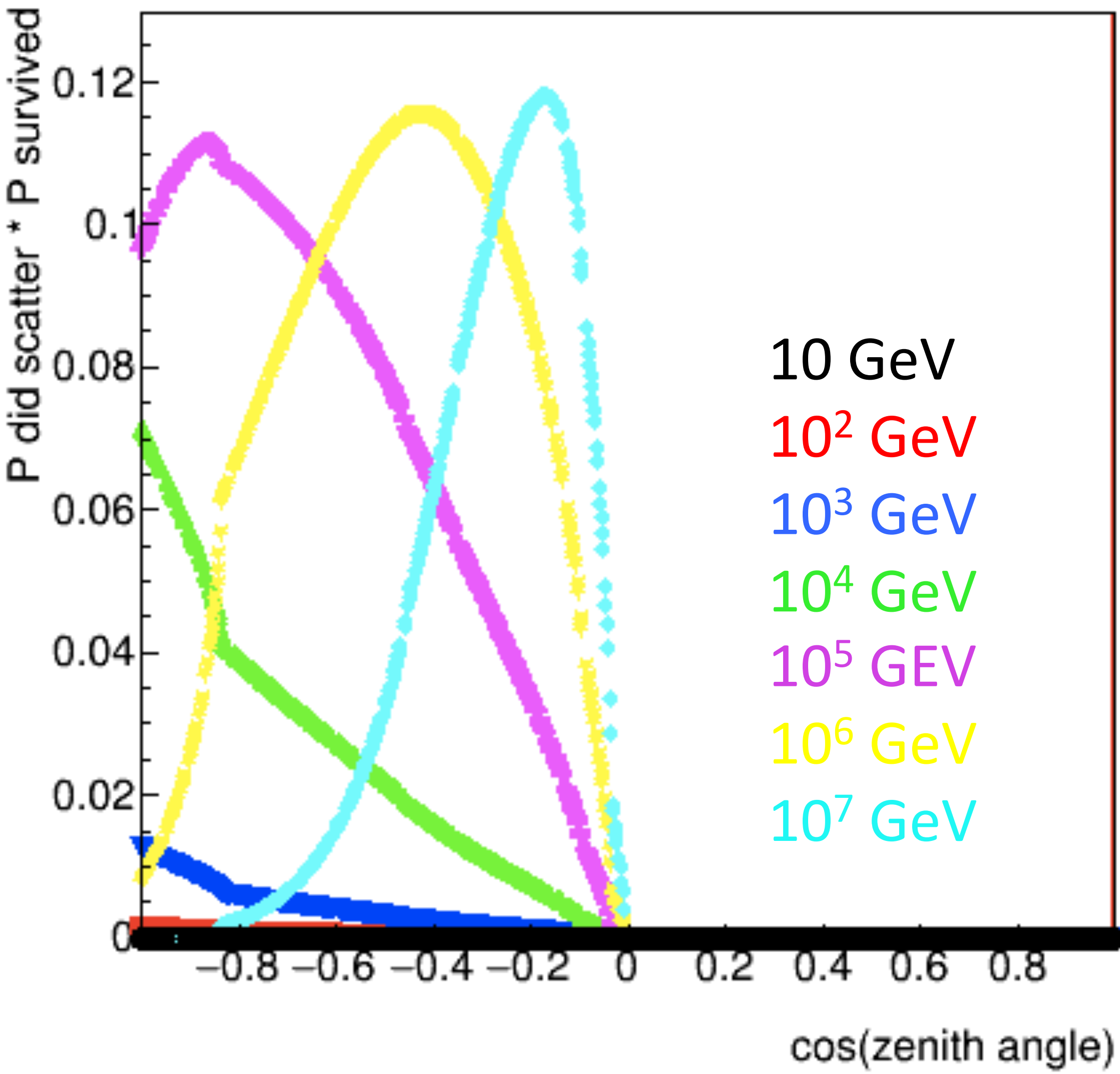


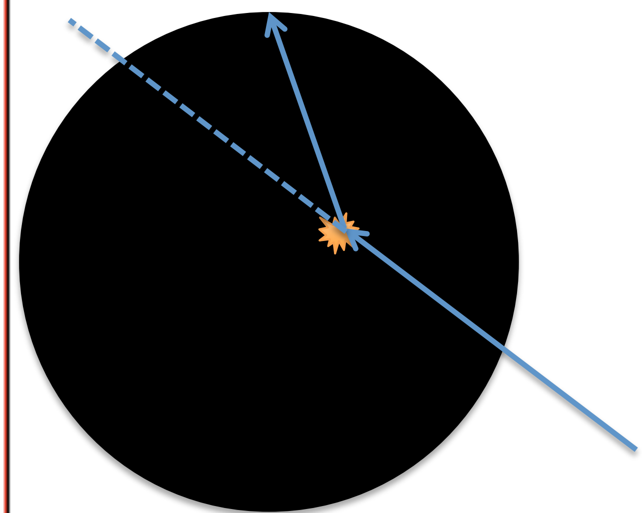
Nu, NC

10 GeV
10^2 GeV
10^3 GeV
10^4 GeV
10^5 GEV
10^6 GeV
10^7 GeV

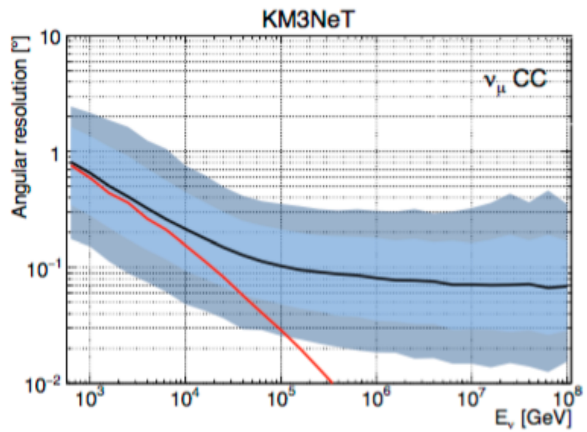


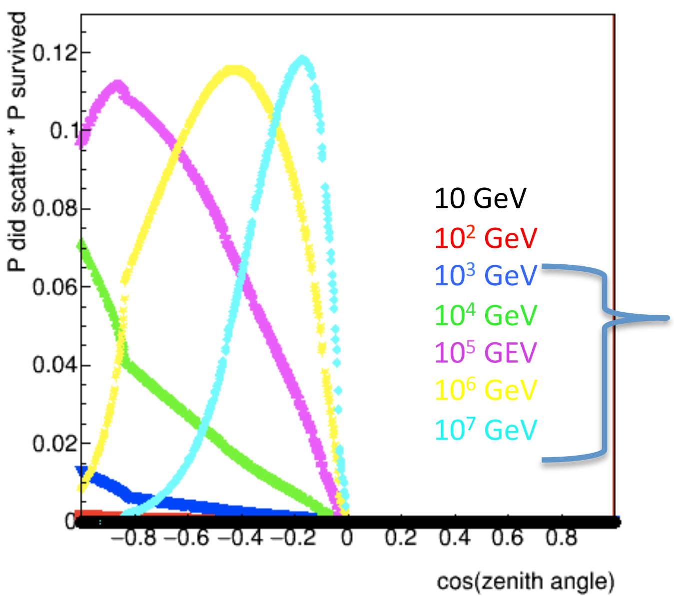
Neutrino NC Scattering (2)

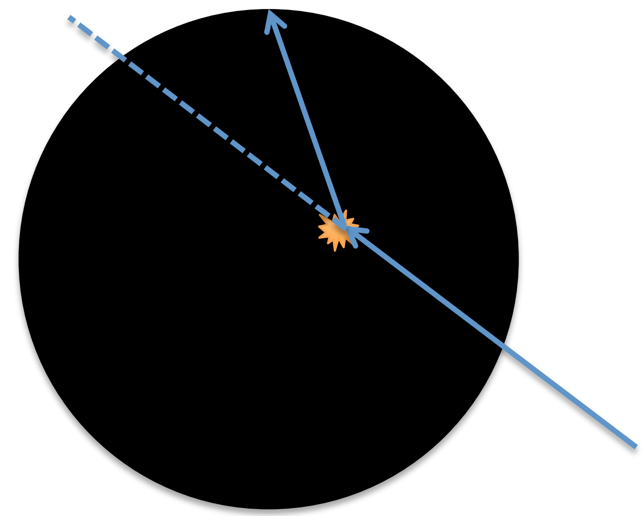




Neutrino NC Scattering (3)







- Change in direction: <≈ 0.6 degrees
  for Enu > 10^3 GeV
- Change in Energy???

Effects on expected atm. Neutrino flux neglected

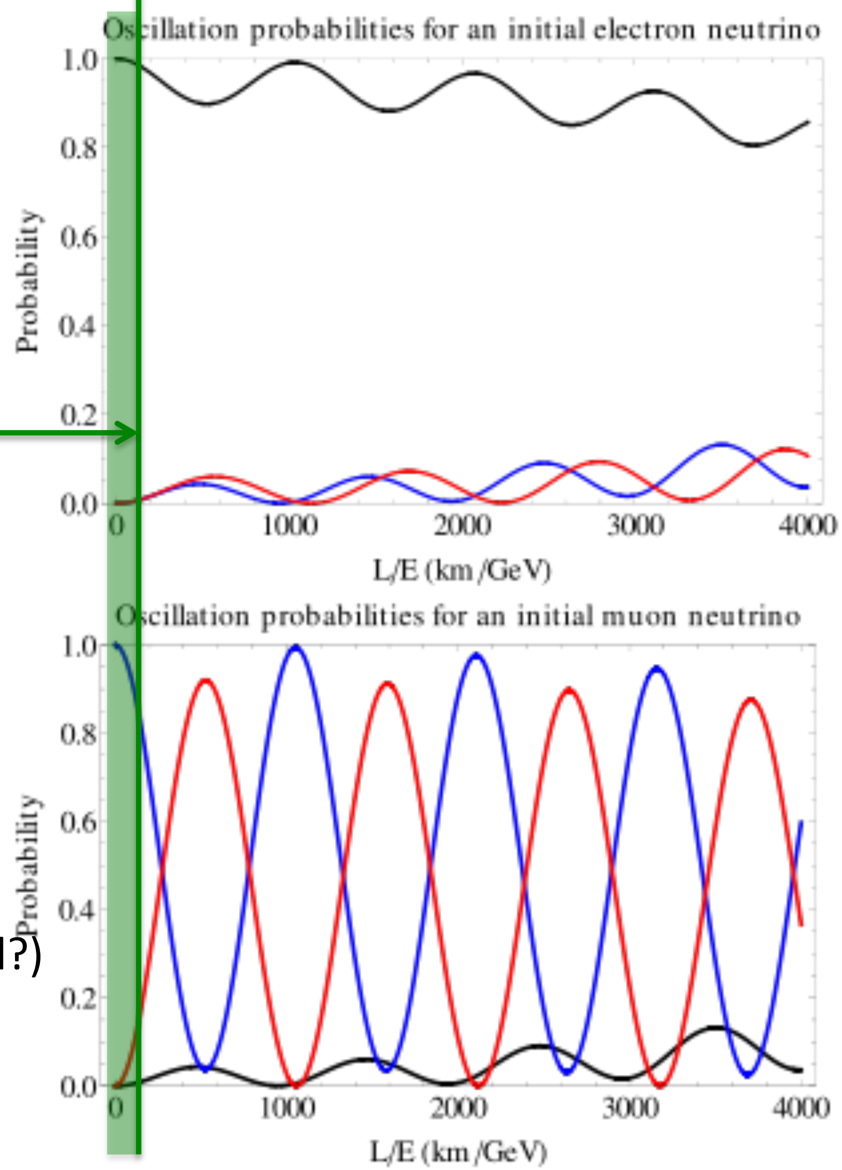
Neutrino Oscillations

Earth radius=6.4 * 10^3 km
100 GeV neutrino

L/E = 1.28 * 10^2 km/GeV

Earth covers one oscillation period
P(oscilate) up to:
~0.01 (electron -> muon/tau)
~0.2 (muon -> tau)
~0.2 (tau -> muon)

For now: Ignore…. (to be continued?)



Likelihood Ingredients

$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true} \mid H)\, dx_{true}\right] - \mu^{tot}(H)$$

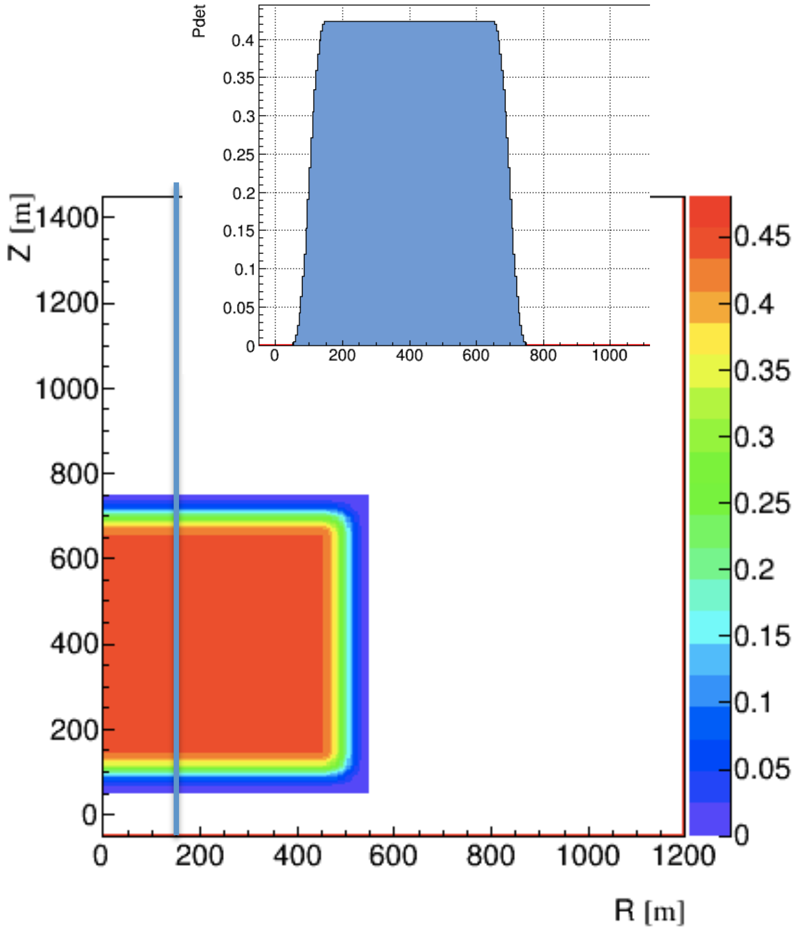
$\mu(x_{true} \mid H)$ Number of expected background or signal events in our detector (can)

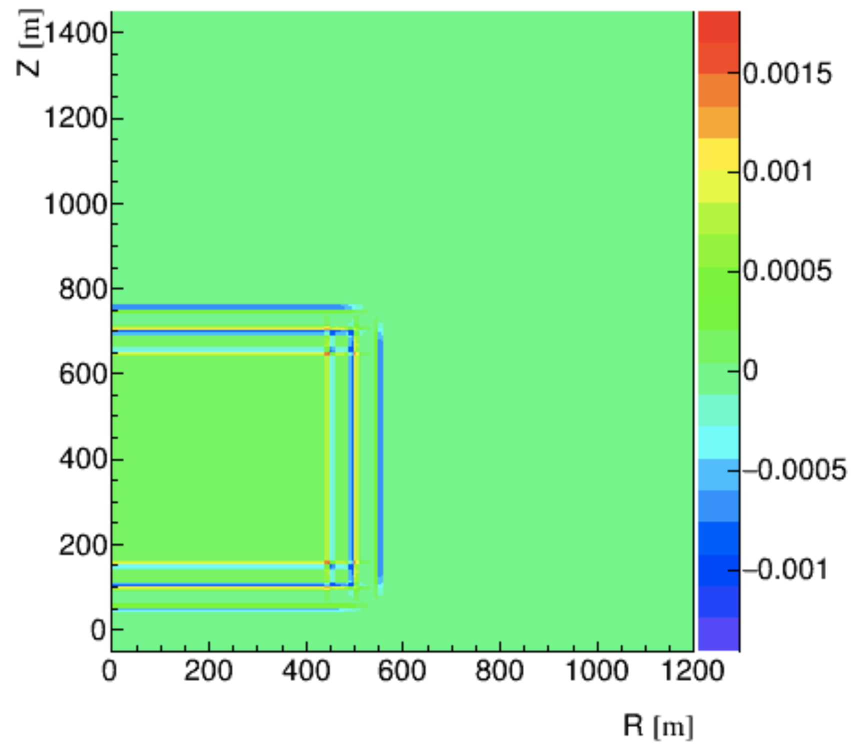
$P^{det}(x_{true})$ Probability to detect (=trigger) and select event

6-D Interpolation from tabulated values -> fast

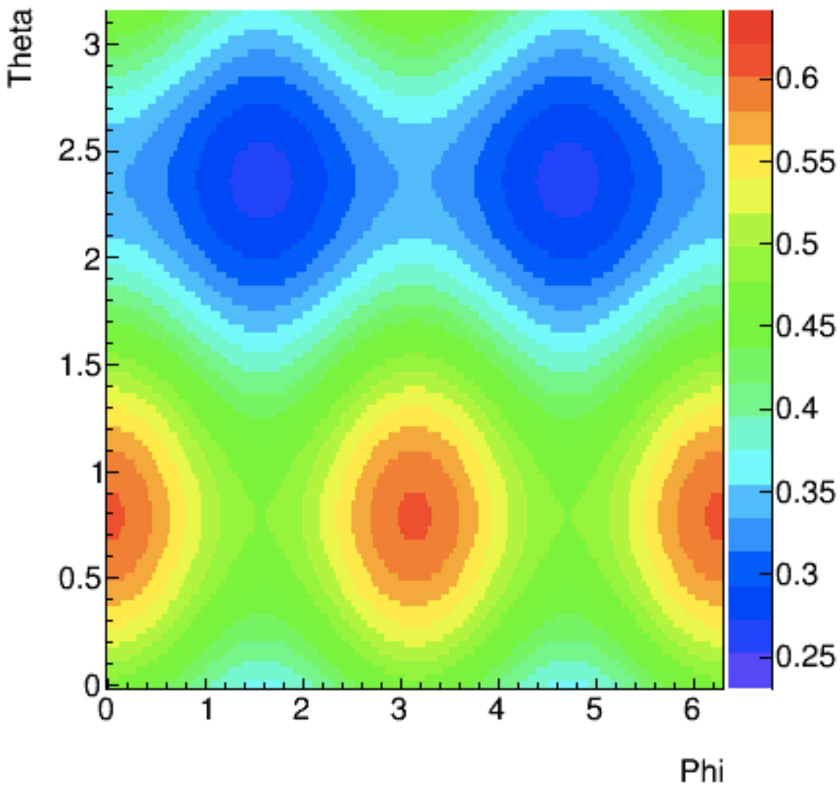
$P(ev_i|\, x_{true})$

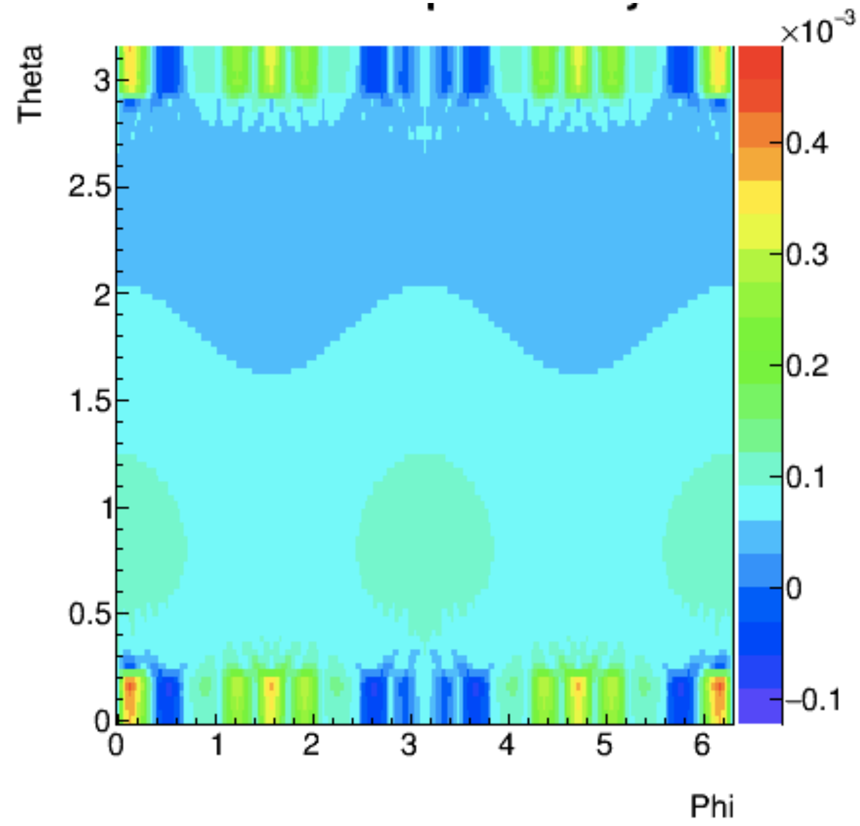
Detection Efficiency (1)





Detection Efficiency (2)





What is Pdet?

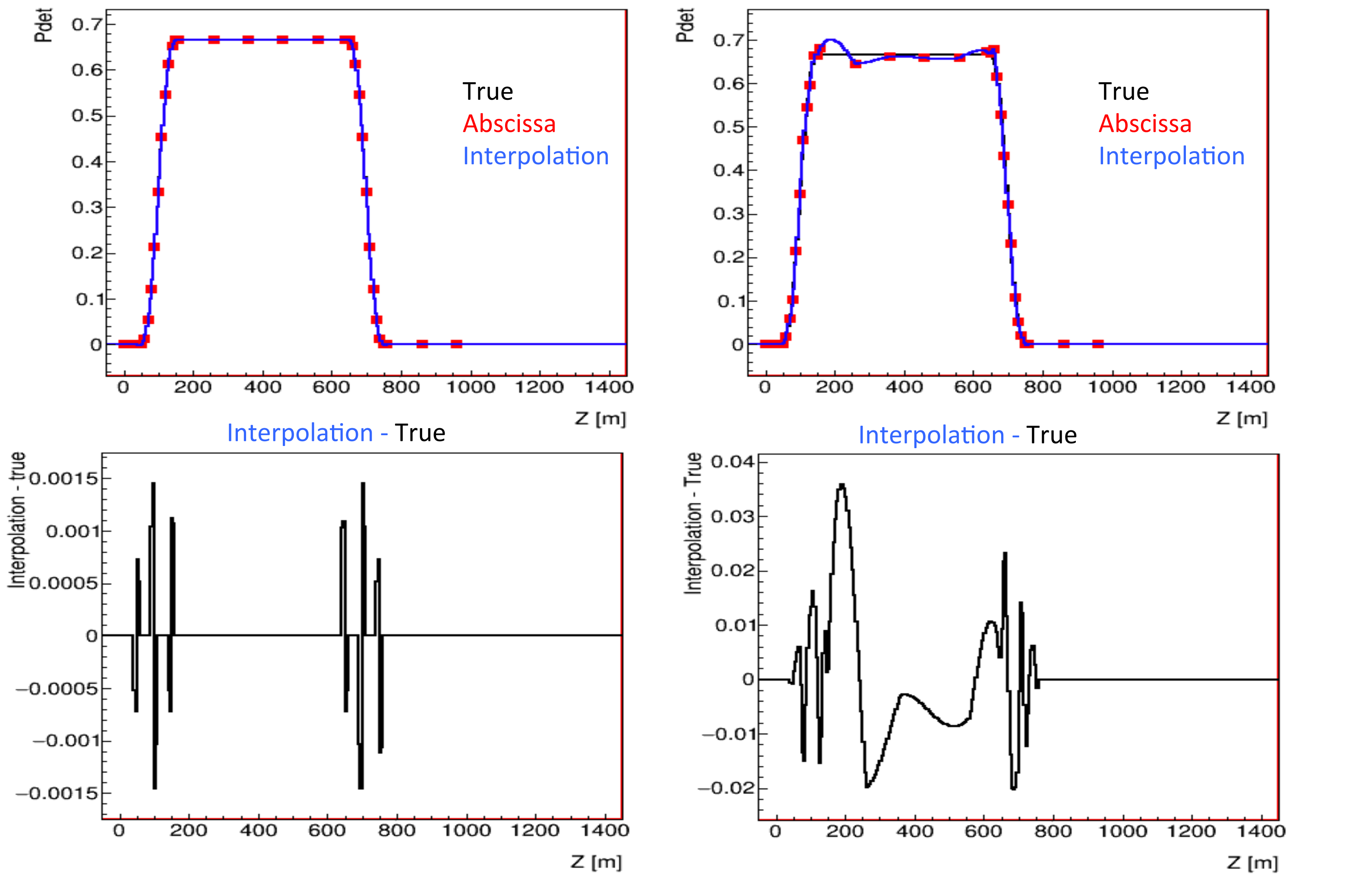
- Probability that an event:
  - Causes hits in detector: Jsirene
  - Leads to a trigger: JTriggerEfficiency
  - Is selected (reject atm. Muons): ??

