Machine learning neutrino-nucleus cross sections

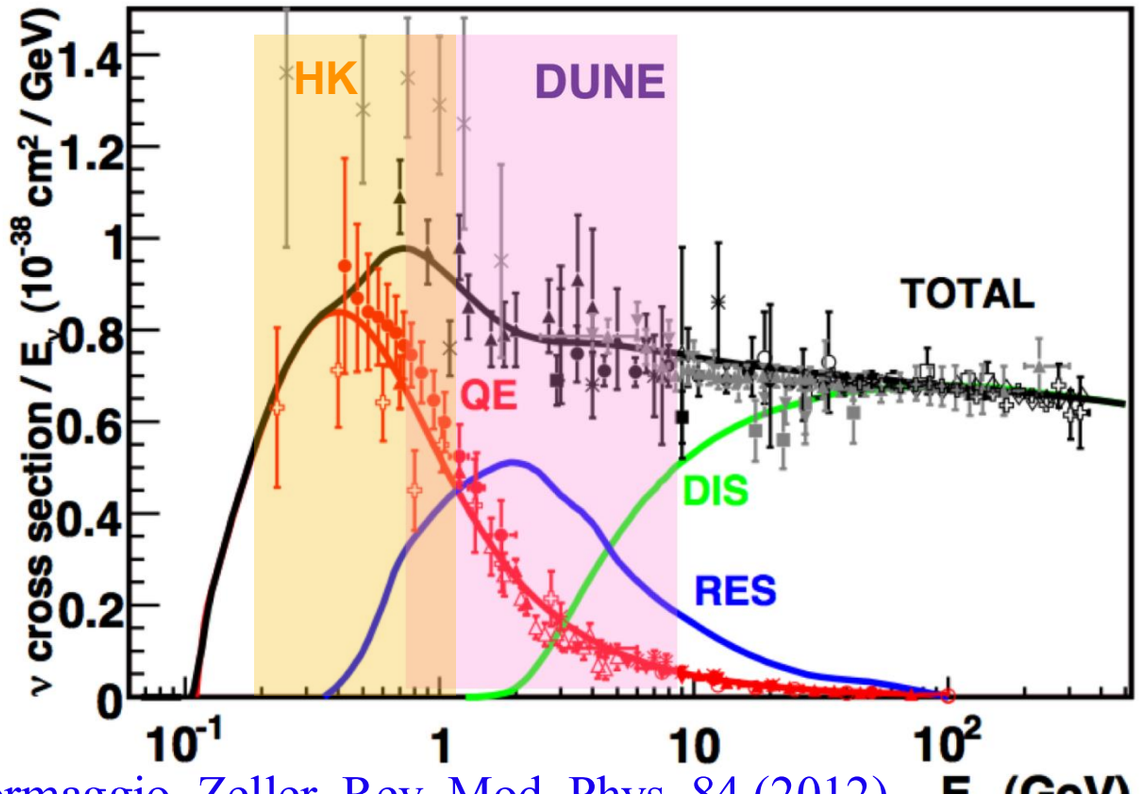
Michael Wagman

KEK Workshop: Neutrino Interactions for Current and Future Experiments

October 28, 2025

Hackett, Isaacson, Li, Tame Narvaez, MW arXiv:2412.16303

Neutrino-nucleus scattering



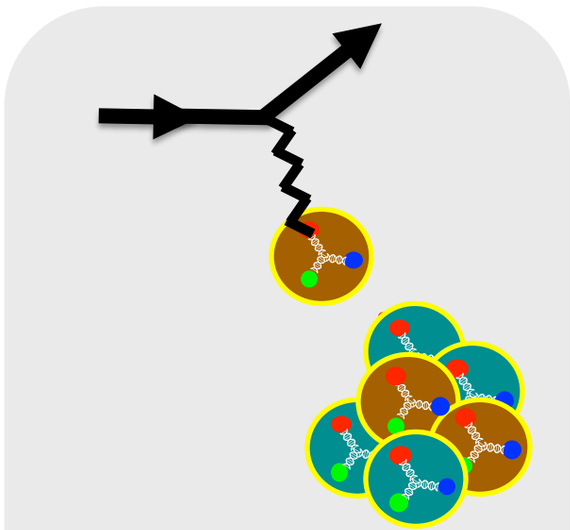
Formaggio, Zeller, Rev. Mod. Phys. 84 (2012)