- Get Pdet(x_{true}) by running MC events

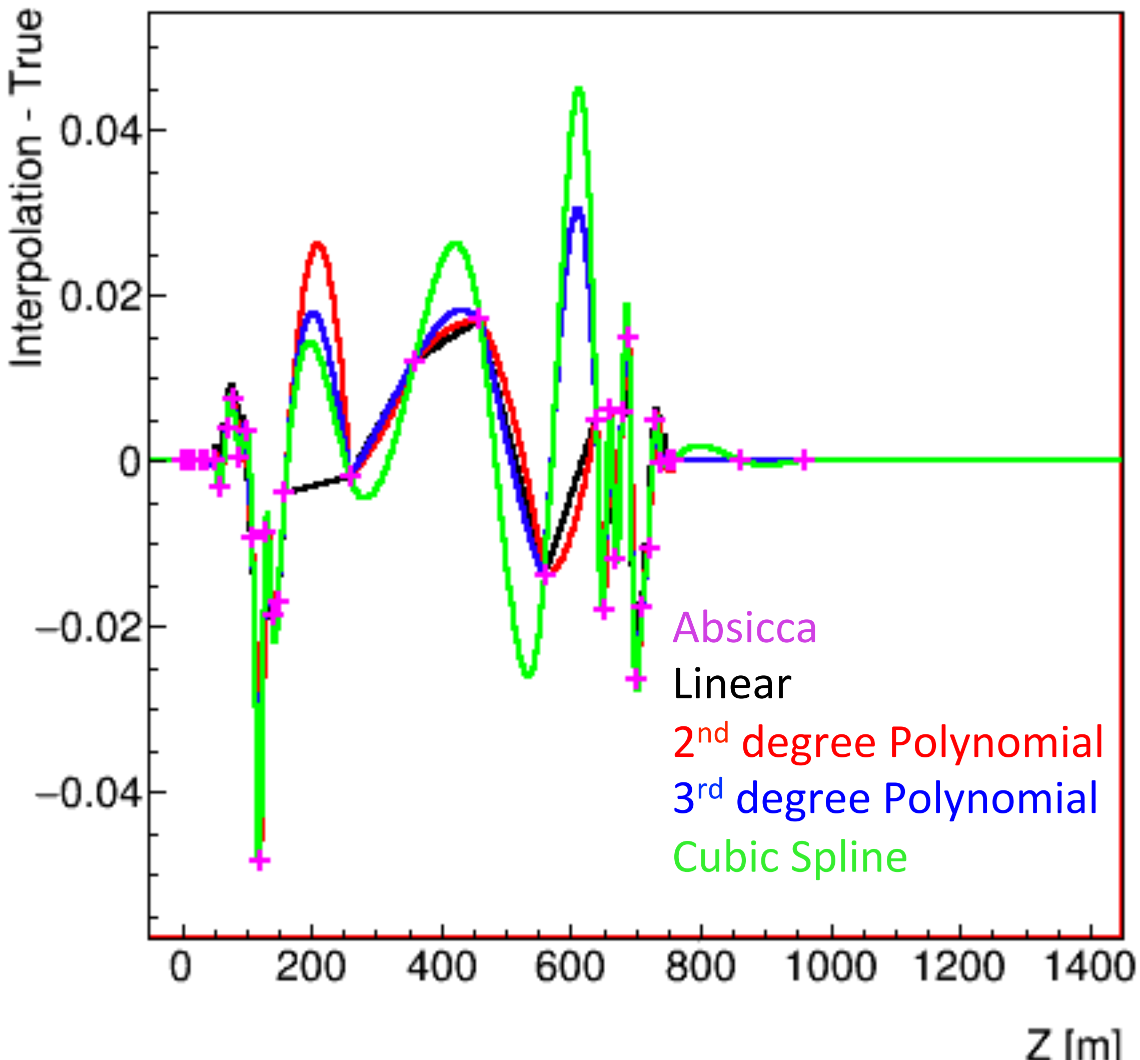
Statistical Fluctuations

Ideal case: Infinite statistics

Realistic case: 1000 simulated events



Different Interpolation Techniques

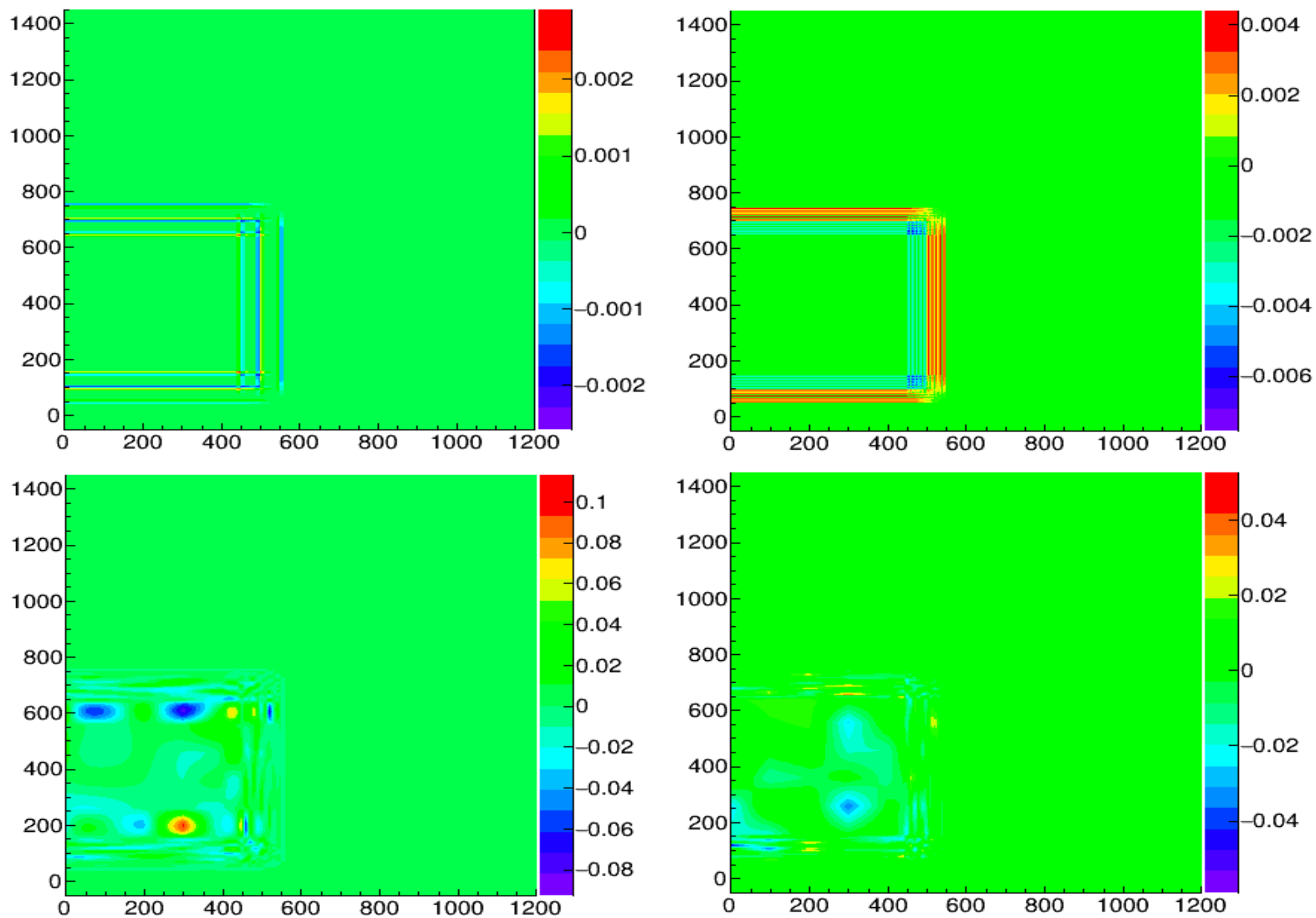


Polynomial vs linear fit

3rd degree polynomial | Linear fit

No stat. fluctuations

With stat. fluctuations



Time Consumption

```
Scanning over 72000 Positions * 98 Directions * 1 Energy-bins = 7056000 points... Done in

624169.543 ms elapsed
623814.165 ms user
    12.998 ms system
 99%CPU
```

3rd degree polynomial interpolation of 7 million points in 10 minutes

```
Scanning over 72000 Positions * 98 Directions * 1 Energy-bins = 7056000 points... Done in

 16068.632 ms elapsed
 16057.558 ms user
     4.999 ms system
 99%CPU
```

Linear interpolation of 7 million points in 16 seconds

Likelihood Ingredients

$$P(data|H) = \sum_i \left[\log \int P(ev_i |\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true} \mid H)\, dx_{true}\right] - \mu^{tot}(H)$$

$\mu(x_{true} \mid H)$ Number of expected background or signal events in our detector (can)

$P^{det}(x_{true})$ Probability to detect (=trigger) and select event

$P(ev_i |\, x_{true})$ Reconstruction, loop over PMTs. Phit * Ptime -> to do

Conclusions

- New method seems promising
- Most ingredients in place
- ‘Reconstruction’ part to be done

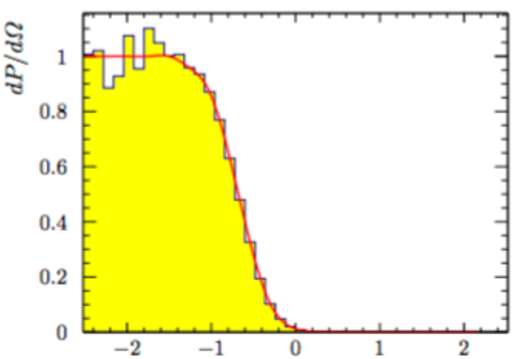
Recap

$$\lambda = \log\left[\frac{P(\mathrm{data}|H_1)}{P(\mathrm{data}|H_0)}\right]$$

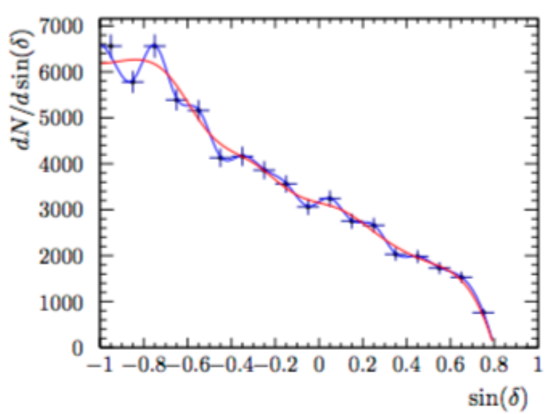
Test Statistic (Conventional)

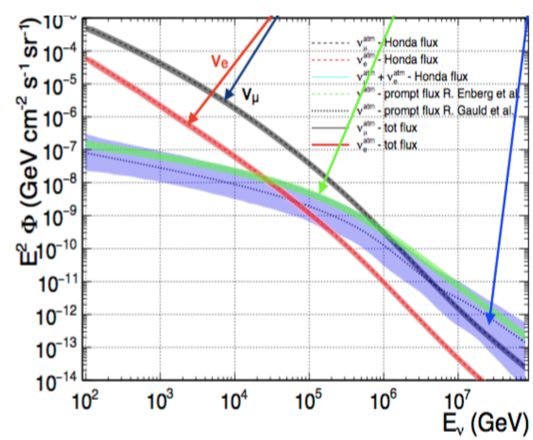
- Given detected (and selected) events $\{ev_i\}$

$$P(data|H) = \sum_i \left[\log \int \underbrace{P(x_{reco,i}|\, x_{true})}_{\text{Reconstruction}} \cdot \underbrace{P^{det}(x_{true})}_{\text{Detection efficiency}} \cdot \underbrace{\mu(x_{true}\mid H)}_{\text{Expected flux}}\, dx_{true}\right] - \mu^{tot}(H)$$



log(angle reco,true)





$$\lambda = \log\left[\frac{P(\mathrm{data}|H_1)}{P(\mathrm{data}|H_0)}\right]$$

Test Statistic

- Given detected (and selected) events $\{ev_i\}$

$$P(data|H) = \sum_i \left[\log \int \underbrace{P(x_{reco,i}|\, x_{true})}_{\text{Reconstruction}} \cdot \underbrace{P^{det}(x_{true})}_{\text{Detection efficiency}} \cdot \underbrace{\mu(x_{true}\,|\,H)}_{\text{Expected flux}}\, dx_{true}\right] - \mu^{tot}(H)$$

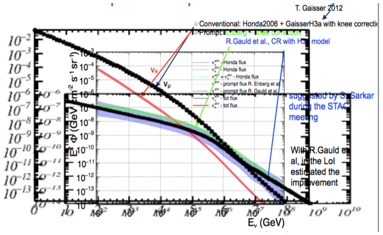
- New method:

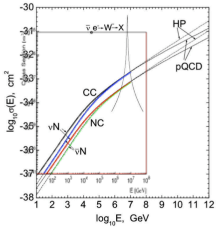
$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true}\,|\,H)\, dx_{true}\right] - \mu^{tot}(H)$$

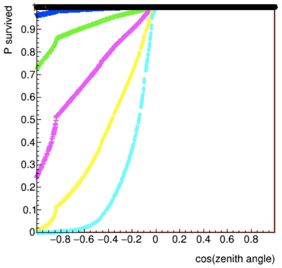
Recap: Likelihood Ingredients

$$P(data|H) = \sum_i \left[\log \int P(ev_i |\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true} \,|\, H)\, dx_{true}\right] - \mu^{tot}(H)$$

$\mu(x_{true} \,|\, H)$ Number of expected background or signal events in our detector (can)







$P^{det}(x_{true})$ Probability to detect (=trigger) and select event

$P(ev_i |\, x_{true})$ Reconstruction, loop over PMTs.

Detection Efficiency

- For each neutrino energy, bjorken-y, position, direction (6 parameters), DO:

- (Very fast) Monte Carlo generator:
  - Secondary particles
  - Photon propagation (JSirene)
  - Trigger

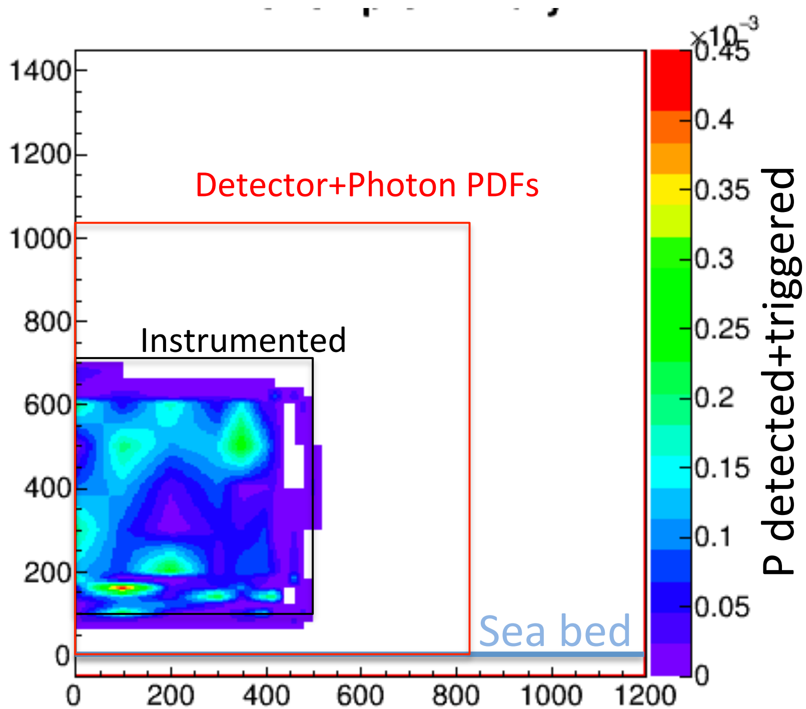
BONUS bij AH Willems

- Count fraction of trig. ev.
- Store in 6D interpolatable PDF table

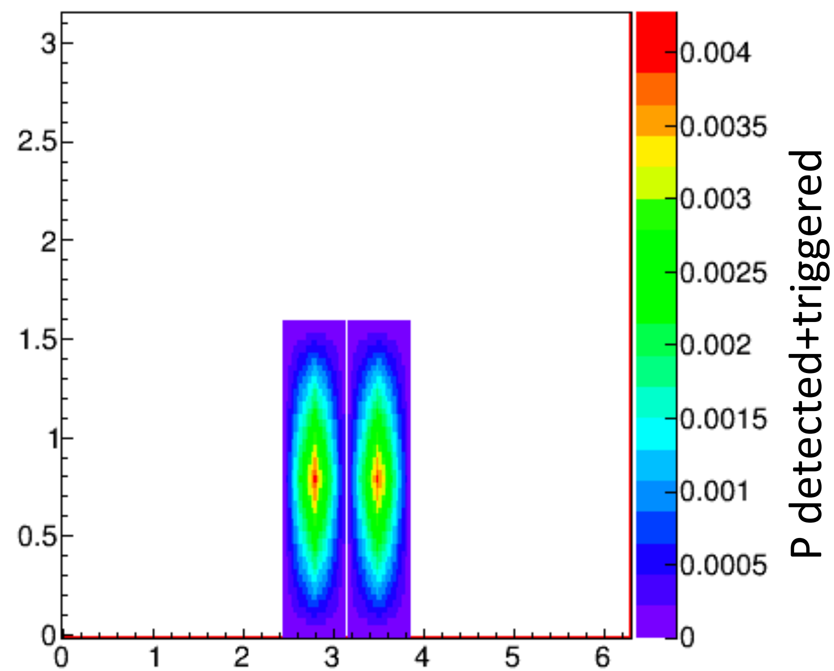
Detection Efficiency @ 10^2 GeV

NC electron-neutrino (only single hadronic shower)

Position



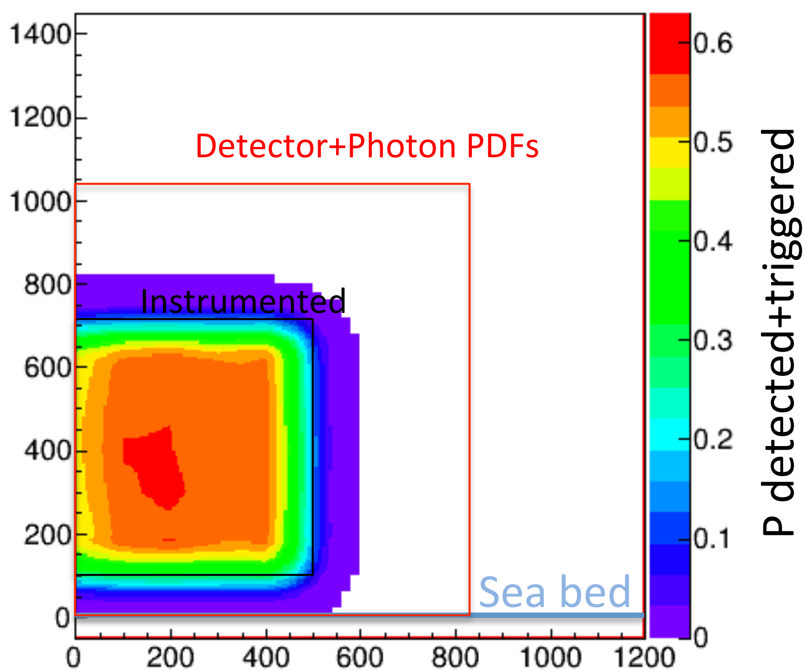
Direction @R=0, Z=400



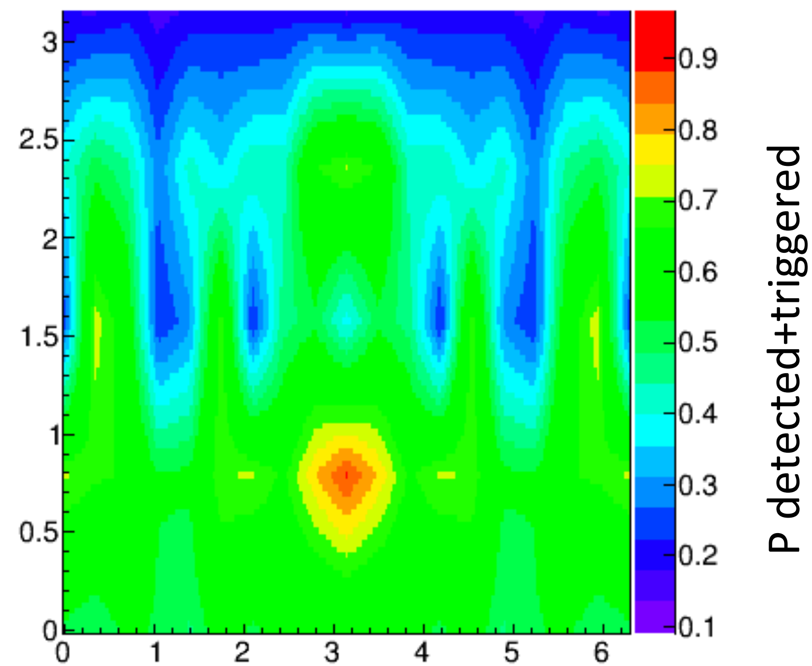
Detection Efficiency @ 10^3 GeV

NC electron-neutrino (only single hadronic shower)

Position



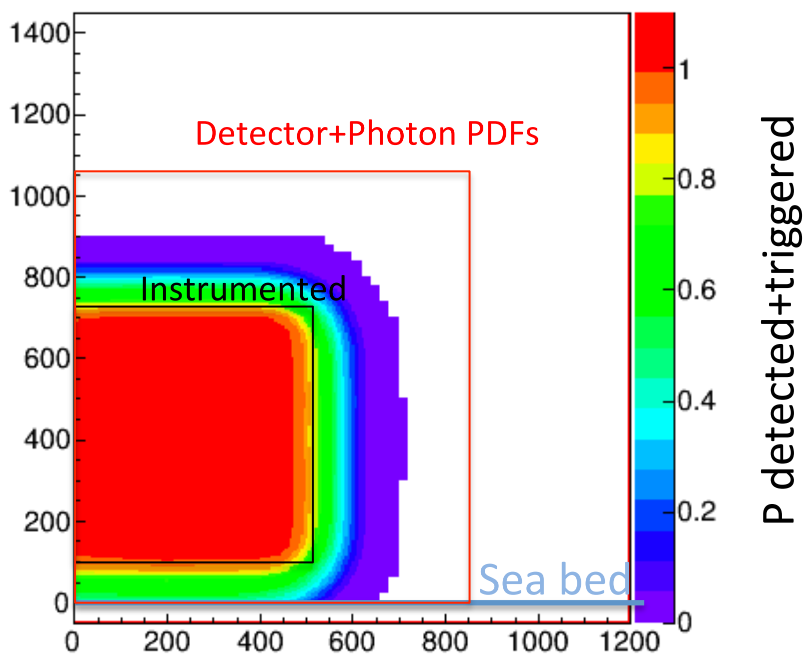
Direction @R=0, Z=400



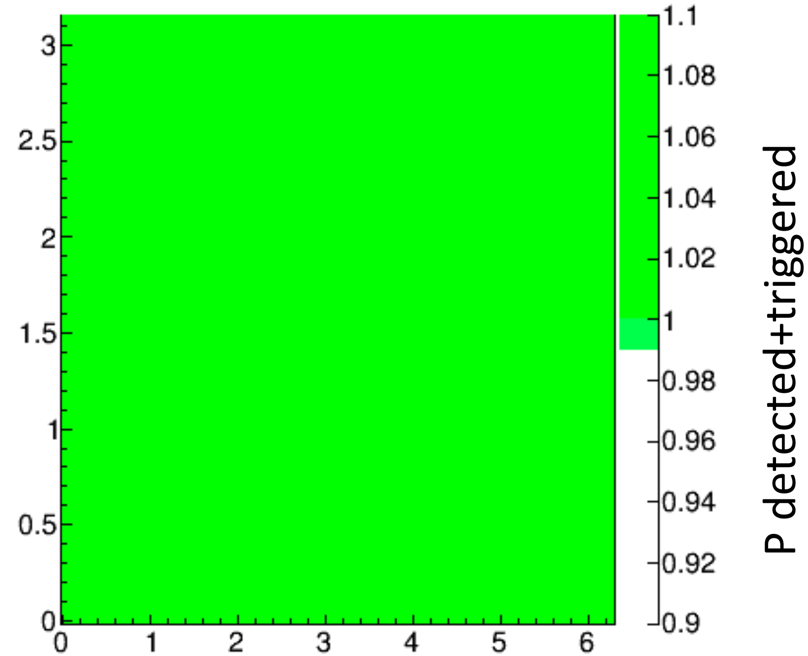
Detection Efficiency @ 10^4 GeV

NC electron-neutrino (only single hadronic shower)