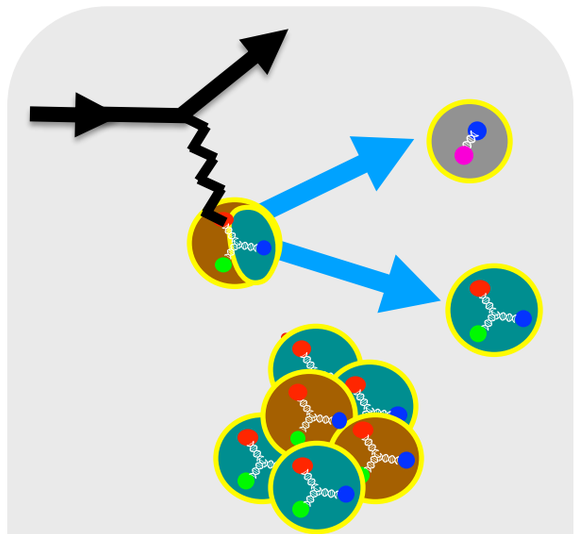
Accelerator neutrino fluxes cover a wide range of energies where different processes dominate cross-section:

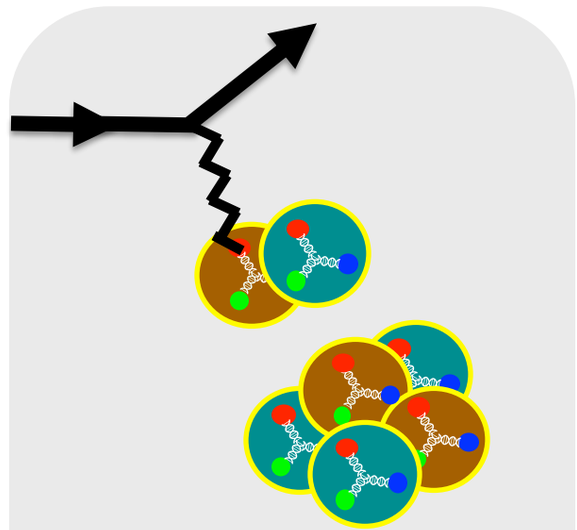
- Quasi-elastic nucleon scattering
- Resonance production
- Deep inelastic scattering

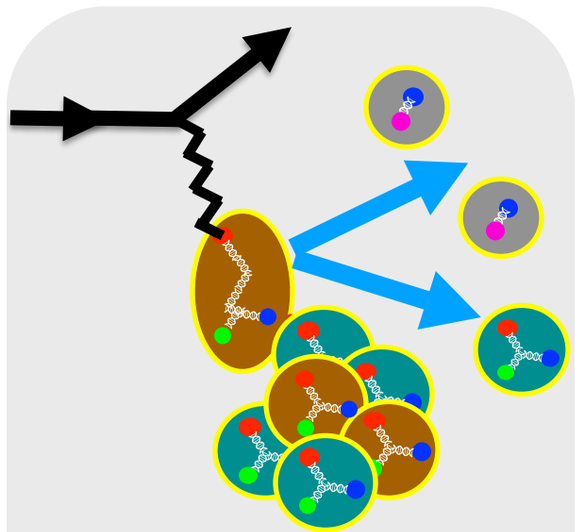
Theory input required to decompose cross section into such processes and therefore predict its energy dependence