Position



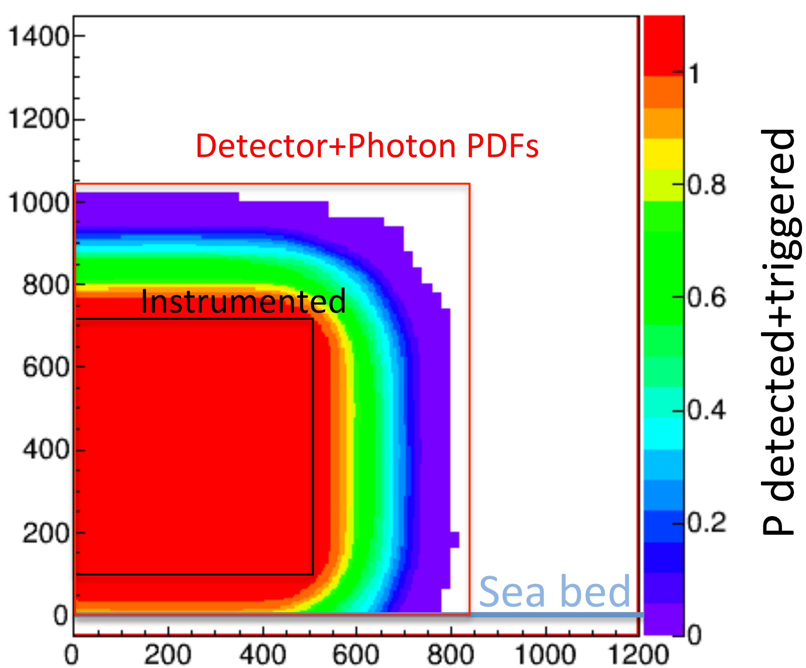
Direction @R=0, Z=400



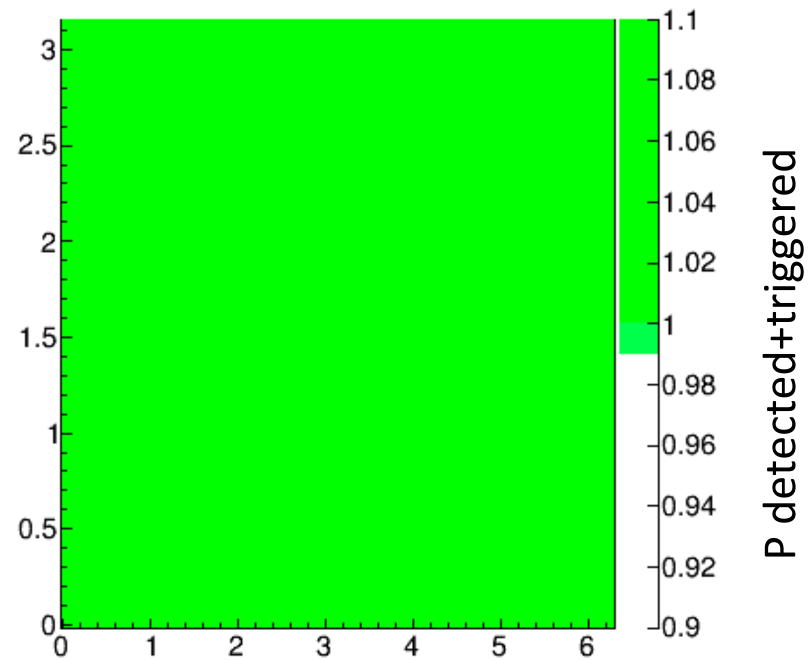
Detection Efficiency @ 10^5 GeV

NC electron-neutrino (only single hadronic shower)

Position



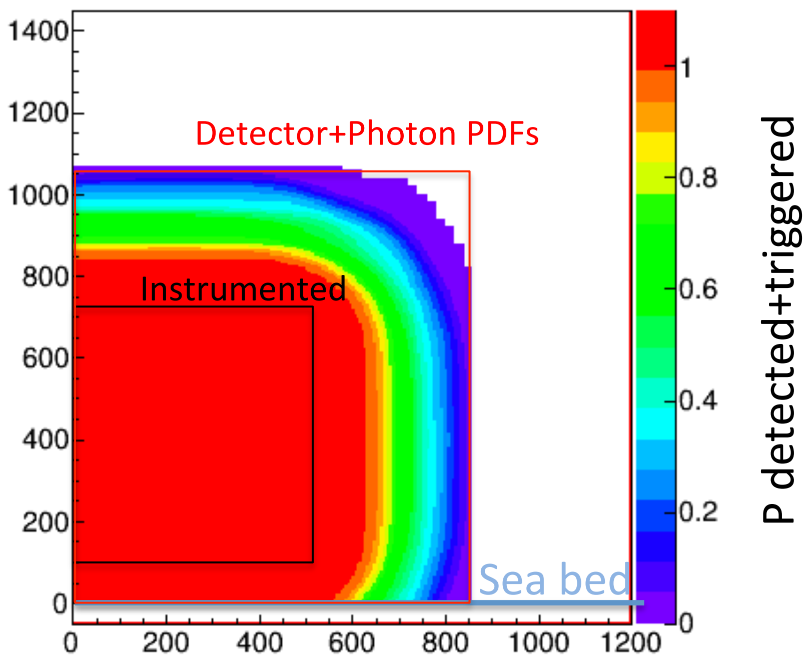
Direction @R=0, Z=400



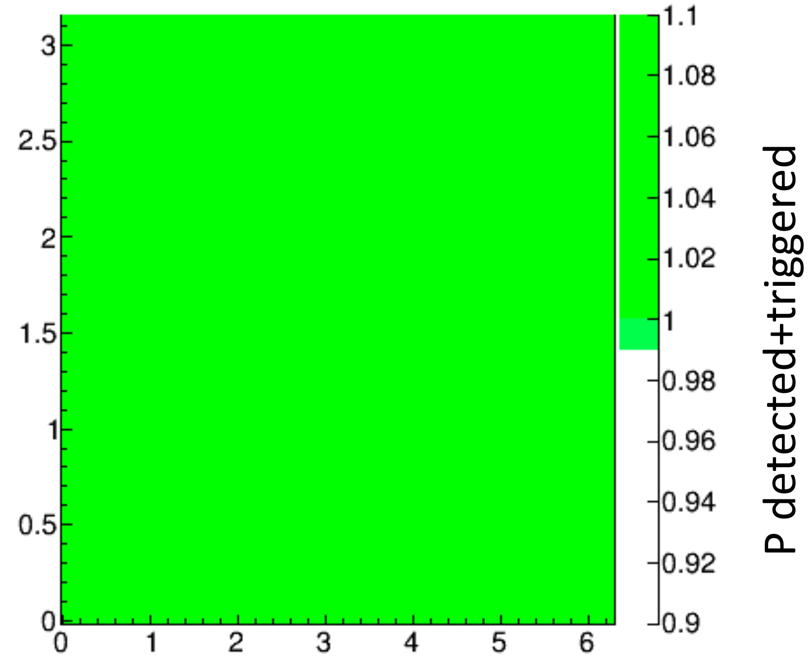
Detection Efficiency @ 10^6 GeV

NC electron-neutrino (only single hadronic shower)

Position



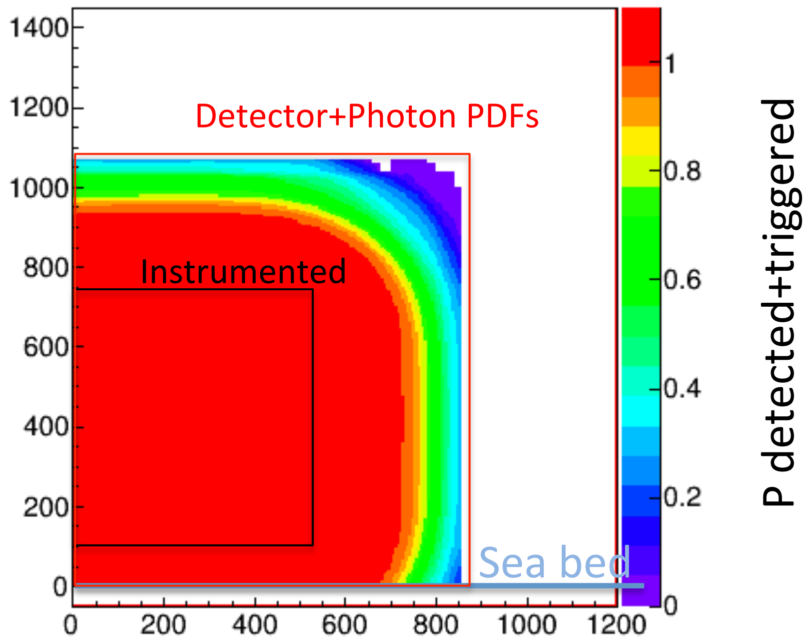
Direction @R=0, Z=400



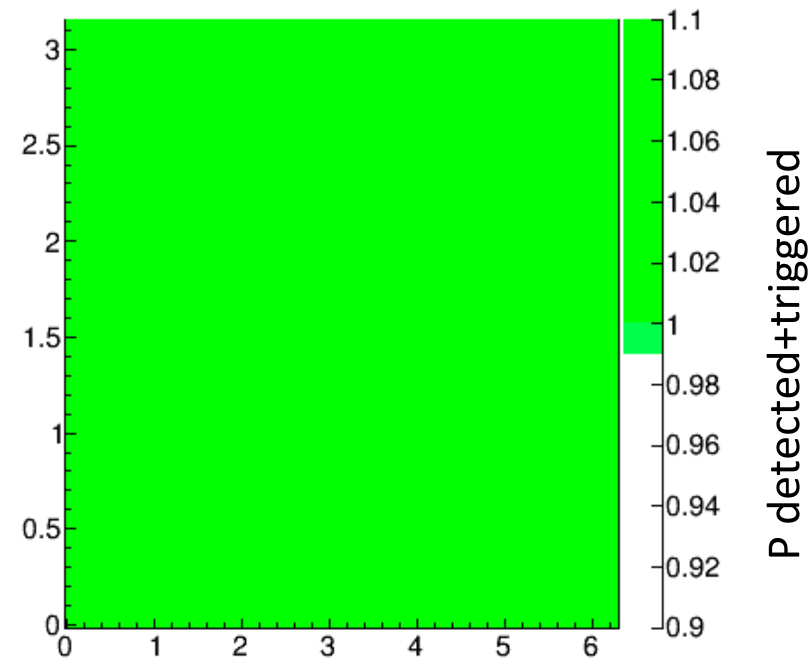
Detection Efficiency @ 10^7 GeV

NC electron-neutrino (only single hadronic shower)

Position



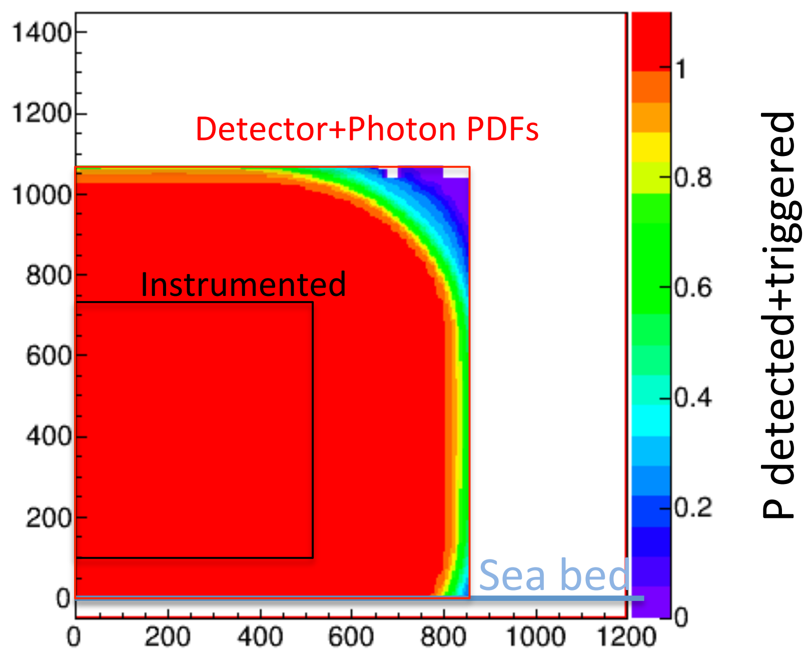
Direction @R=0, Z=400



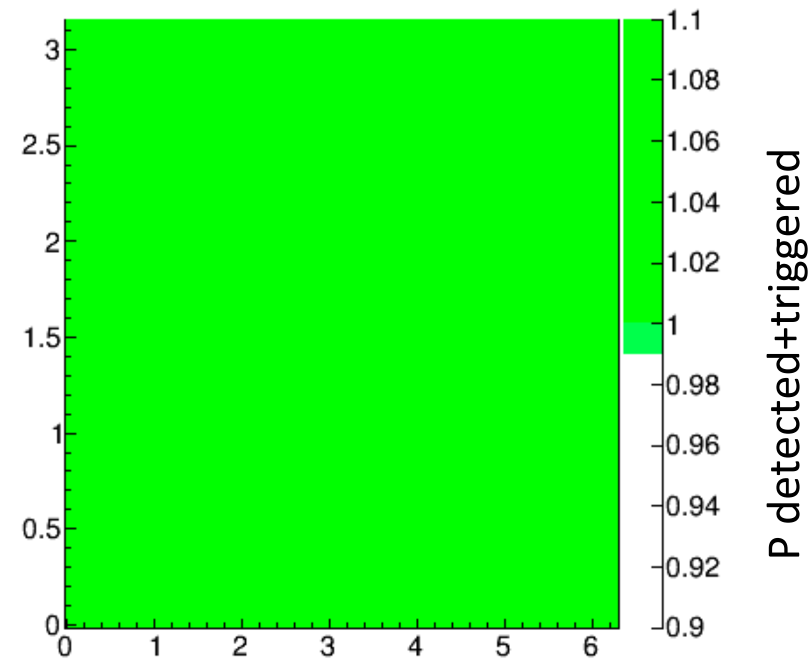
Detection Efficiency @ 10^8 GeV

NC electron-neutrino (only single hadronic shower)

Position



Direction @R=0, Z=400

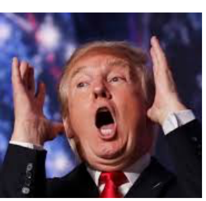


Likelihood Ingredients

$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true}\ |H)\ dx_{true}\right] - \mu^{tot}(H)$$

$\mu(x_{true}\ |H)$ Number of expected background or signal events in our detector (can)

$P^{det}(x_{true})$ Probability to detect (=trigger) and select event



$P(ev_i|\, x_{true})$ Probability to obtain measured event ev_i
given a certain neutrino hypothesis x_{true}

Event Probability D.F.

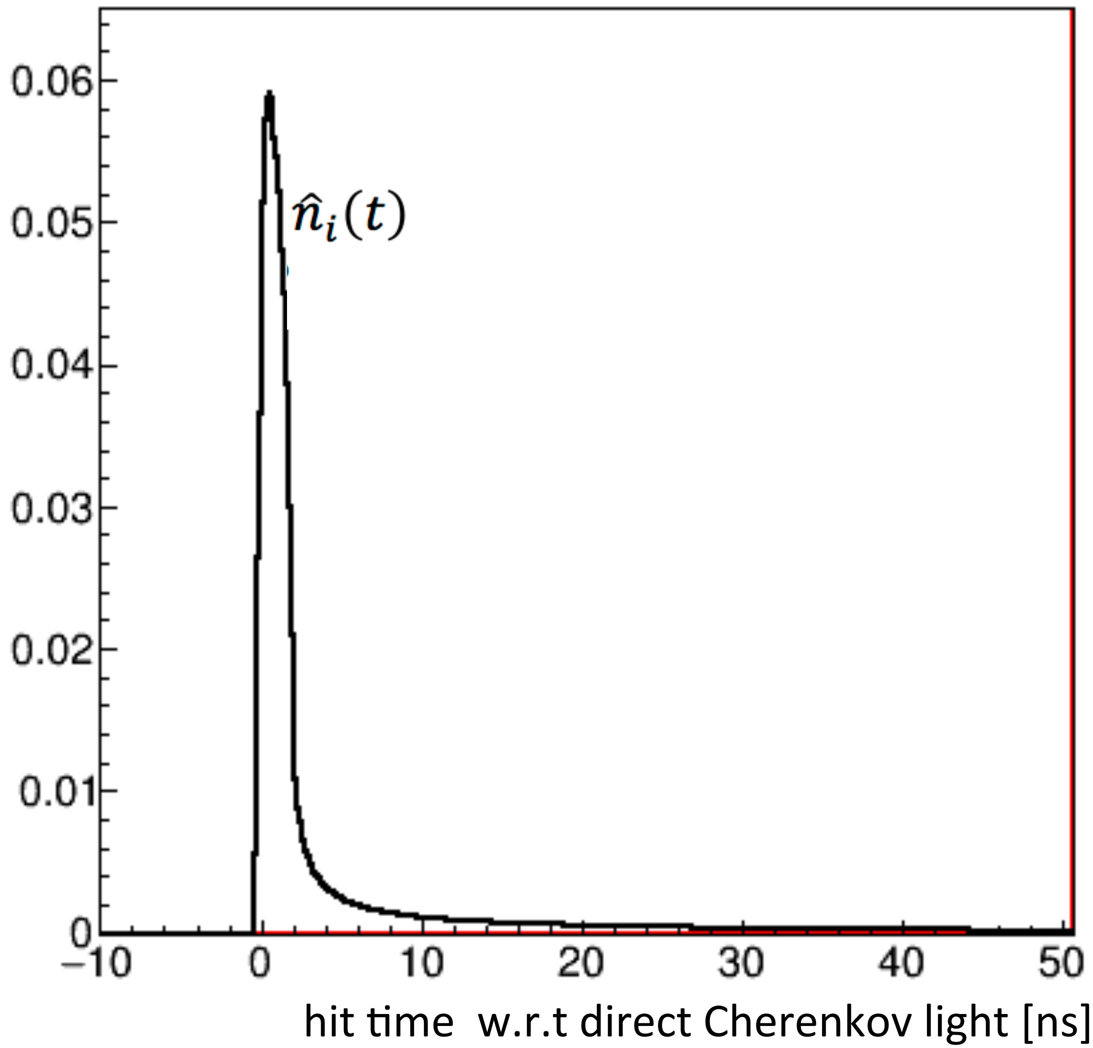
$$P(ev \mid x) = \prod_{hit\ PMTs} [P_i^{hit} \cdot P_i^{t\ 1st}] \cdot \prod_{non\ hit\ PMTs} [1 - P_i^{hit}]$$

$$P_i^{hit} = 1 - \exp\left(-\int_{-\infty}^{\infty} \hat{n}_i(t)\, dt\right)$$

Expected number of photons from 40K and shower/track on PMT i at time t

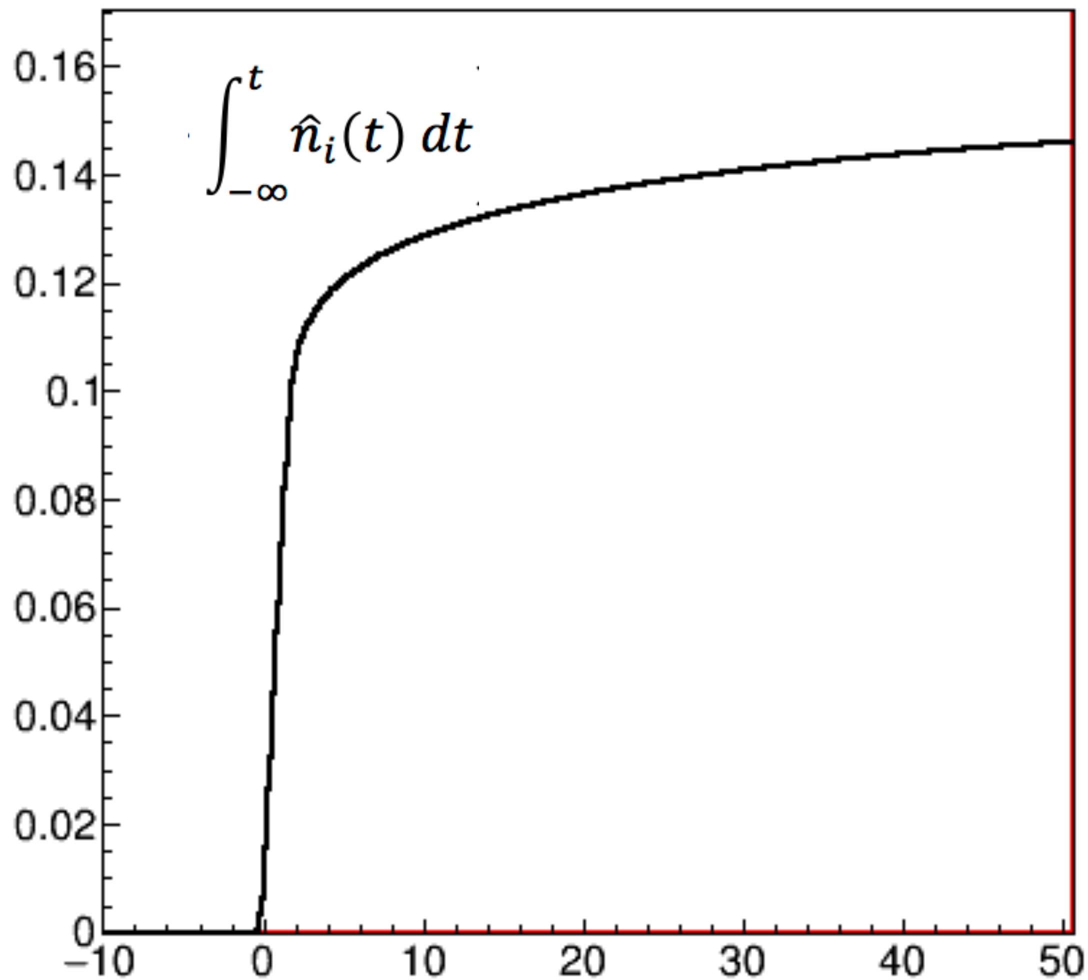
$$P_i^{t\ 1st} \cdot P_i^{hit} = \underbrace{\exp\left(-\int_{-\infty}^{t} \hat{n}_i(t)\, dt\right)}_{\text{P not hit before t}} \cdot \underbrace{(1 - \exp(-\hat{n}_i(t)))}_{\text{P hit at t}}$$

Hit Time PDF



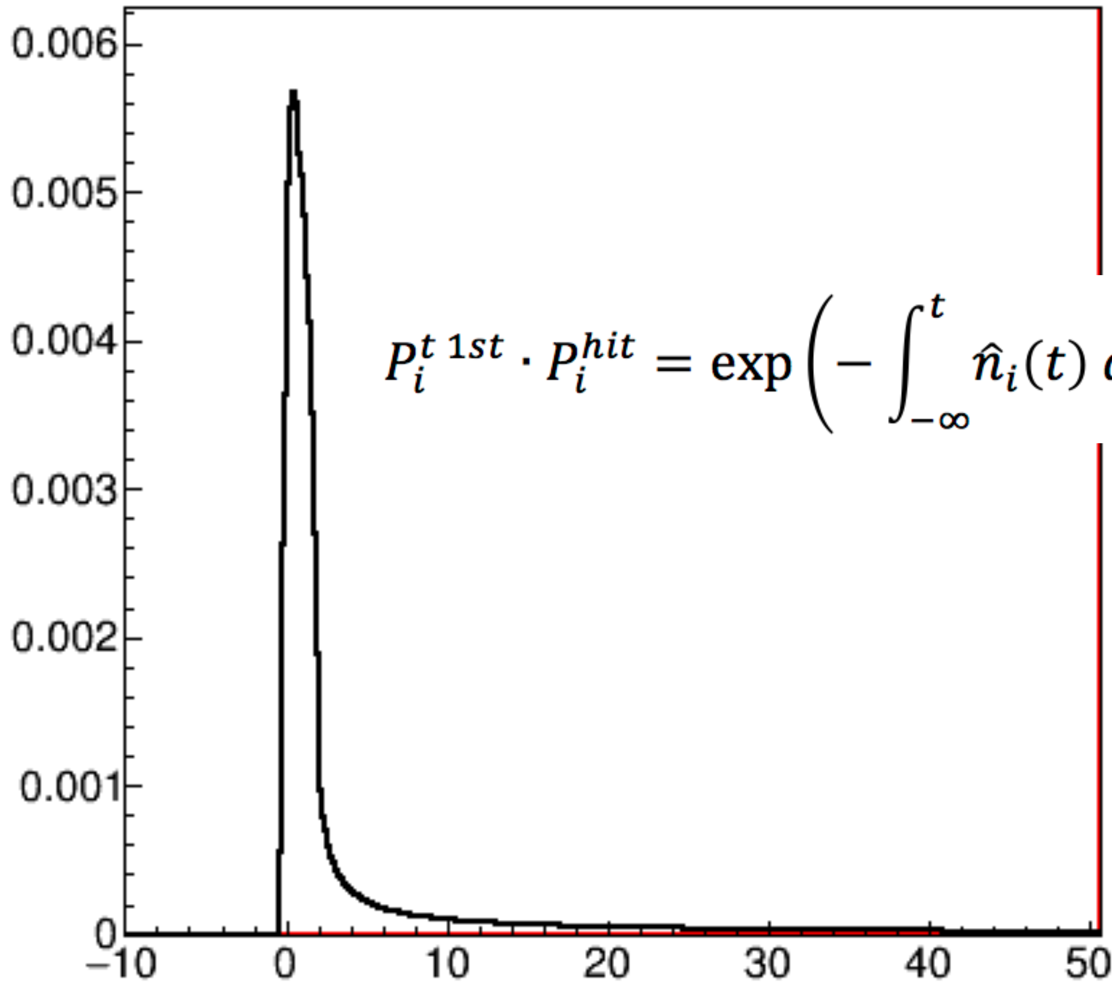
0.06
0.05
0.04
0.03
0.02
0.01
0
−10
0
10
20
30
40
50
$\hat{n}_i(t)$
hit time w.r.t direct Cherenkov light [ns]

Hit Time PDF



hit time w.r.t direct Cherenkov light [ns]

Hit Time PDF



$$P_i^{t\ 1st} \cdot P_i^{hit} = \exp\left(-\int_{-\infty}^{t} \hat{n}_i(t)\, dt\right) \cdot (1 - \exp(-$$

Hits in ~~Theory~~ Practice

- Presence of ^{40}K background hits

$$P_i^{t\ 1st} \cdot P_i^{hit} = \exp\left(-\int_{-\infty}^{t} \hat{n}_i(t)\, dt\right) \cdot (1 - \exp(-\hat{n}_i(t)))$$

- If all hit times are selected: signal will be overwhelmed by background

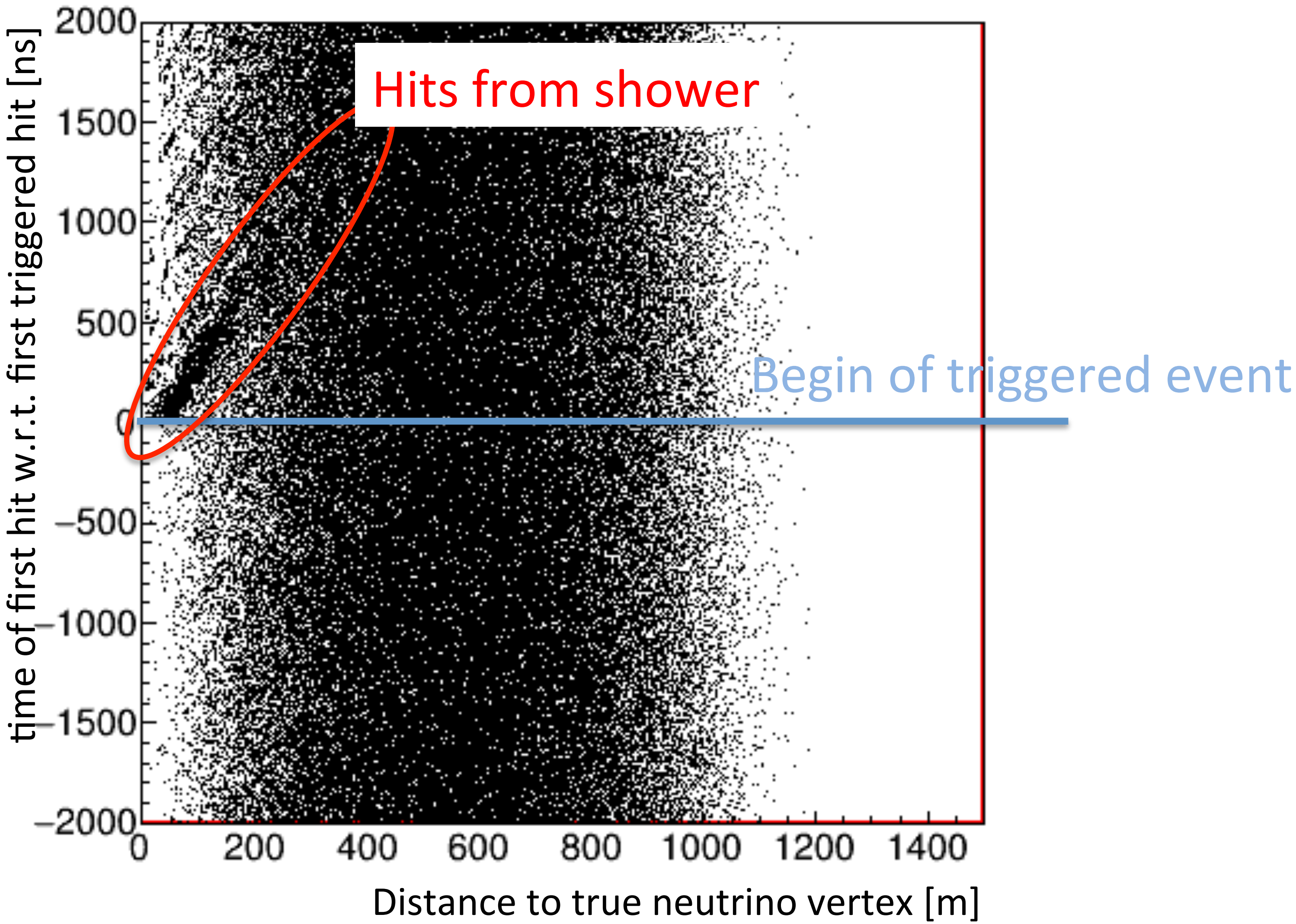
- Solution: only select hits in certain time window

Hit Selection Time Window

- Select Hits around expected hit time from given hypothesis
  - Advantage: Very pure selection
  - Drawback: biassed selection

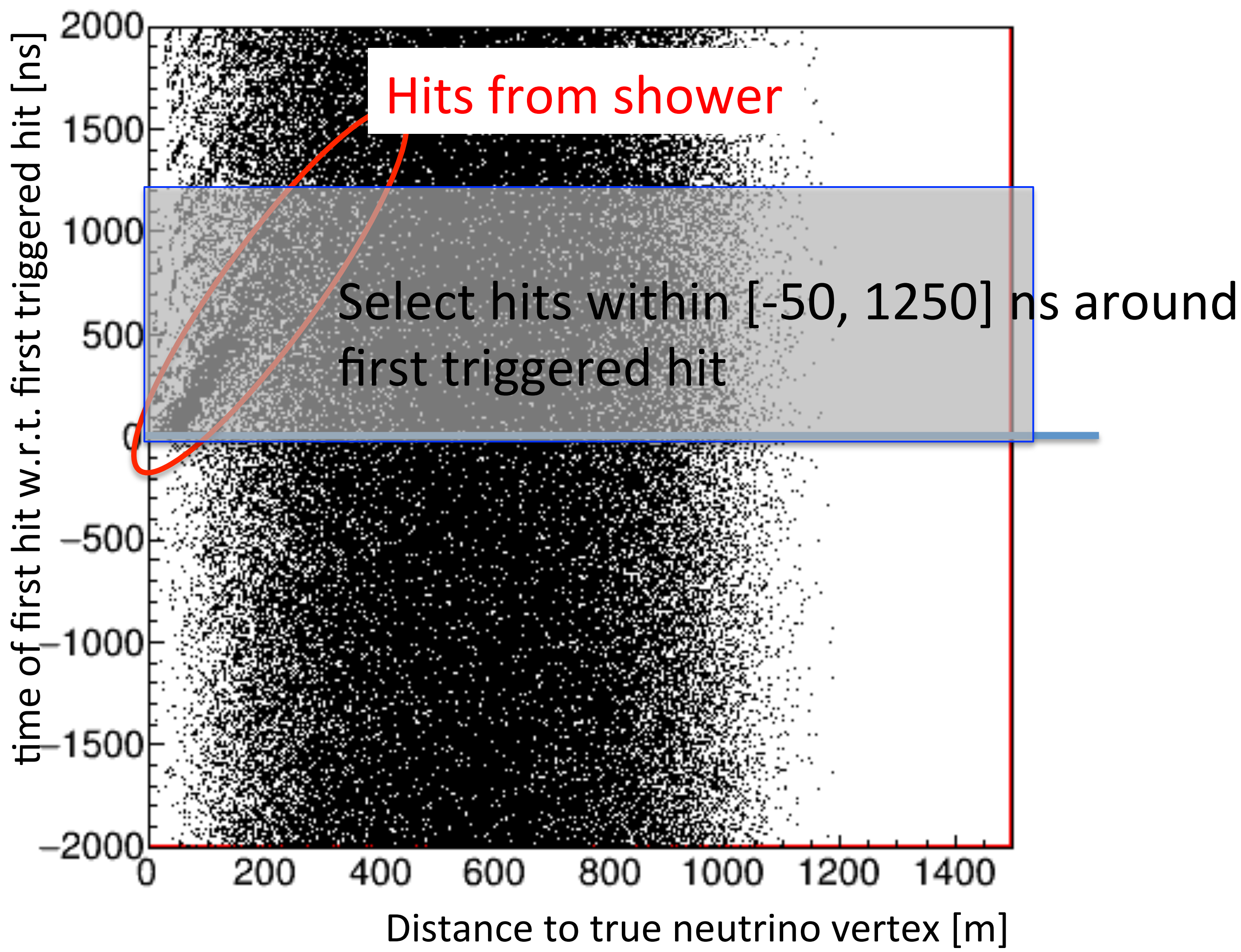
- Solution: Select hits around triggered event

Hit Times



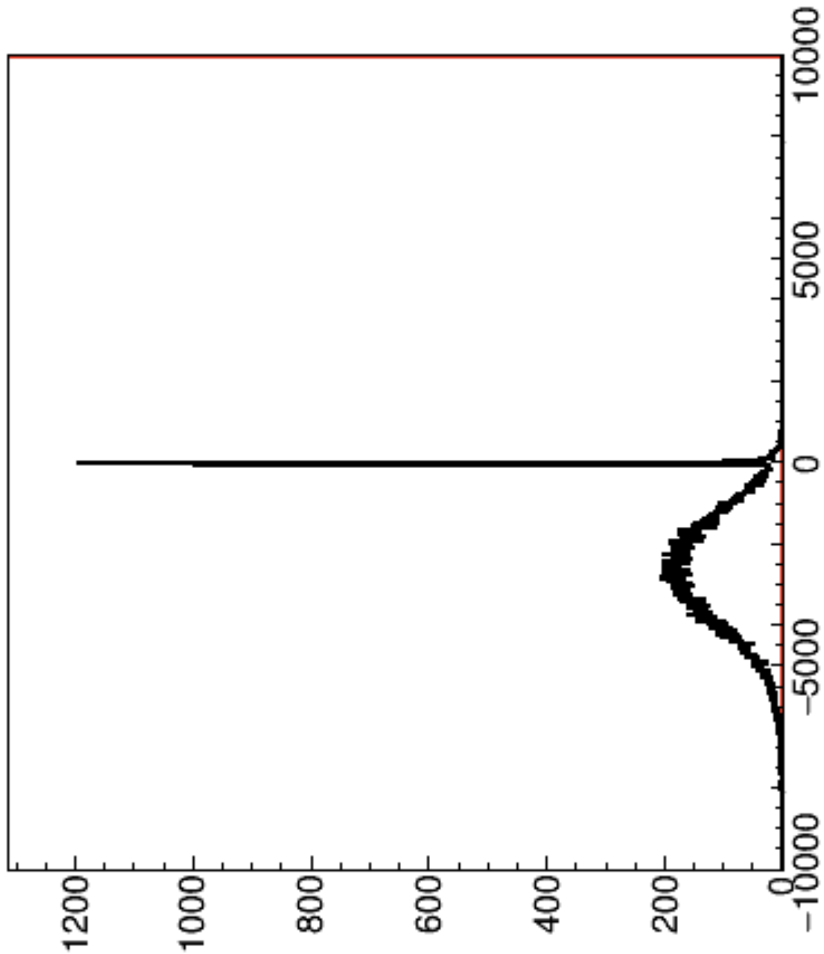
Hits from shower
Begin of triggered event
time of first hit w.r.t. first triggered hit [ns]
2000
1500
1000
500
0
−500
−1000
−1500
−2000
0
200
400
600
800
1000
1200
1400
Distance to true neutrino vertex [m]

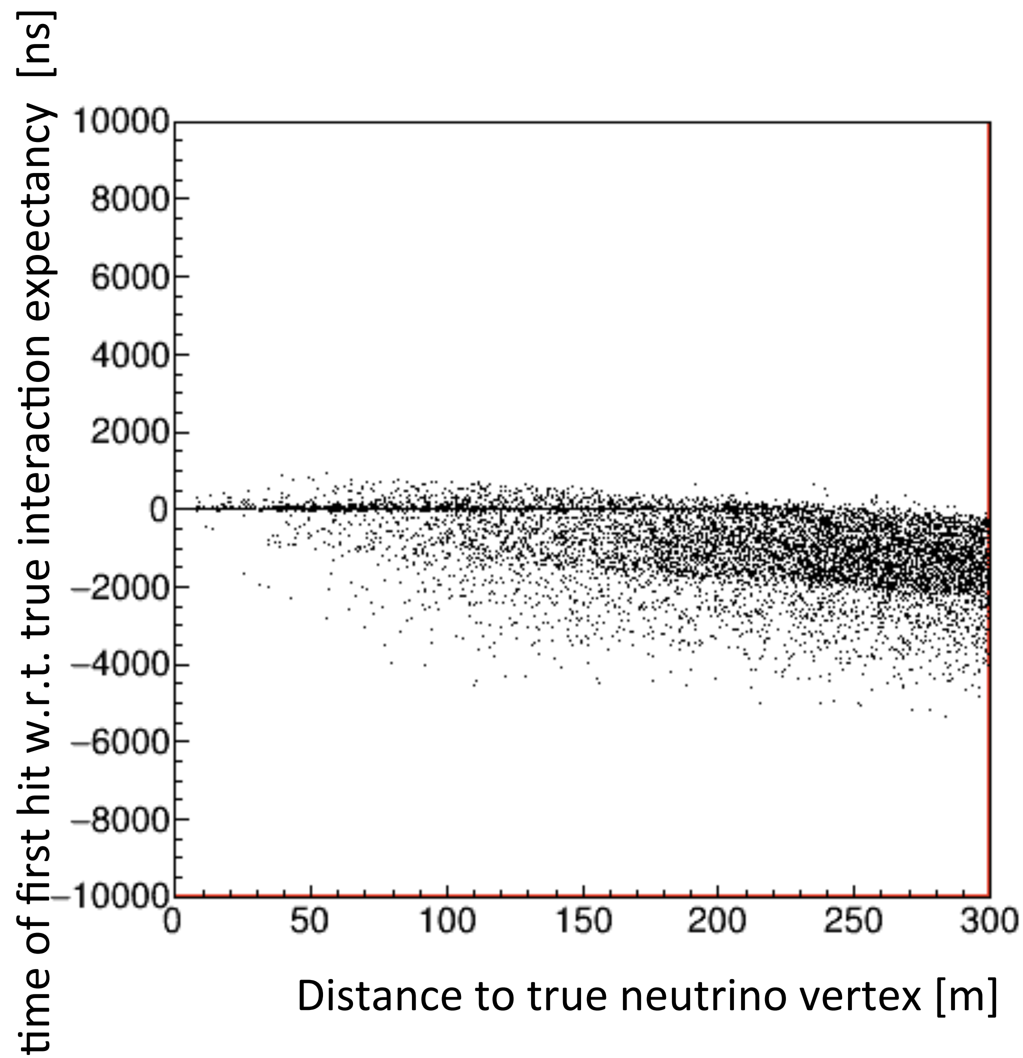
Hit Times



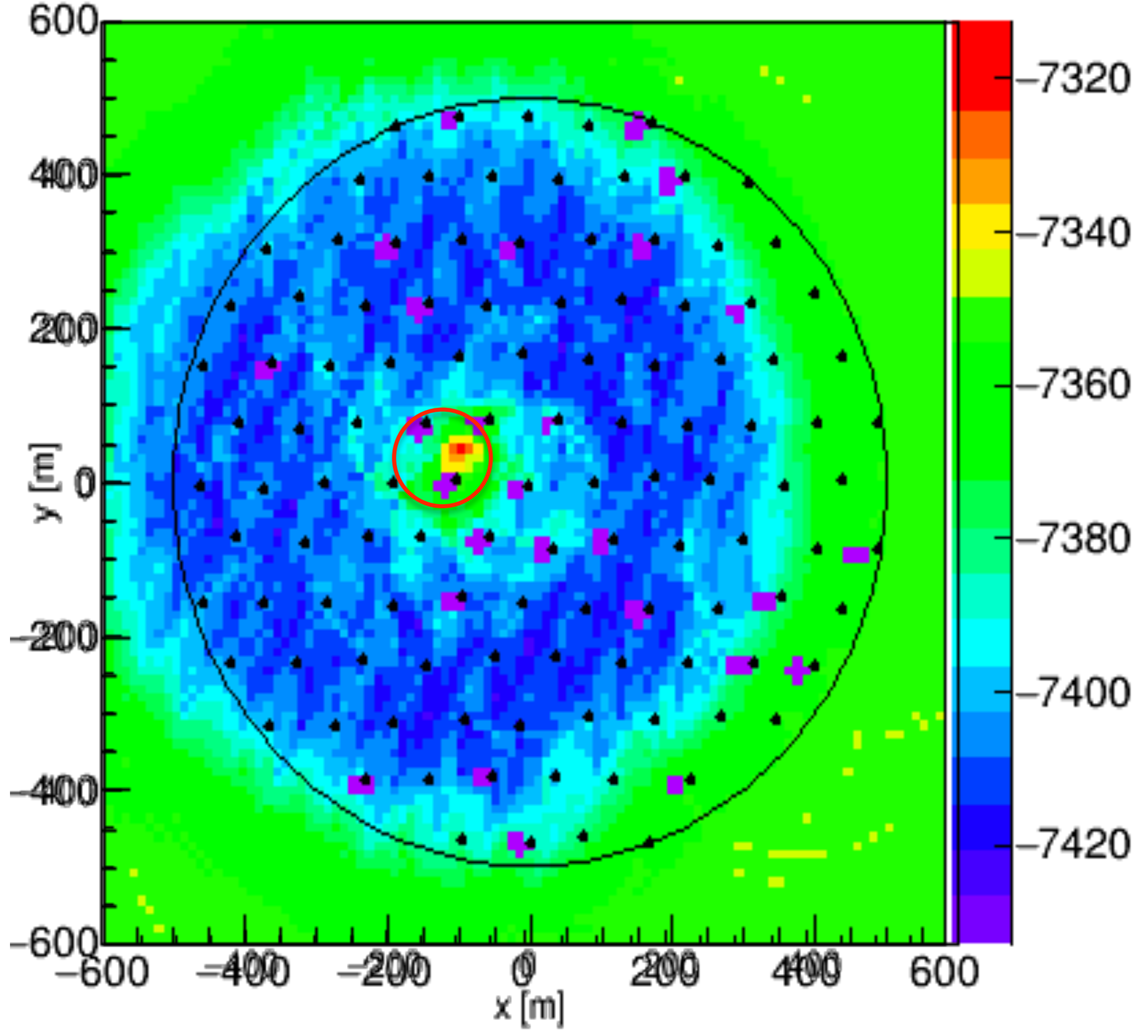
Hits from shower
Select hits within [-50, 1250] ns around first triggered hit
time of first hit w.r.t. first triggered hit [ns]
2000
1500
1000
500
0
−500
−1000
−1500
−2000
0
200
400
600
800
1000
1200
1400
Distance to true neutrino vertex [m]

Hit Times w.r.t. direct Cher. light





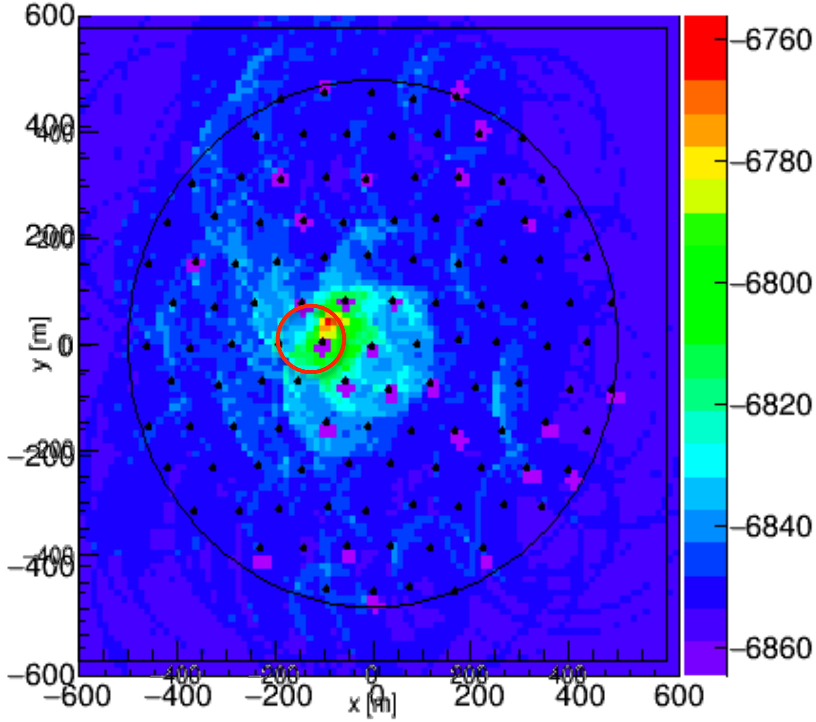
$$P(ev \mid x) = \prod_{hit\ PMTs} \left[P_i^{hit} \cdot P_i^{t\ 1st}\right] \cdot \prod_{non\ hit\ PMTs} \left[1 - P_i^{hit}\right]$$



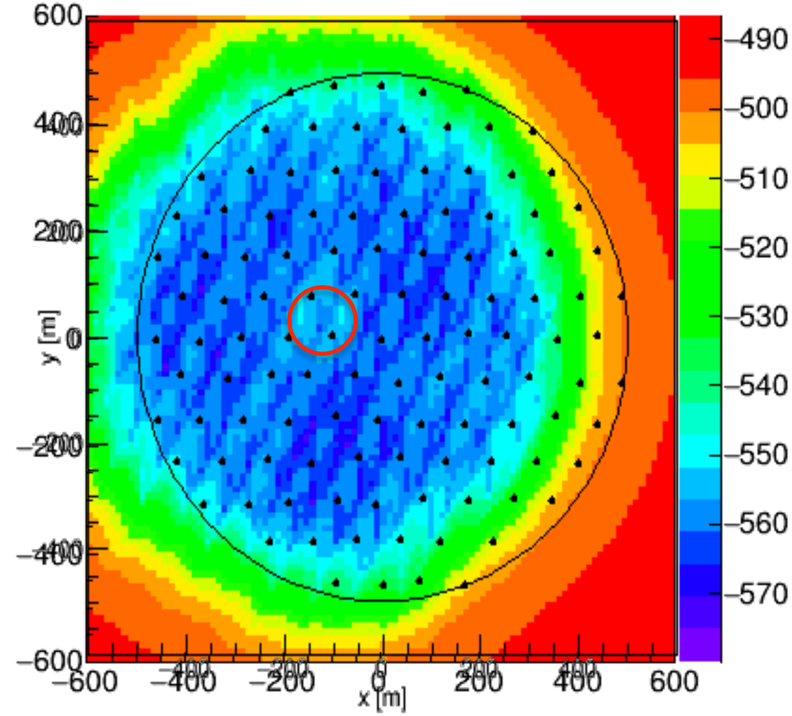
$logP(ev \mid x)$

$$logP(ev \mid x) = \sum_{hit\ PMTs} \left[\log\left(P_i^{hit}\right) + \log\left(P_i^{t\ 1st}\right)\right] + \sum_{non\ hit\ PMTs} \left[\log\left(1 - P_i^{hit}\right)\right]$$

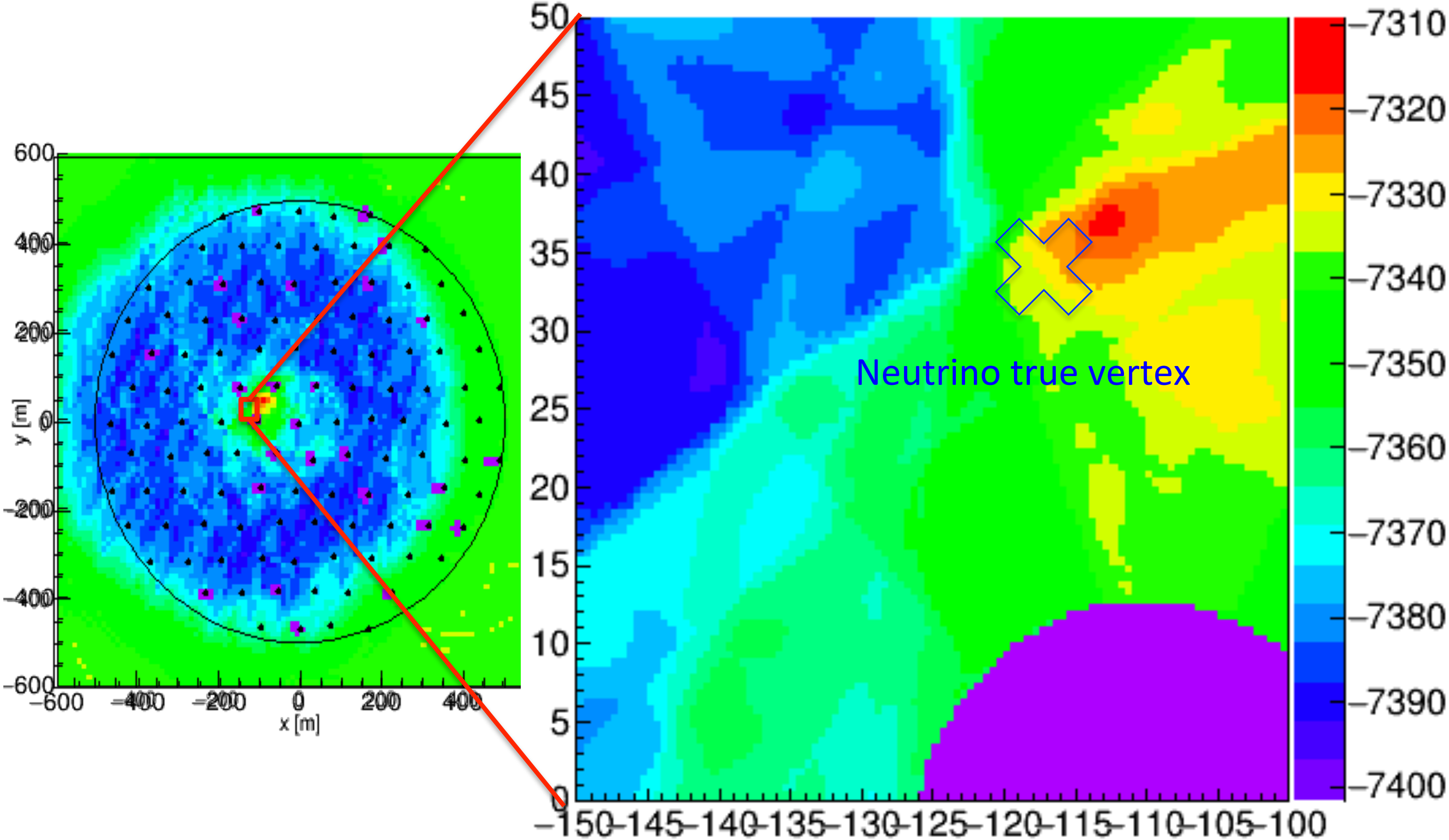
Hit PMTs



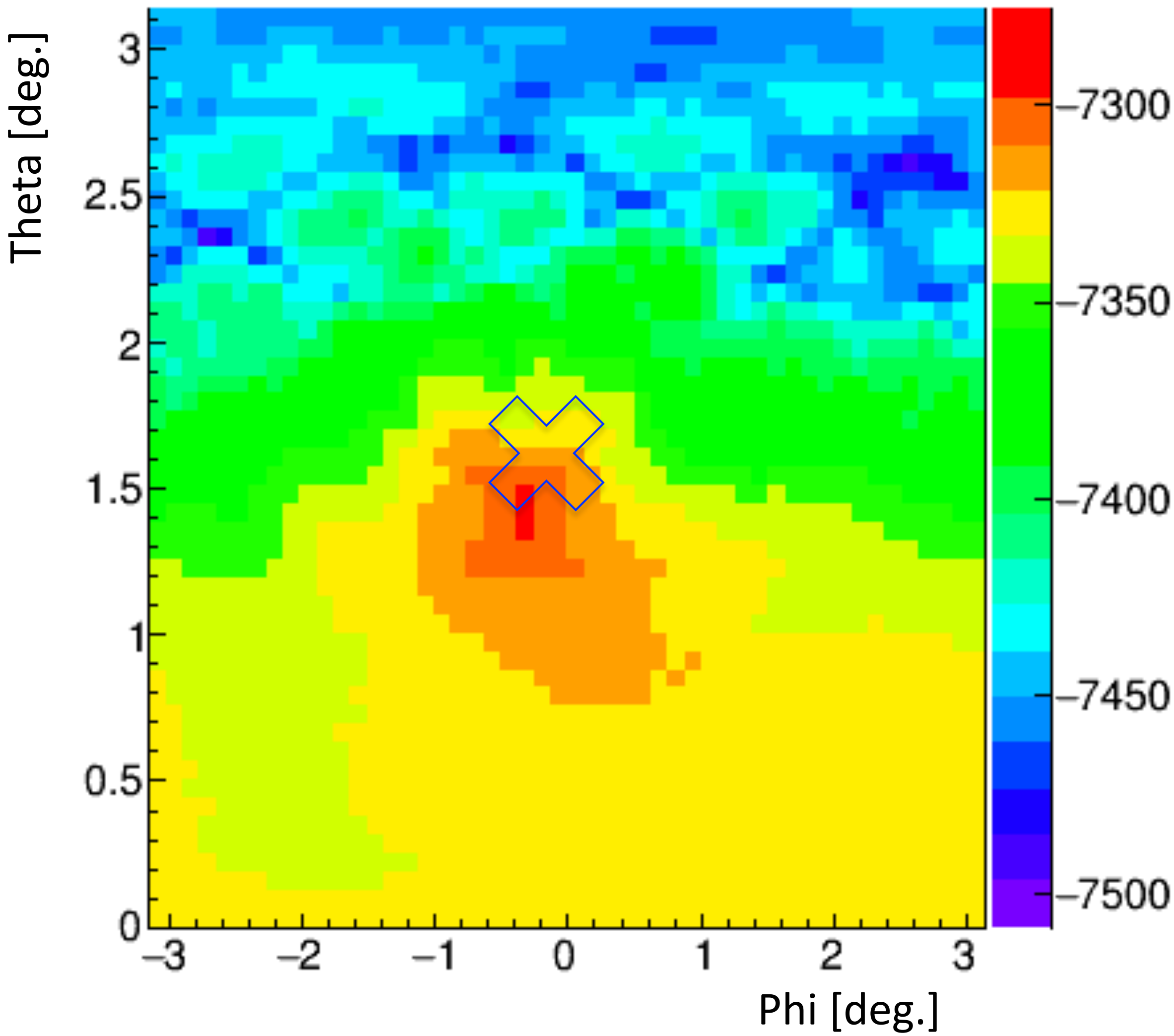
Not hit PMTs



Event Probability: Position



Event Probability: Direction



Likelihood Ingredients

$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true}\,|\,H)\, dx_{true}\right] - \mu^{tot}(H)$$

$\mu(x_{true}\,|\,H)$ Number of expected background or signal events in our detector (can)

$P^{det}(x_{true})$ Probability to detect (=trigger) and select event

$P(ev_i|\,x_{true})$ Probability to obtain measured event ev_i
given a certain neutrino hypothesis x_{true}

How to solve the 8D integral?

$$P(data|H) = \sum_i \left[\log \int P(ev_i |\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true} \mid H)\, dx_{true}\right] - \mu^{tot}(H)$$

- Interaction vertex position (3D)
- Interaction time (1D)
- (Neutrino) Direction (2D)
- Neutrino Energy (1D)
- Bjorken-y (1D)

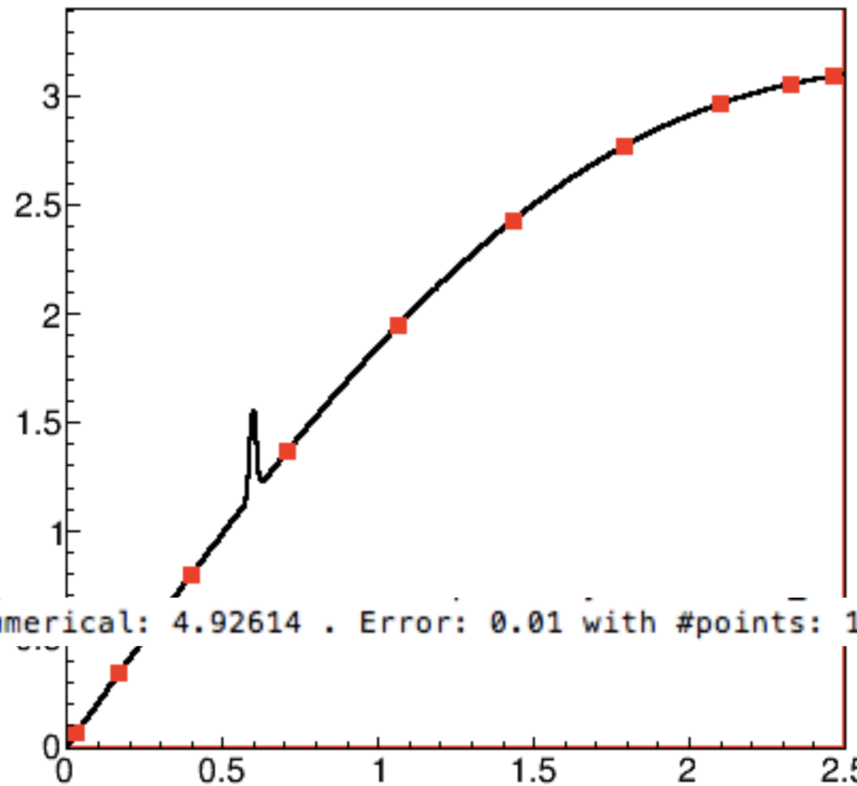
How to solve the ~~8D~~ 6D integral?

$$P(data|H) = \sum_i \left[\log \int P(ev_i|\, x_{true}) \cdot P^{det}(x_{true}) \cdot \mu(x_{true}\,|\, H)\, dx_{true}\right] - \mu^{tot}(H)$$

- Interaction vertex position (3)
- Interaction time (1)
  - Relatively easy once other params. are given?
- (Neutrino) Direction (2)
- Neutrino Energy (1)
  - Analytically?
- Bjorken-y (1)

Difficulties

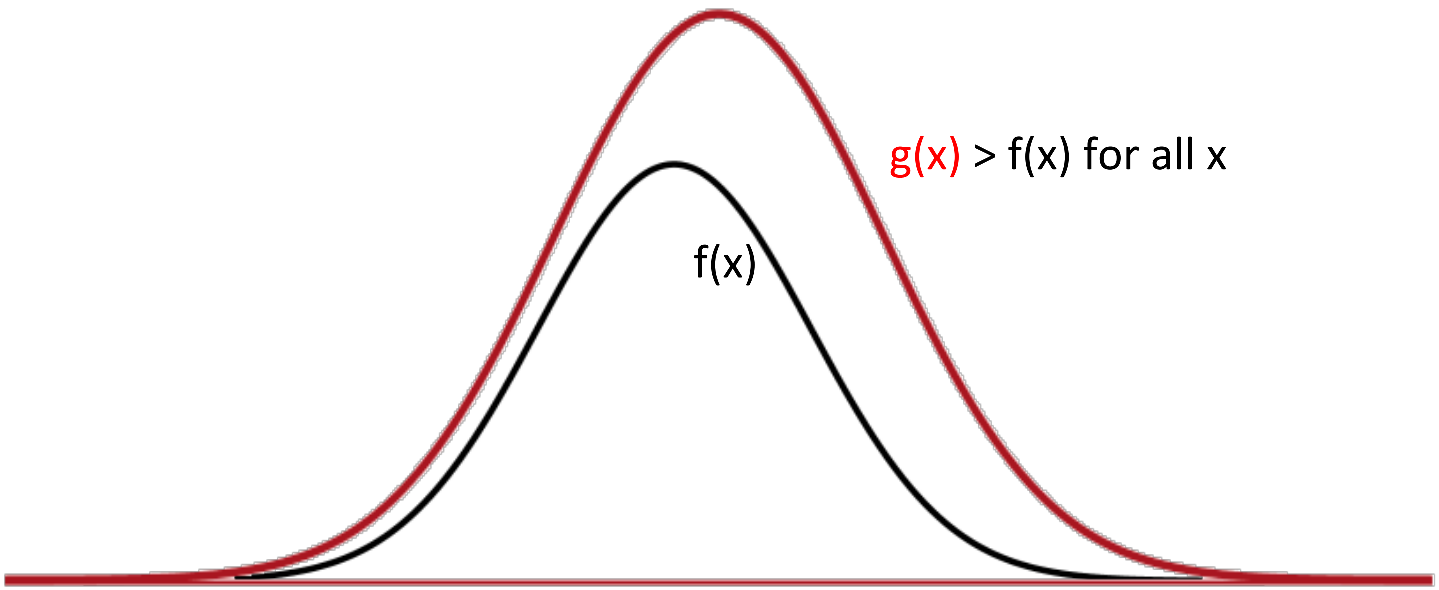
- Event PDF is in general sharply peaked
  - ~1 degree (showers), ~0.1 degree (tracks)
  - ~1m (showers)
- Algorithm generally misses this peak



integral, true: 4.93614 ,numerical: 4.92614 . Error: 0.01 with #points: 10

- Each function evaluation takes time

MC Integration Techniques (1)

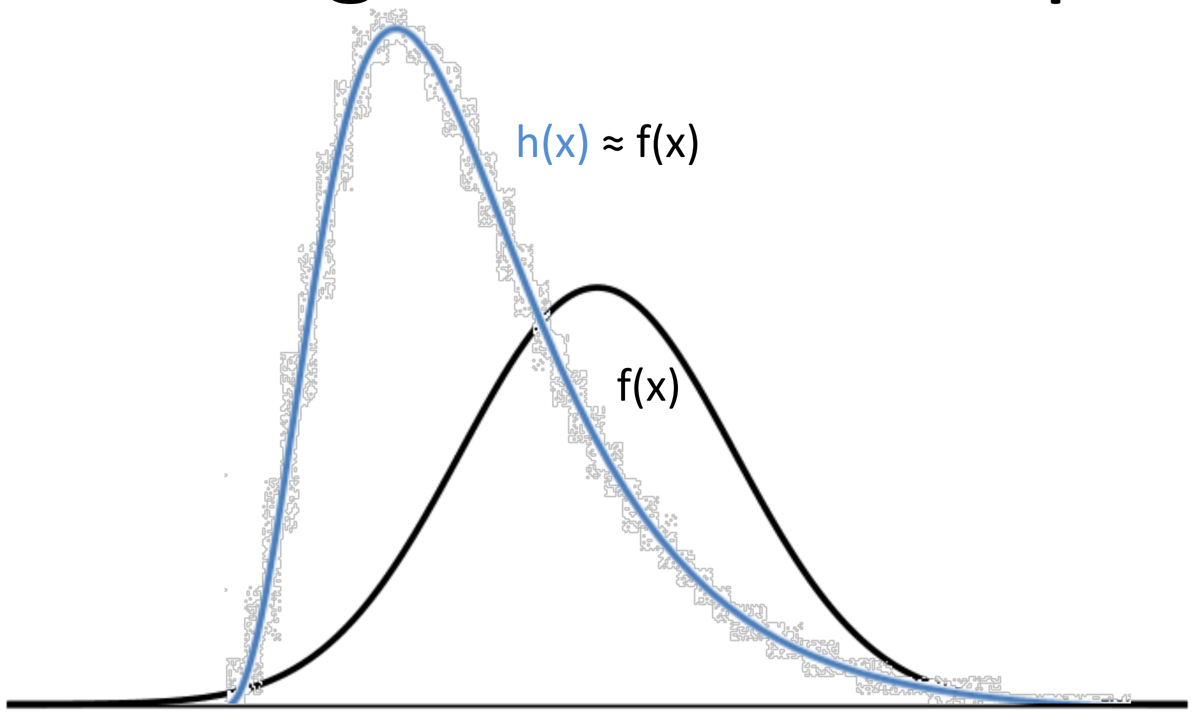


N times:

1) Random x from g(x)
2) Random y, 0<y<g(x)
3) If y <= f(x) { n++ }

I = int f(x) dx = n/N * int g(x) dx

MC Integration Techniques (2)

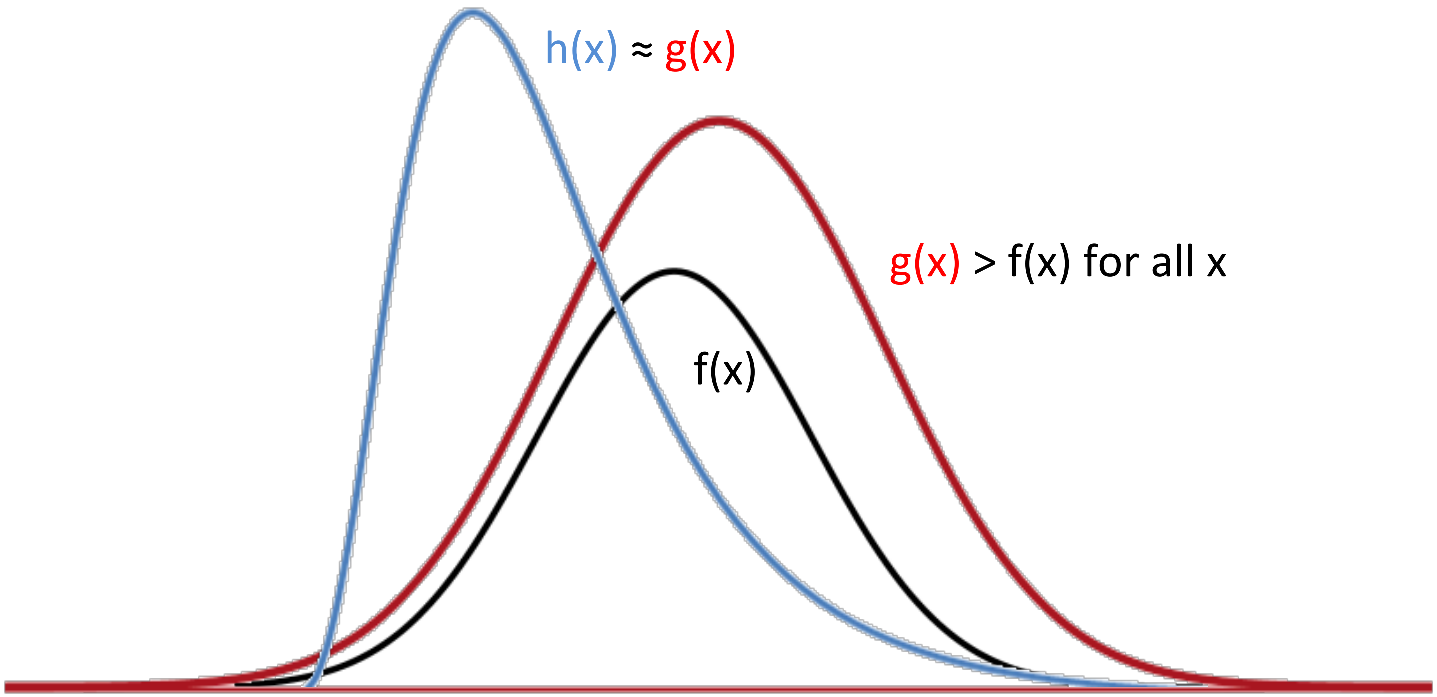


N times:

1) Random x from h(x)
2) A += f(x)/h(x)

I = int f(x) dx = A/N * int h(x) dx

Combined

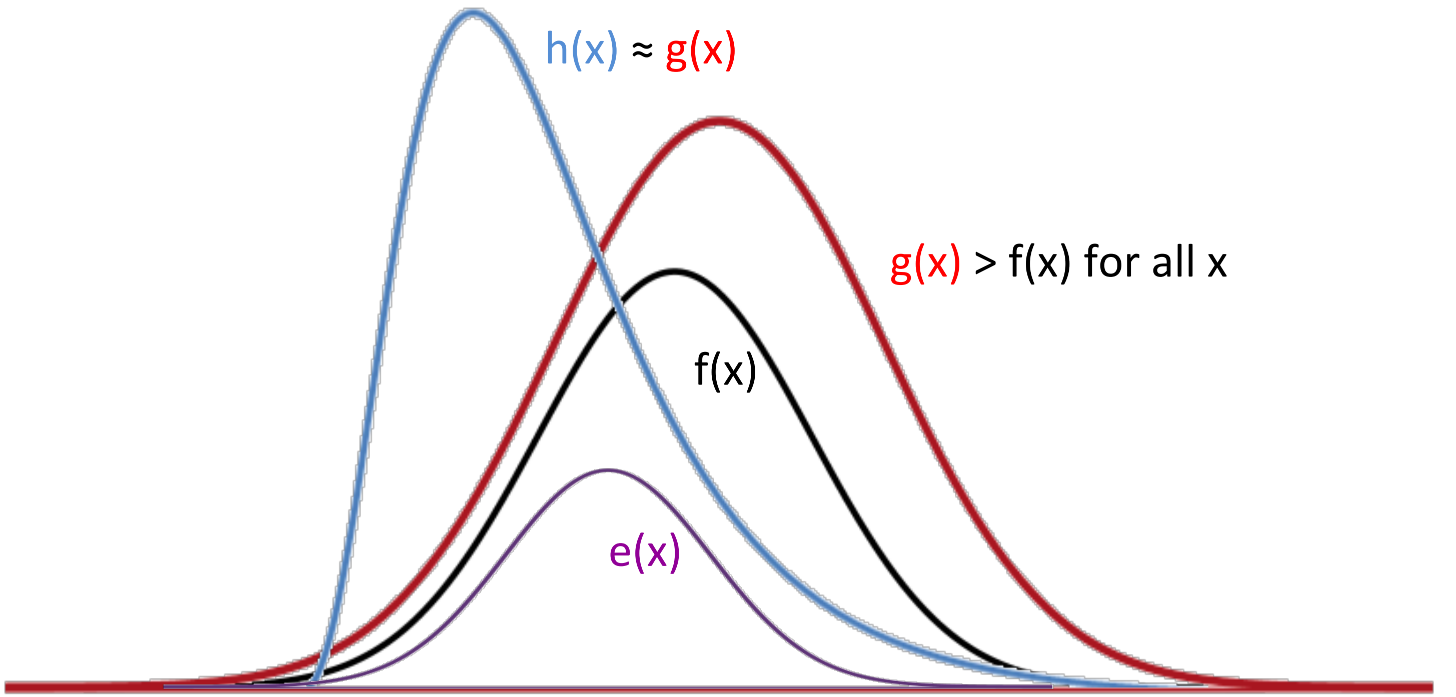


N times:

1) Random x from h(x)
2) A += g(x)/h(x)
3) Random y, 0<y<g(x)
4) If y <= f(x) { n++ }

I = int f(x) dx = n/N^2 * A * int h(x) dx

Combined + Extended

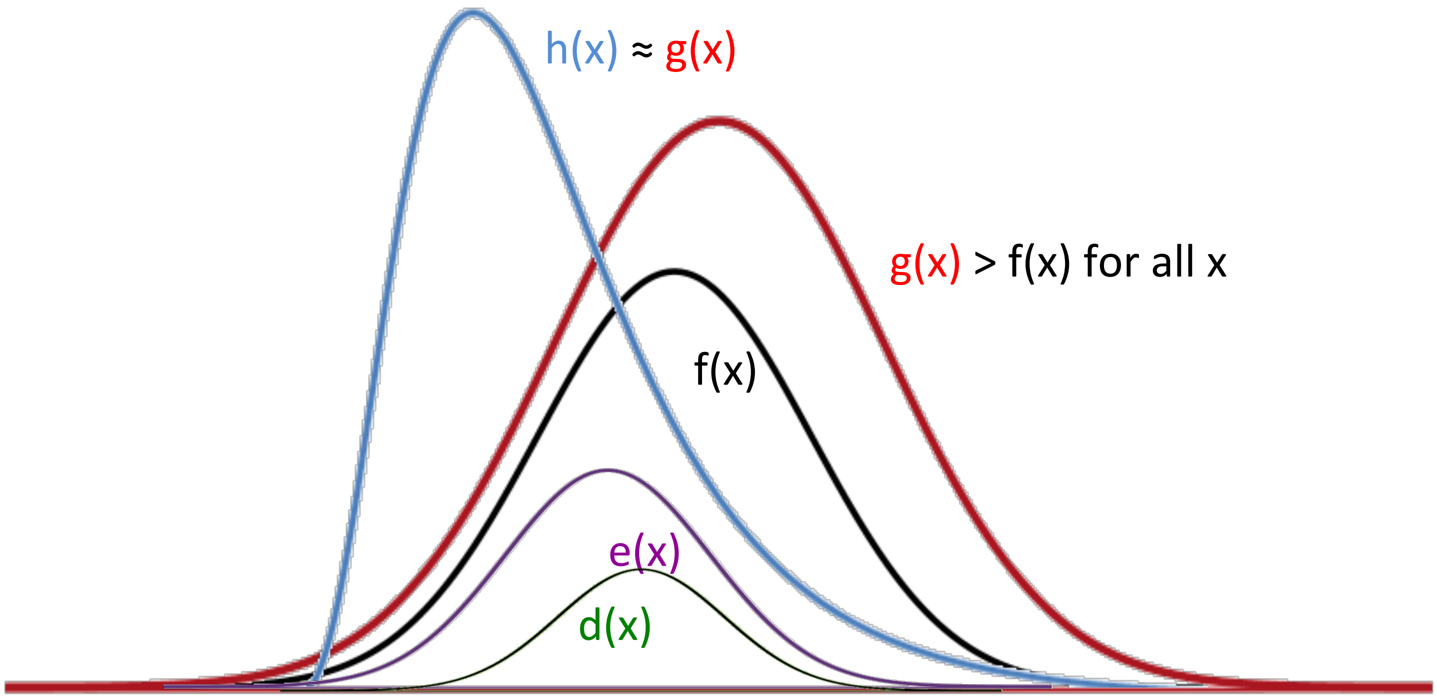


N times:

1) Random x from h(x)
2) A += g(x)/h(x)
3) Random y, 0<y<g(x)
4) If y <= f(x) { if y <= e(x) { n++ } }

$$I = \text{int } e(x)\, dx = n/N^2 * A * \text{int } h(x)\, dx$$

Combined + Extended

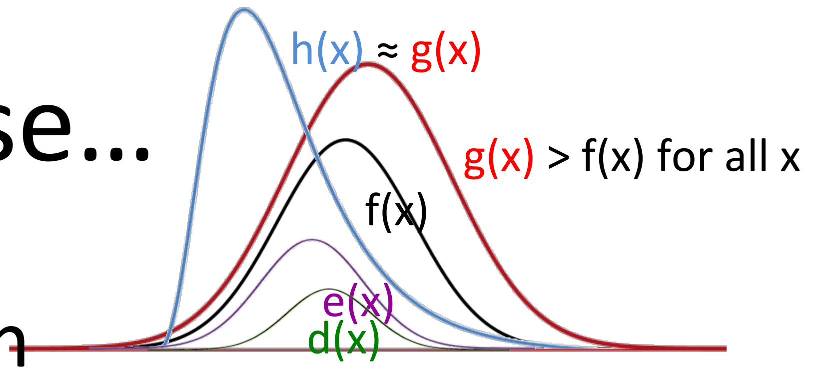


N times:

1) Random x from h(x)
2) A += g(x)/h(x)
3) Random y, 0<y<g(x)
4) If y <= f(x) { if y <= e(x) { if y <= d(x) { n++ } } }

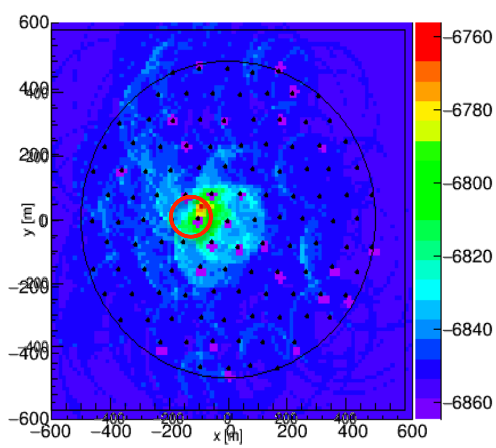
$$I = \int d(x)\,dx = n/N^2 * A * \int h(x)\,dx$$

In our case…

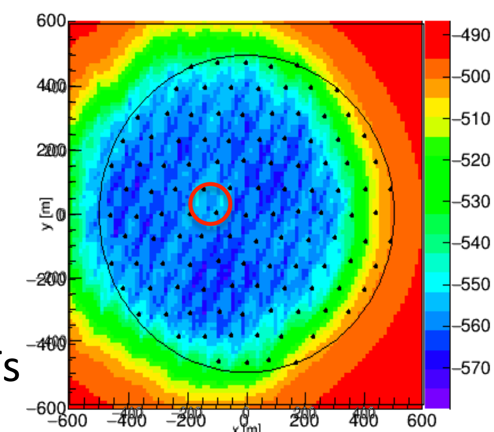


- h(x): some guiding function
- g(x): P(ev | x) over a small subset of PMTs
- f(x): P(ev | x) with slightly more PMTs
- e(x): P(ev | x) over a even more PMTs
- d(x): P(ev | x) with all PMTs

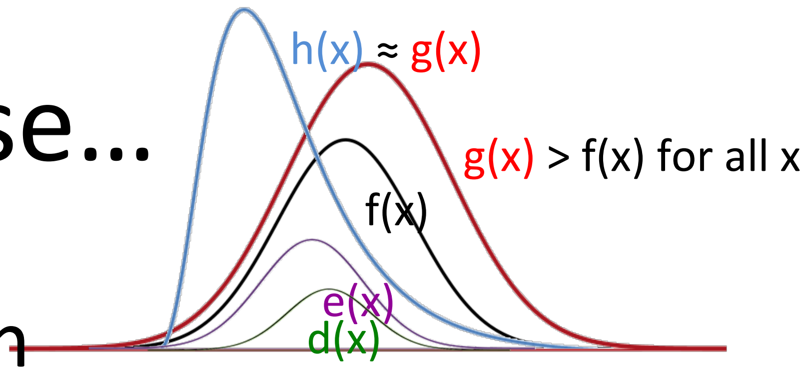
Hit PMTs



Non hit PMTs



In our case…



- h(x): some guiding function
- g(x): P(ev | x) over a small subset of PMTs
- f(x): P(ev | x) with slightly more PMTs
- e(x): P(ev | x) over a even more PMTs
- d(x): P(ev | x) with all PMTs

N times:

1) Random x from h(x)
2) A += g(x)/h(x)
3) Random y, 0<y<g(x)
4) If y <= f(x) { if y <= e(x) { if y <= d(x) { n++ } } }

I = int d(x) dx = n/N^2 * A * int h(x) dx

Guiding function

- Convenient choice: multivariate normal distribution
  - tracks: JPrefit PDF
  - Showers: ????

Guiding function

- Convenient choice: multivariate normal distribution
  - tracks: JPrefit PDF
  - Showers: ????

MASTER THESIS

Reconstruction of High-energy Neutrino-induced Particle Showers in KM3NeT.

BY

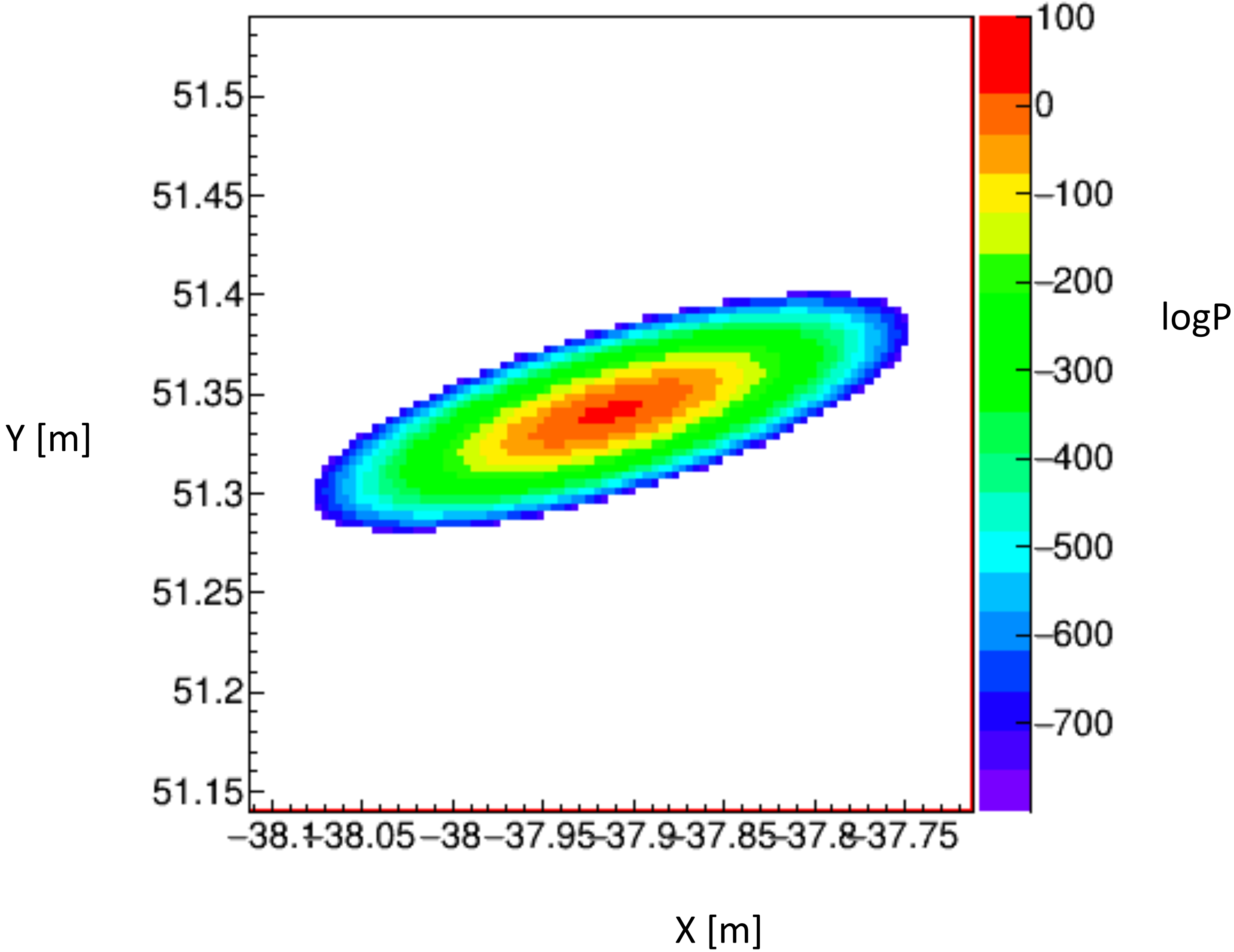
www.nikhef.nl/~kmelis/Masters_Thesis.pdf/Thesis.pdf

Shower Vertex PDF

- Basically a Chi^2 distribution
  - Very sensitive to outliers (i.e. ^{40}K hits)

- Use first triggered hit on each DOM
- Hit clustering algorithm:
  - If many hits: iteratively remove worst hit
  - If #hits <=16: Try all combinations

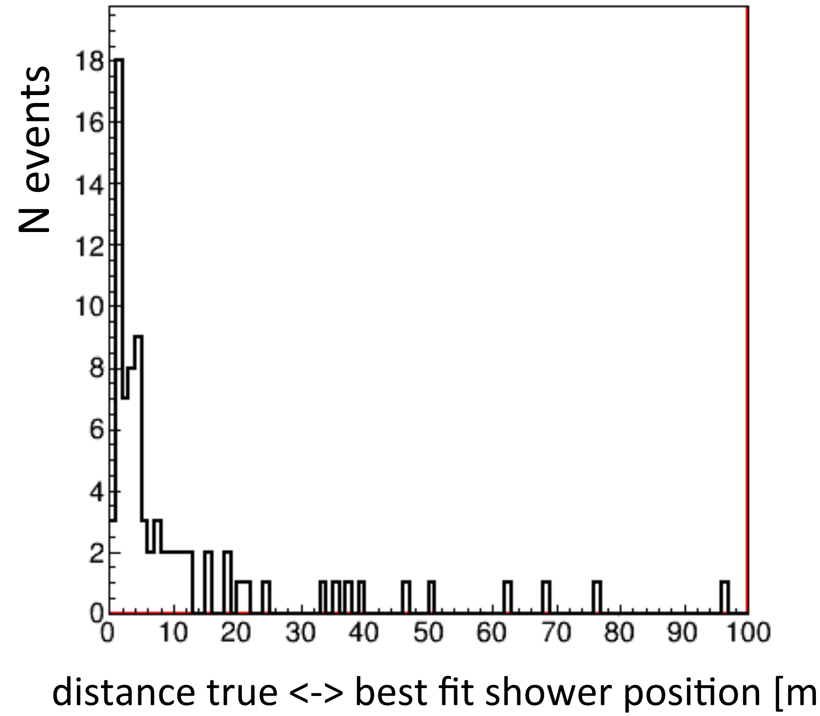
Shower Vertex PDF



Bonus: Shower Position Reconstruction

- Reasonable resolution:
  - median ~1m
- Principle (cluster+ chi^2) usable for tau double bang prefit?

Preliminary



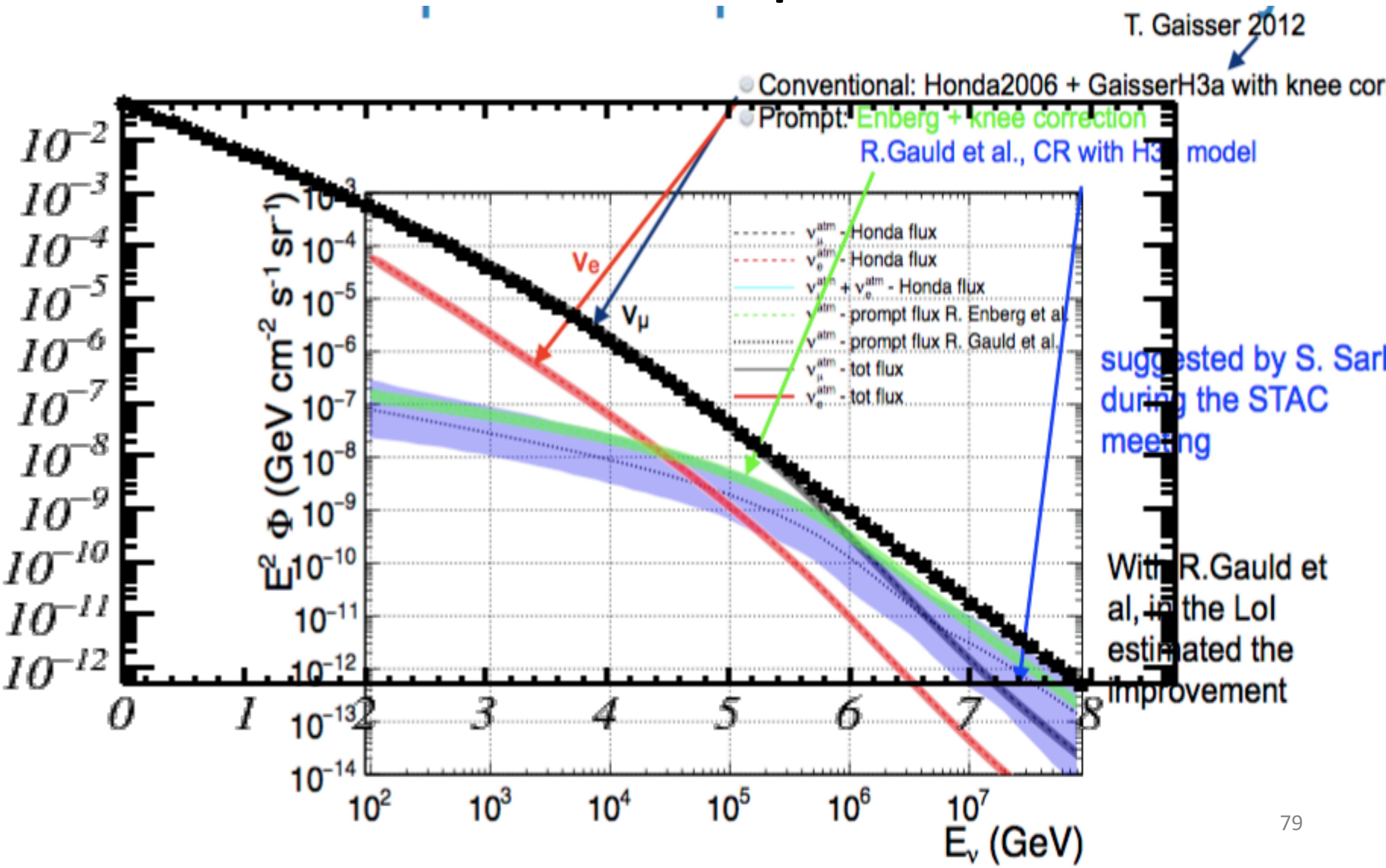
BONUS bij AH Willems

Conclusions

- Most ingredients seem 
  - Neutrino background and signal fluxes
  - Detection efficiency tables
  - Event probability
- Integral evaluation seems feasible with MC techniques
- BONUS bij AH Willems :
  - Shower prefit (+tau double bang prefit?)
  - Very fast neutrino MC simulator

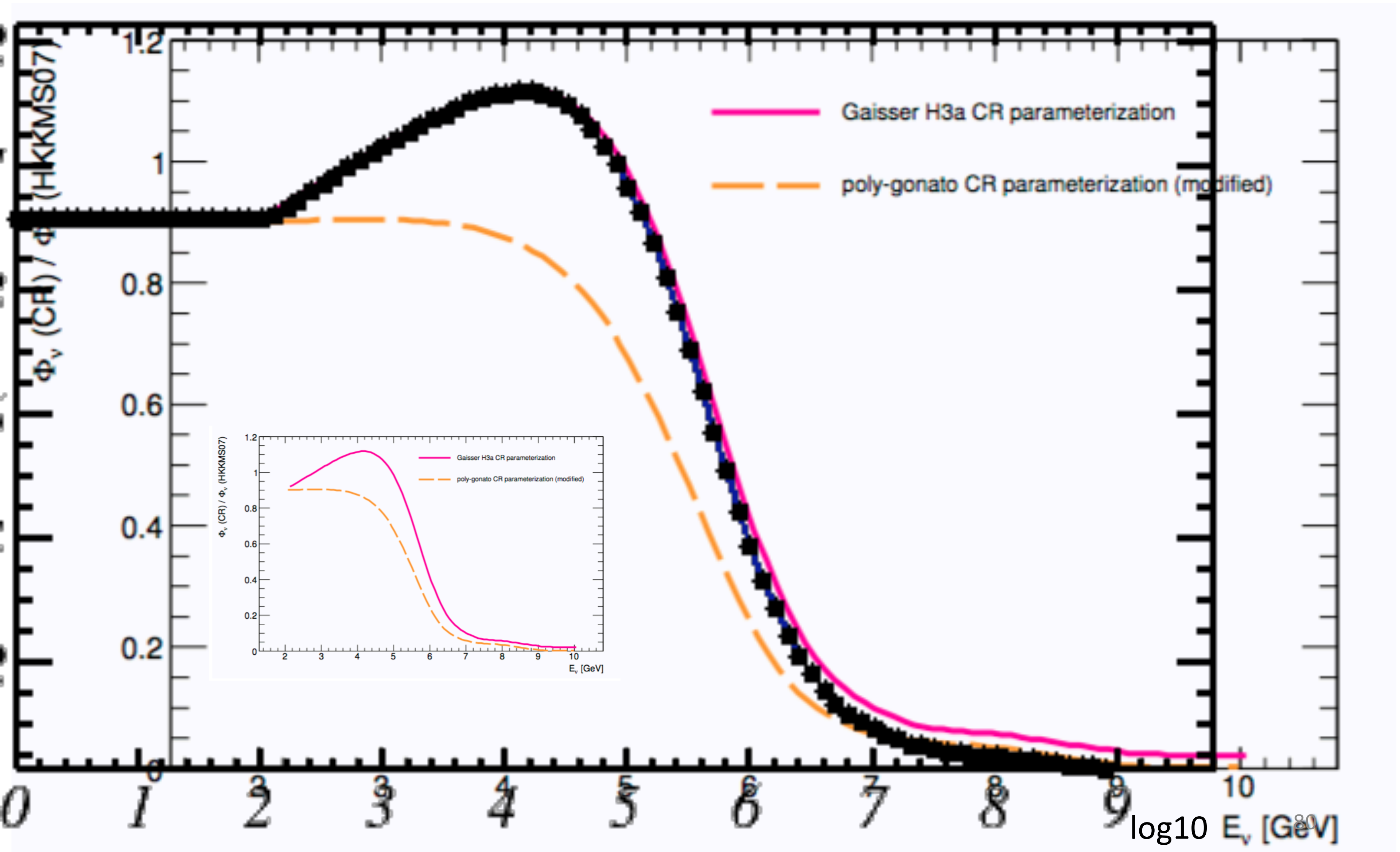
Backup

Honda extrapolated

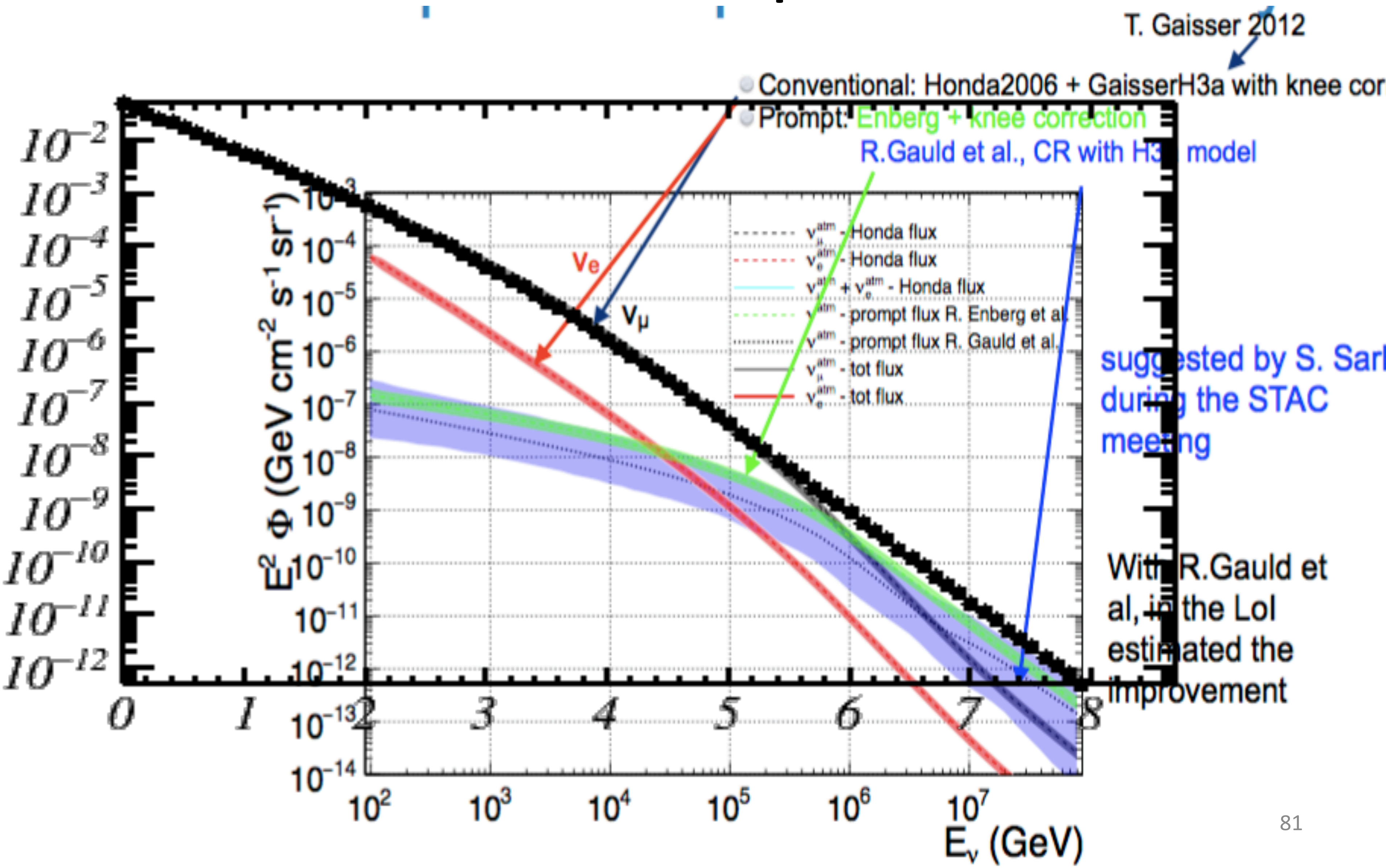


arXiv:1311.7048

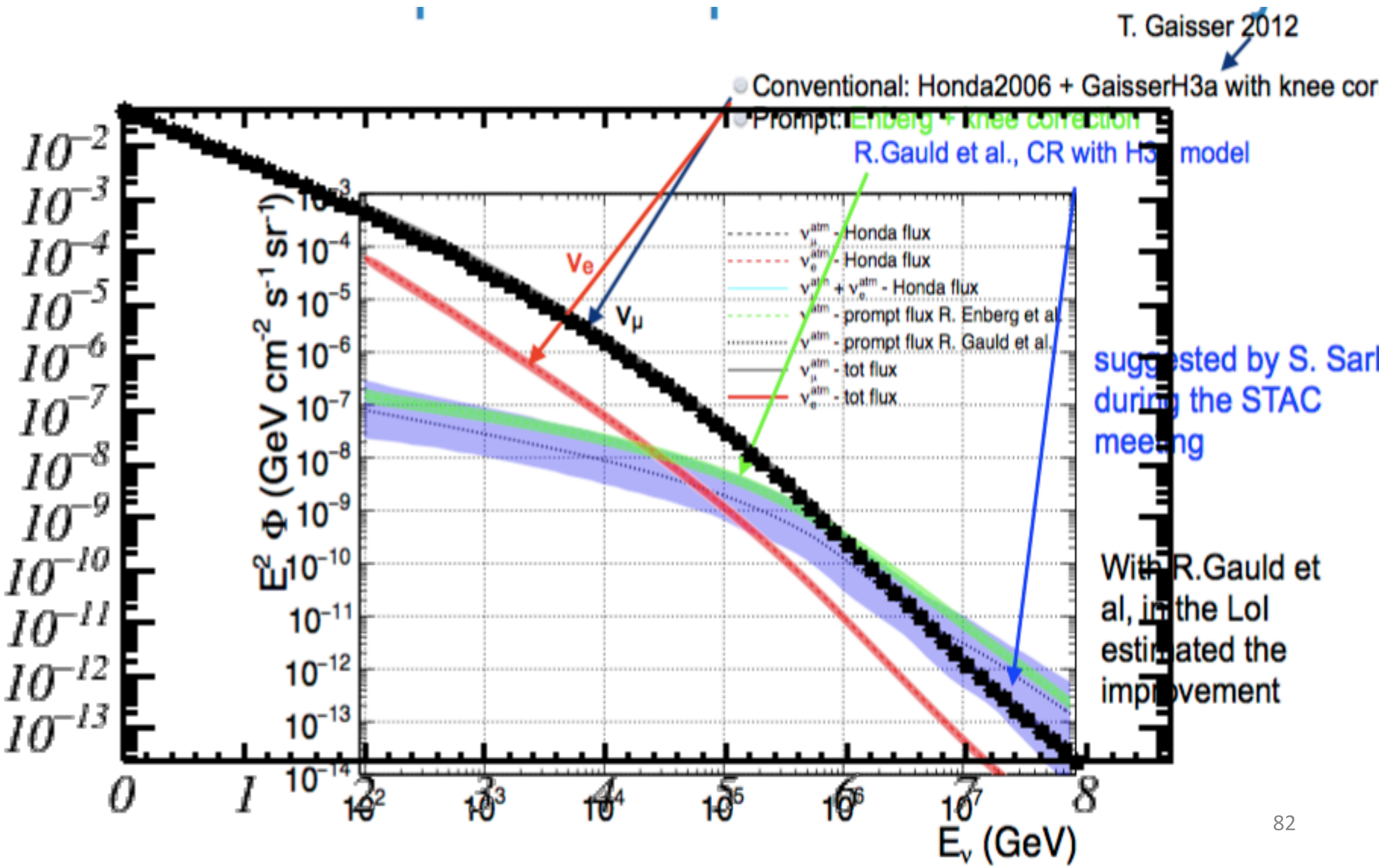
Knee Correction (Gaisser H3a)



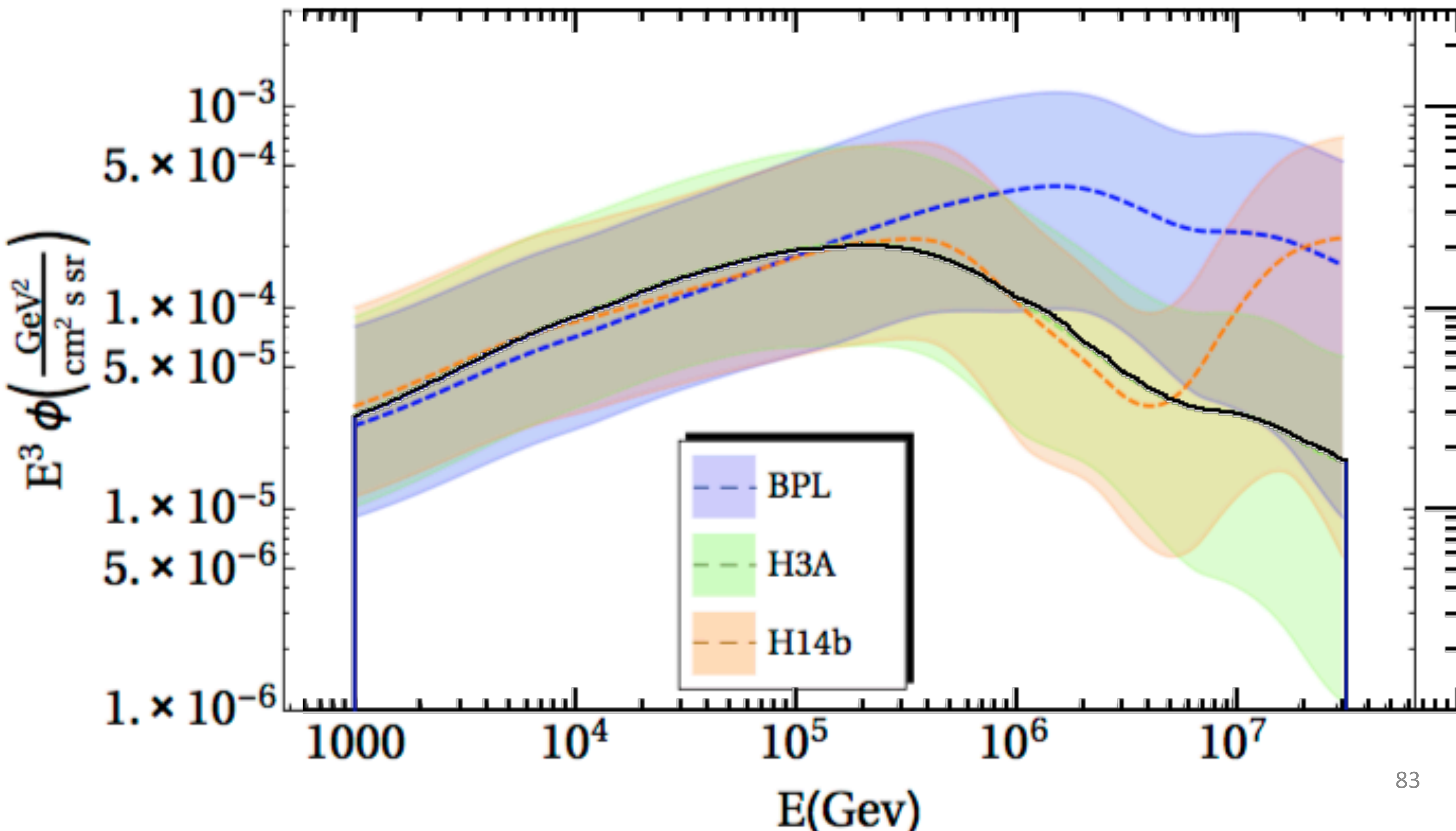
Honda extrapolated



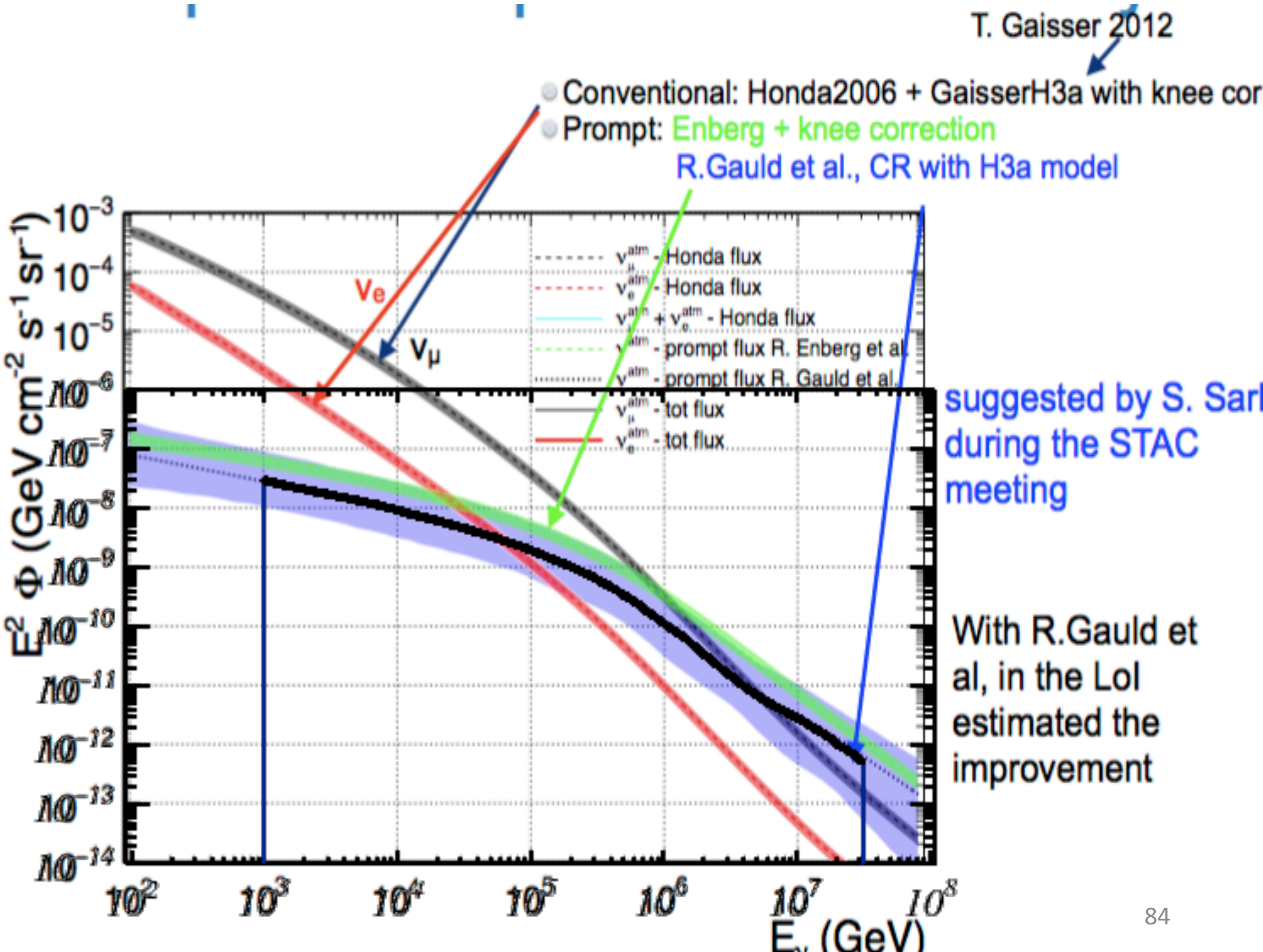
Honda extrapolated + knee correction



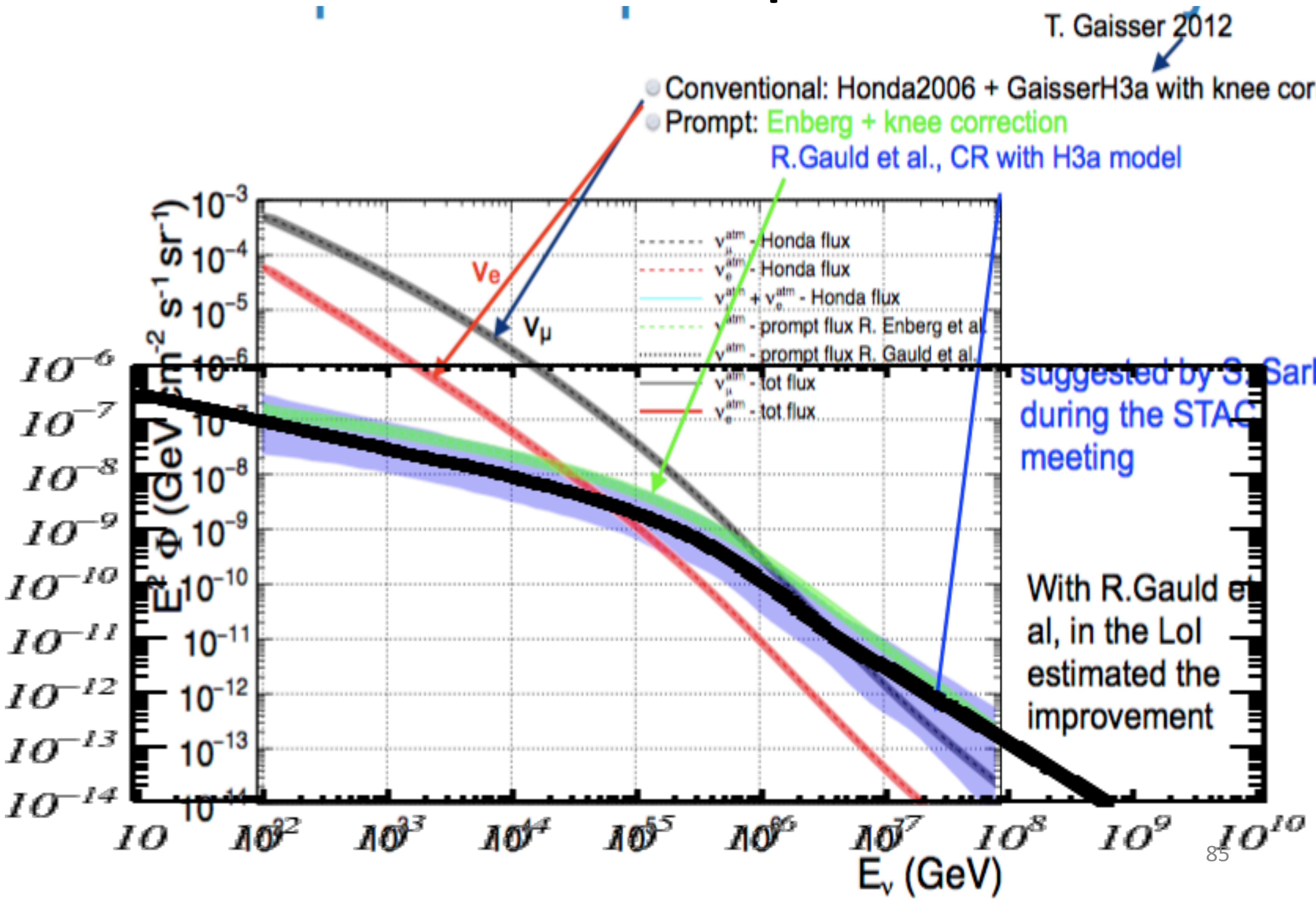
Prompt: Gauld Flux (2016)



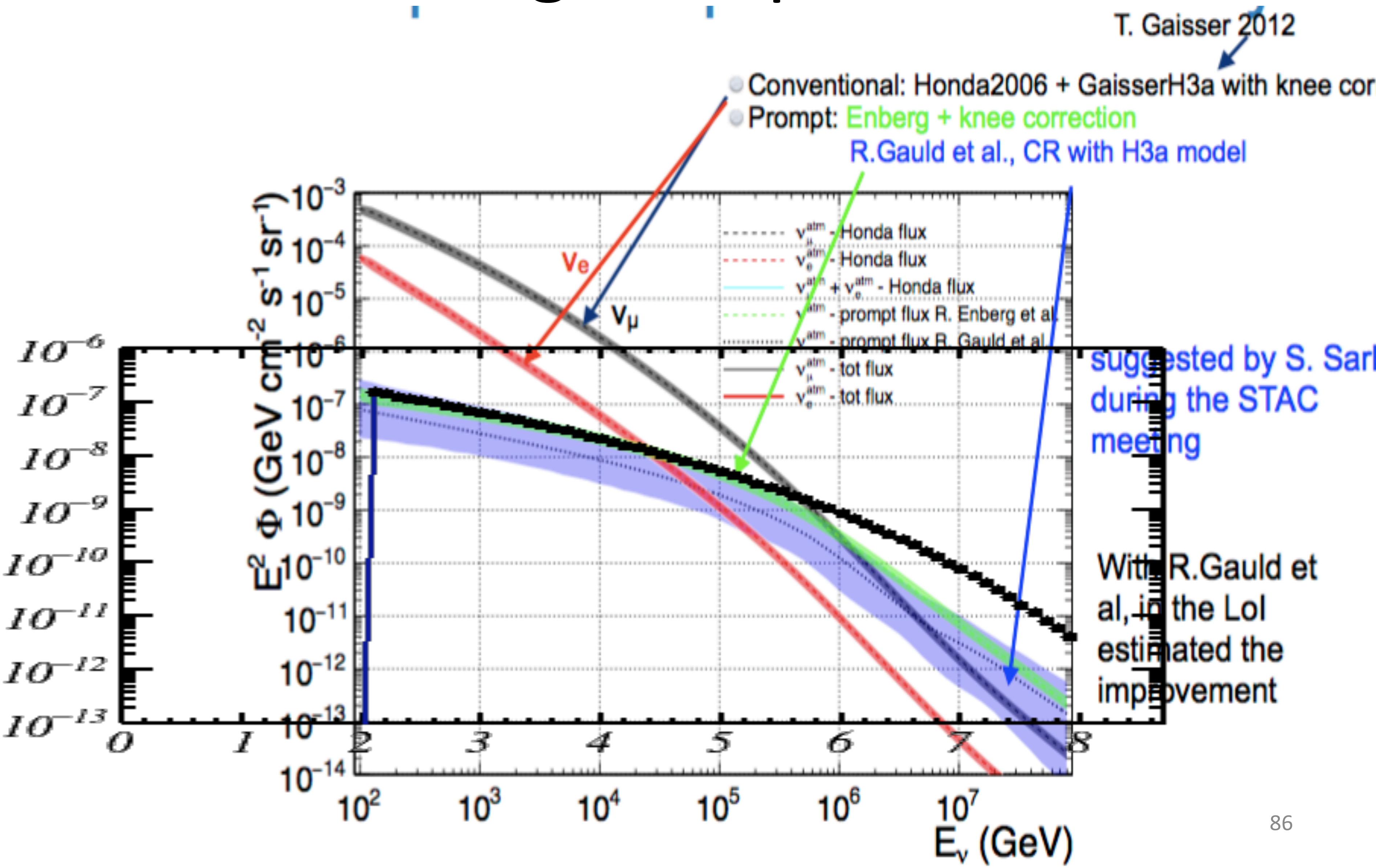
Gauld 2016


T. Gaisser 2012
Conventional: Honda2006 + GaisserH3a with knee cor
Prompt: Enberg + knee correction
R.Gauld et al., CR with H3a model
ν_e
ν_μ
ν_μ^{atm} - Honda flux
ν_e^{atm} - Honda flux
$\nu_\mu^{atm} + \nu_e^{atm}$ - Honda flux
ν^{atm} - prompt flux R. Enberg et al.
ν^{atm} - prompt flux R. Gauld et al.
ν_μ^{atm} - tot flux
ν_e^{atm} - tot flux
$E^2 \Phi$ (GeV cm^{-2} s^{-1} sr^{-1})
E_ν (GeV)
suggested by S. Sarl
during the STAC
meeting
With R.Gauld et
al, in the LoI
estimated the
improvement

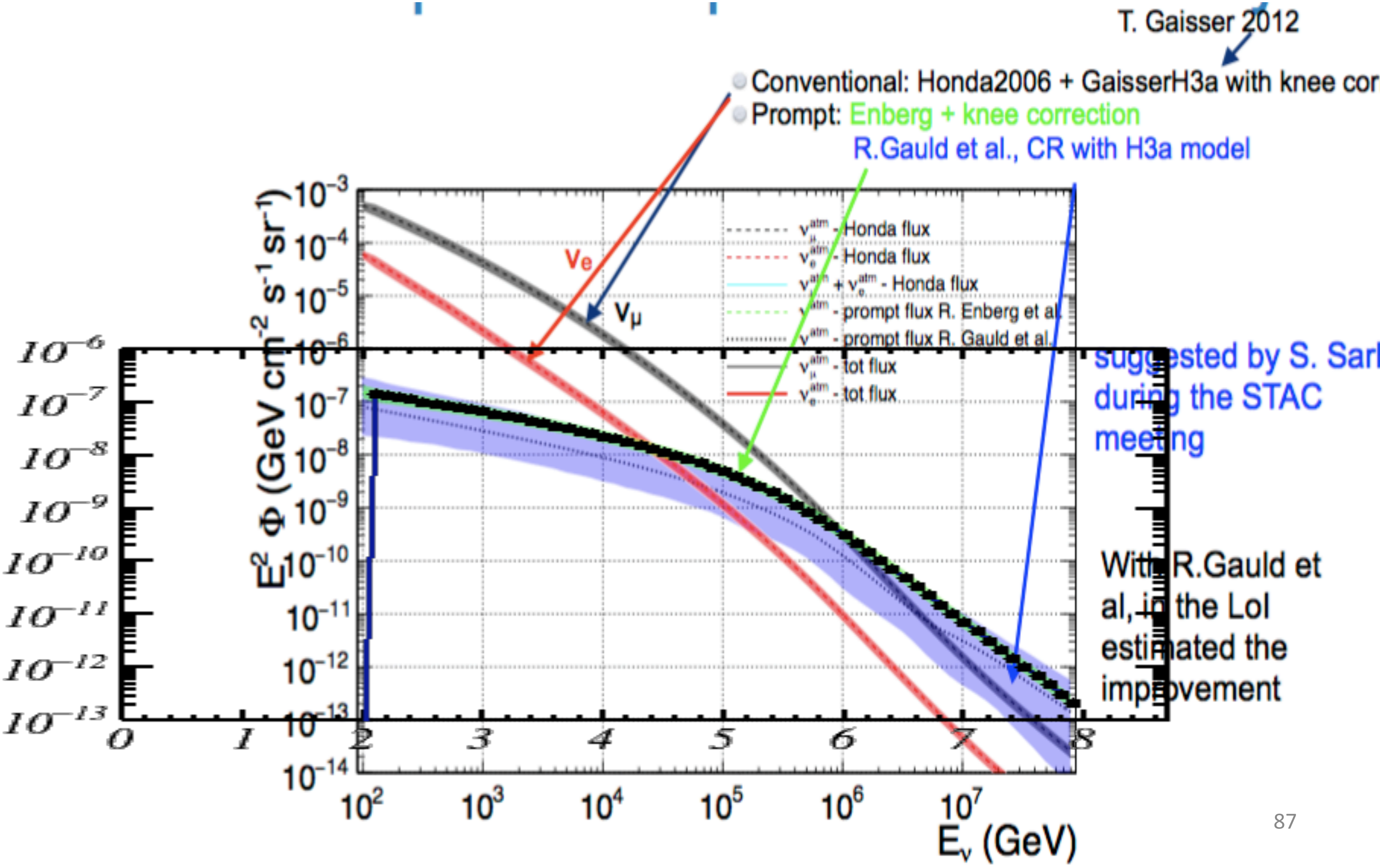
Gauld 2016 extrapolated



Enberg extrapolated

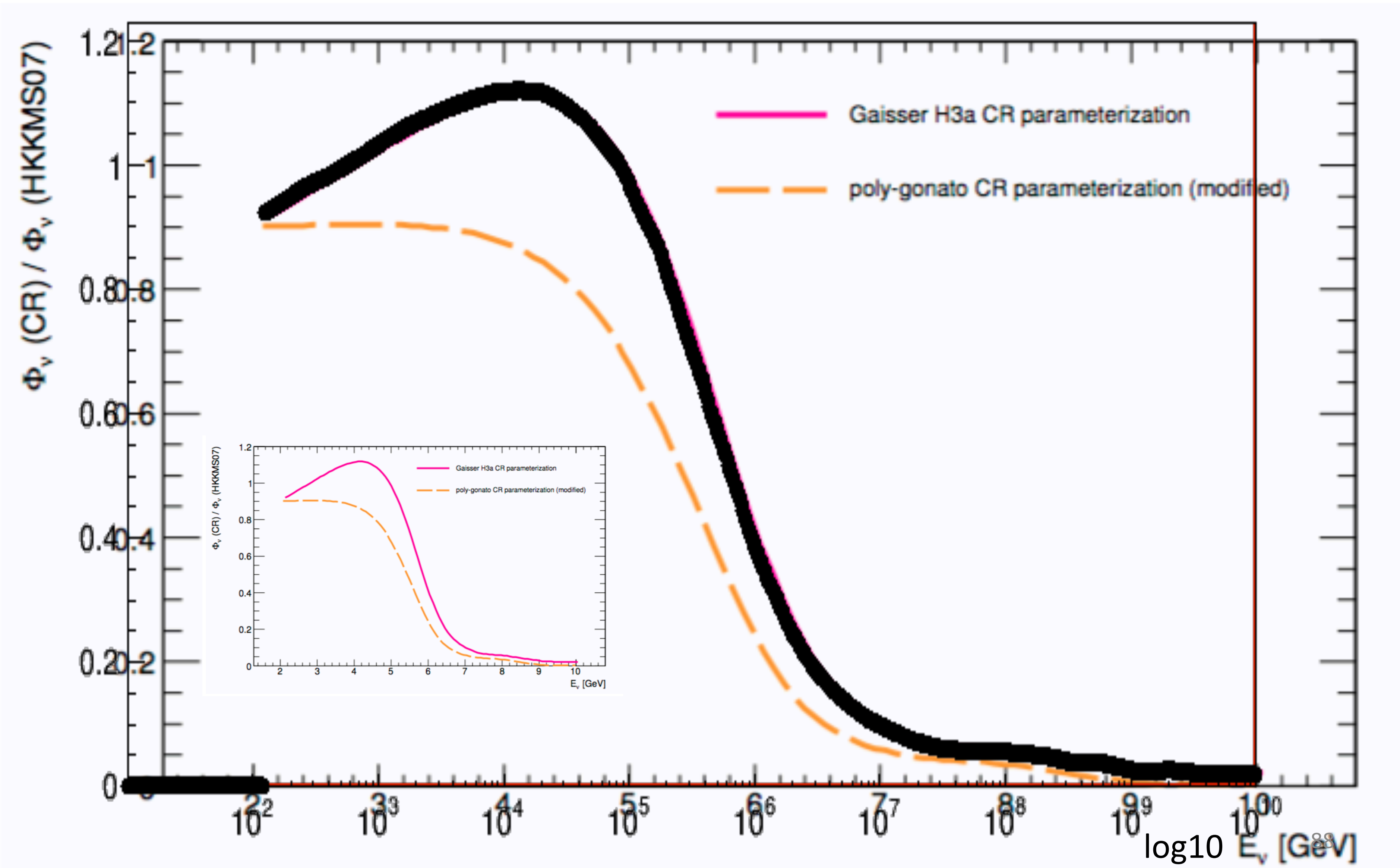


Enberg extrapolated + knee correction



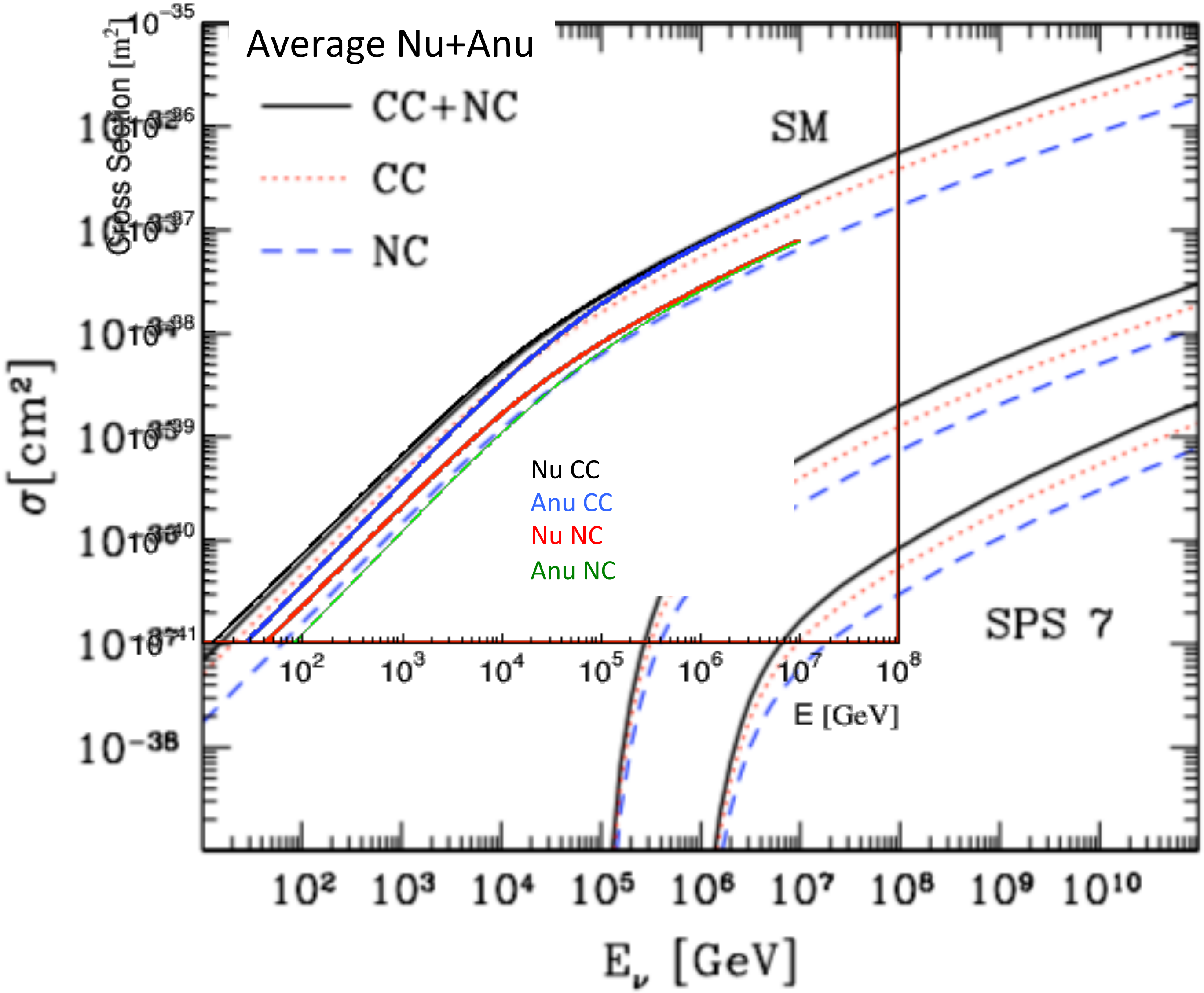
arXiv:1311.7048

Knee Correction (Gaisser H3a)



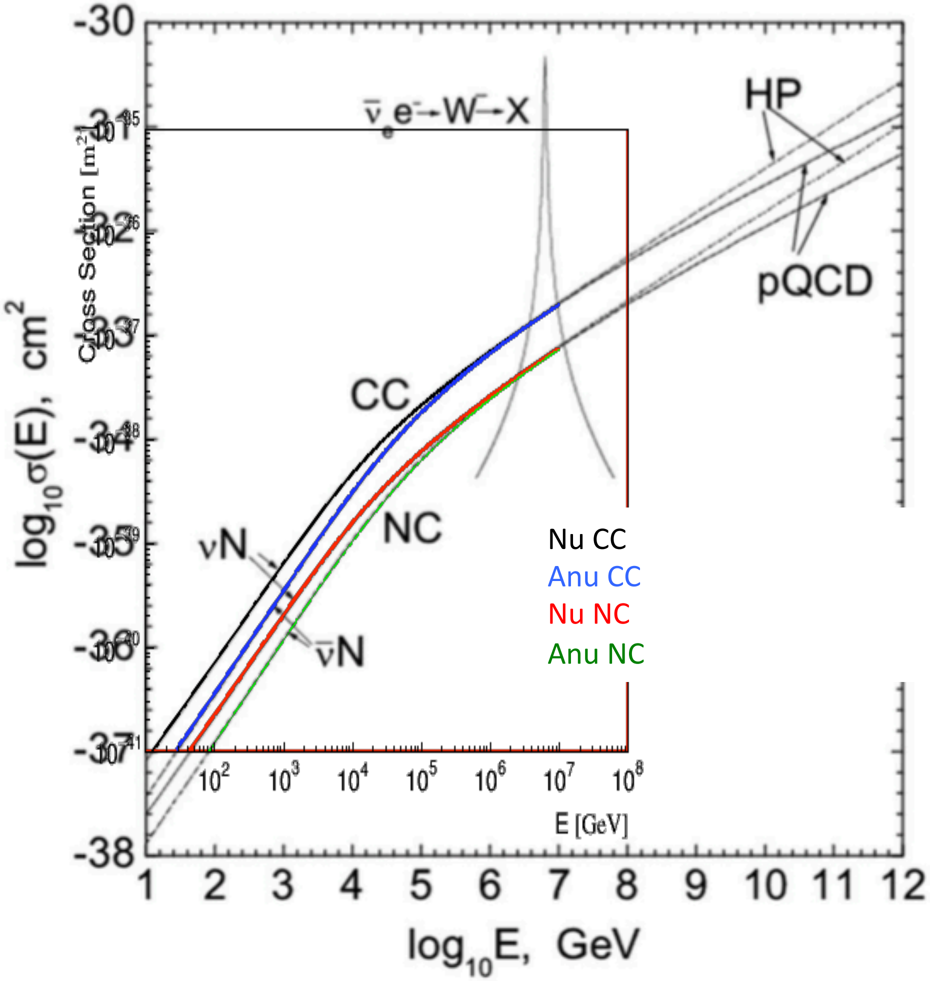
https://arxiv.org/pdf/hep-ph/0604188.pdf

Neutrino Cross Sections

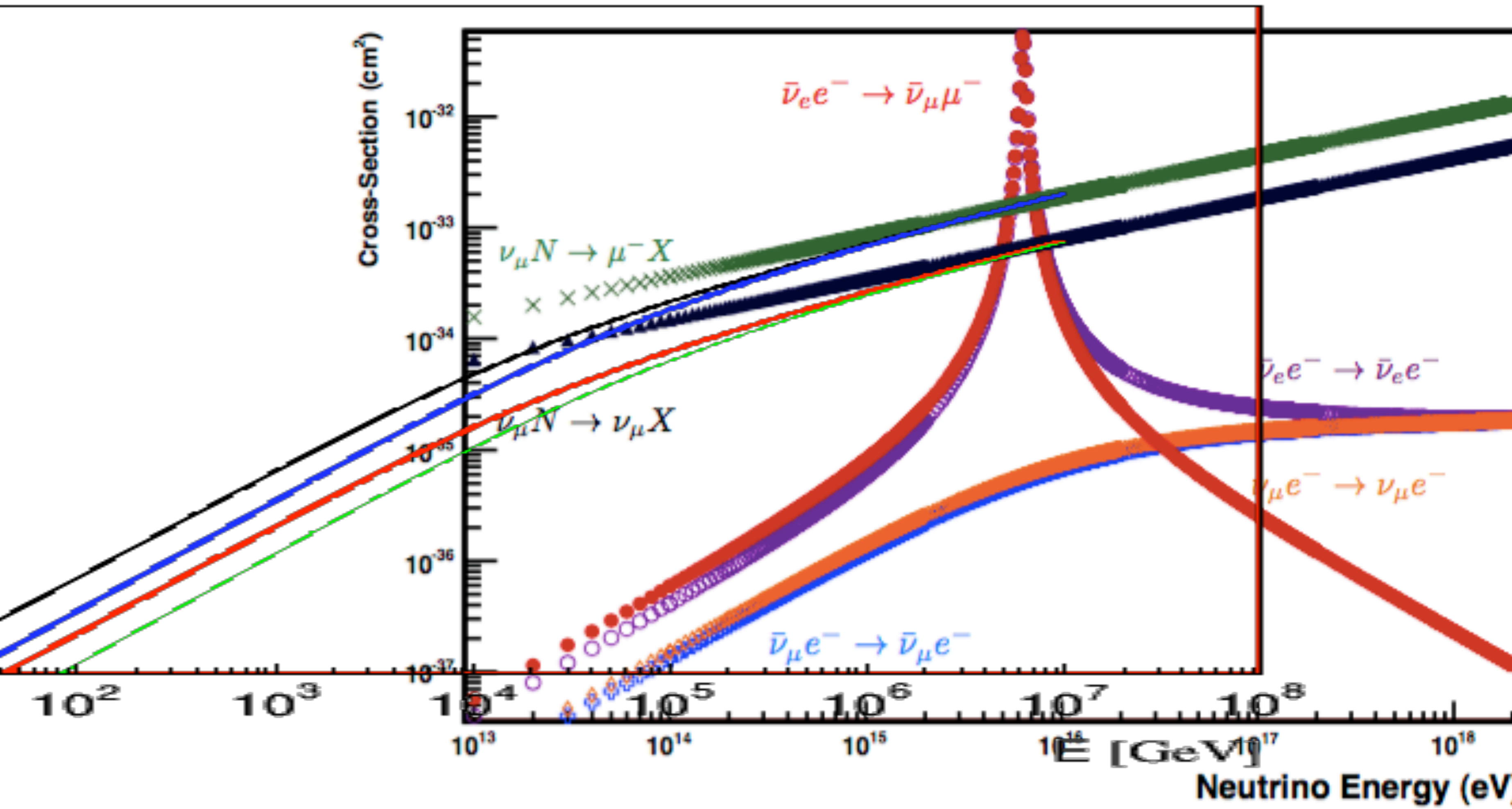


ANIS, Kowalski 2003
Figure from Colnard 2009

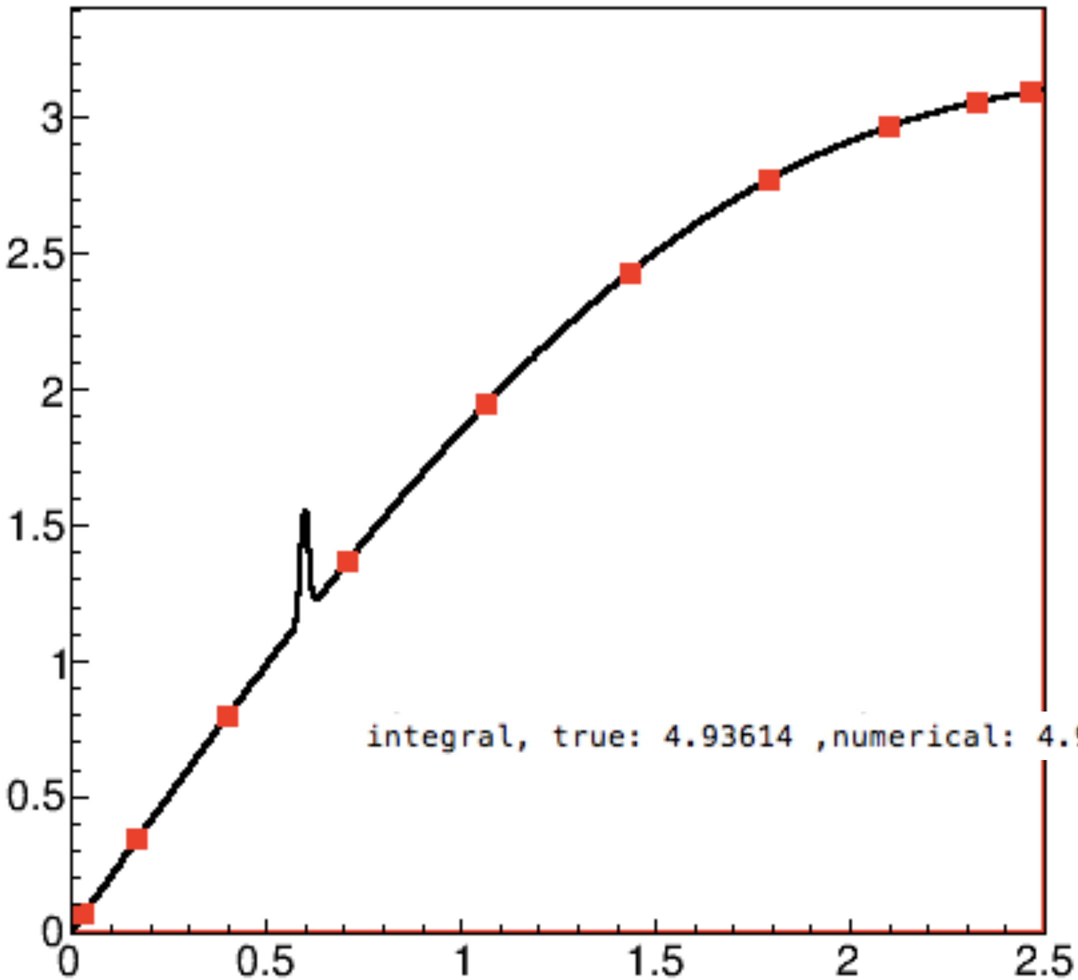
Neutrino Cross Sections



Neutrino Cross Sections

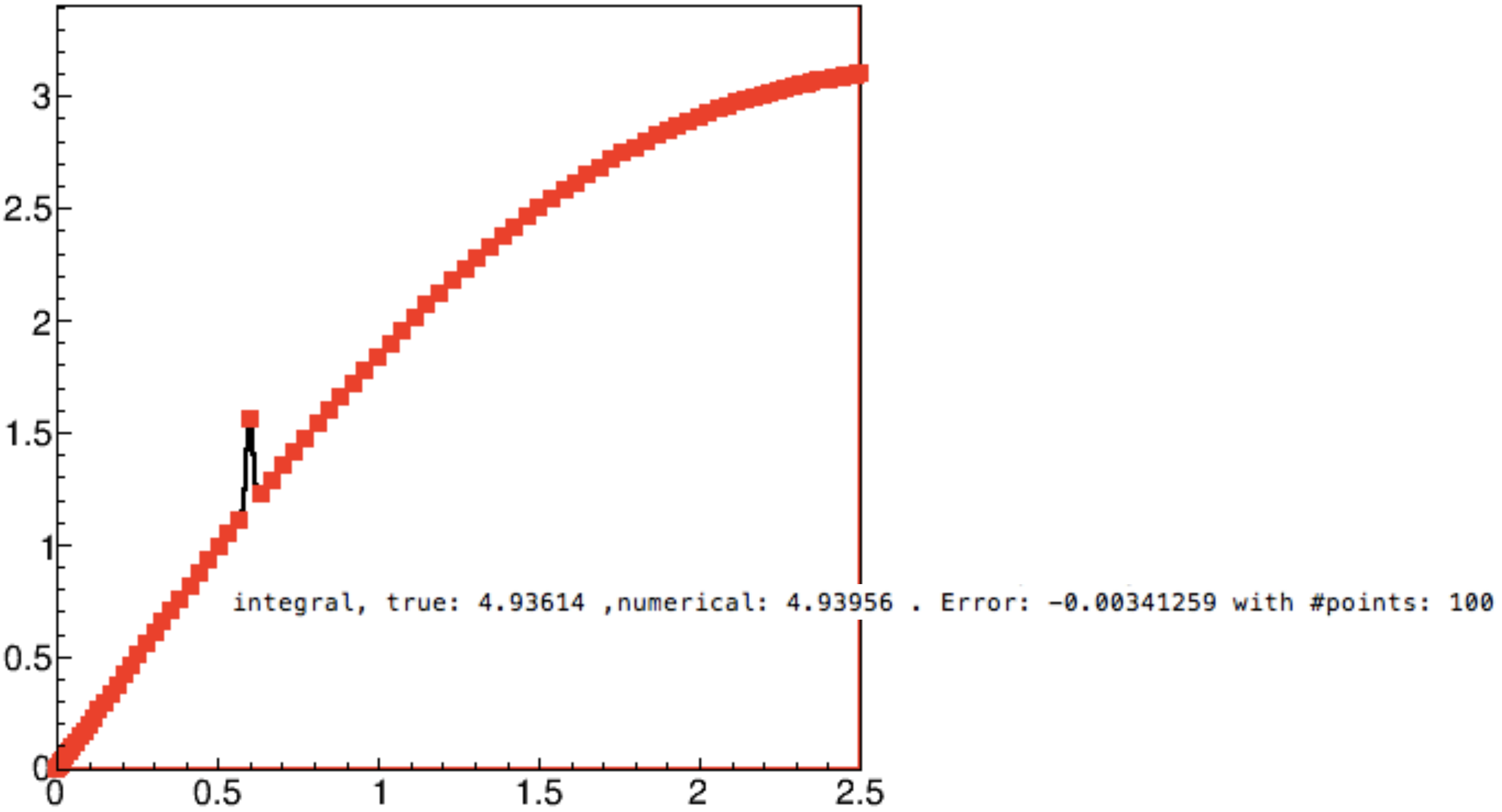


Gaussian Quadrature

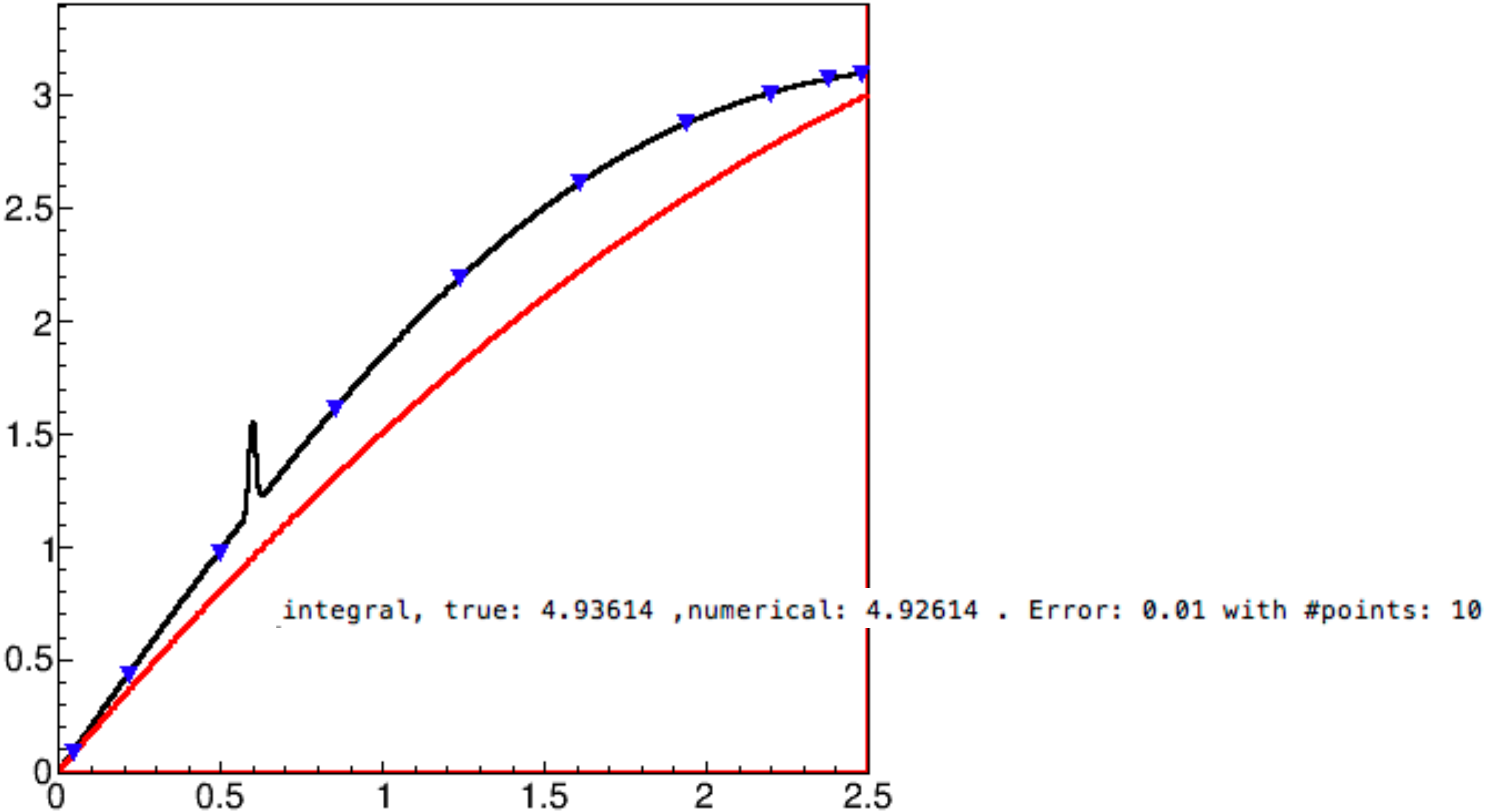


integral, true: 4.93614 ,numerical: 4.92614 . Error: 0.01 with #points: 10

Gaussian Quadrature



Gaussian Quadrature



Event Probability: Direction