Effective theories for different energies require different inputs



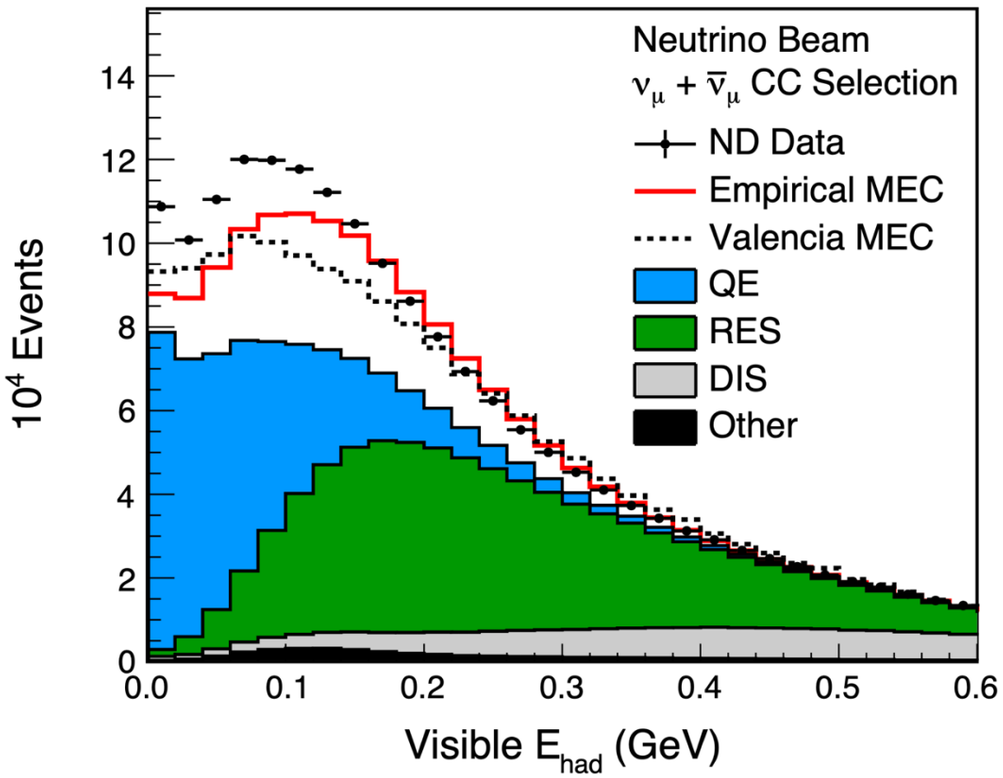






Tuning is hard

Acero et al [NOvA] Eur. Phys. J. C 80 (2020)

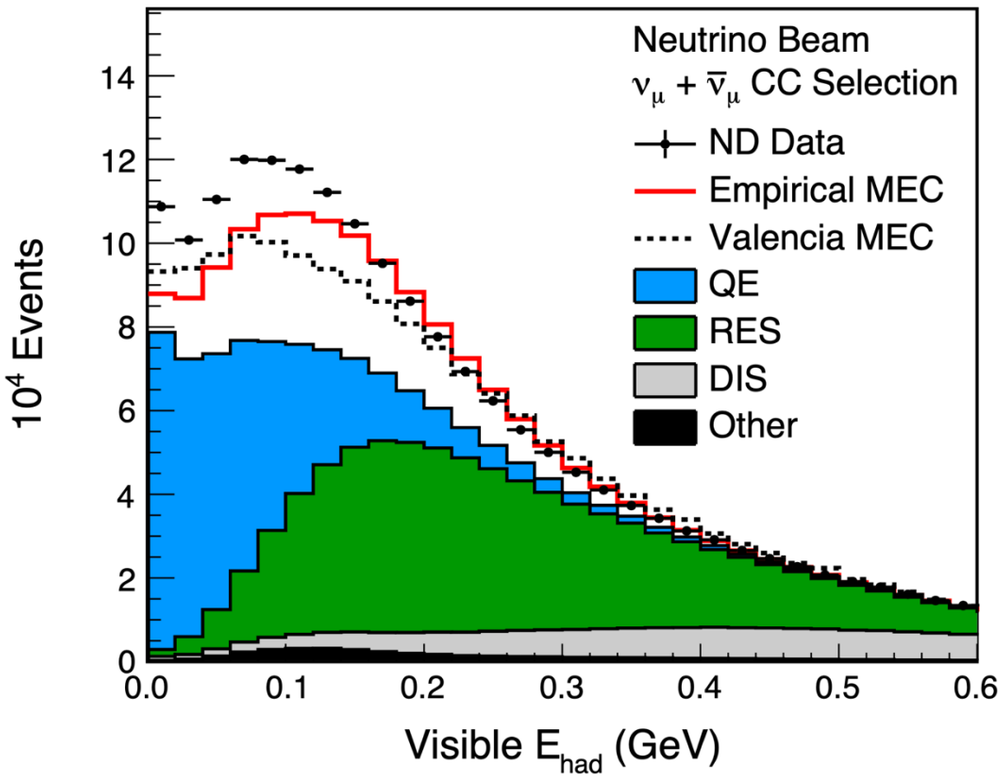


QCD is hard —- event generators must be tuned to agree with experimental data

- Interplay of different reaction mechanisms means there's no unique tuning prescription

Tuning is hard

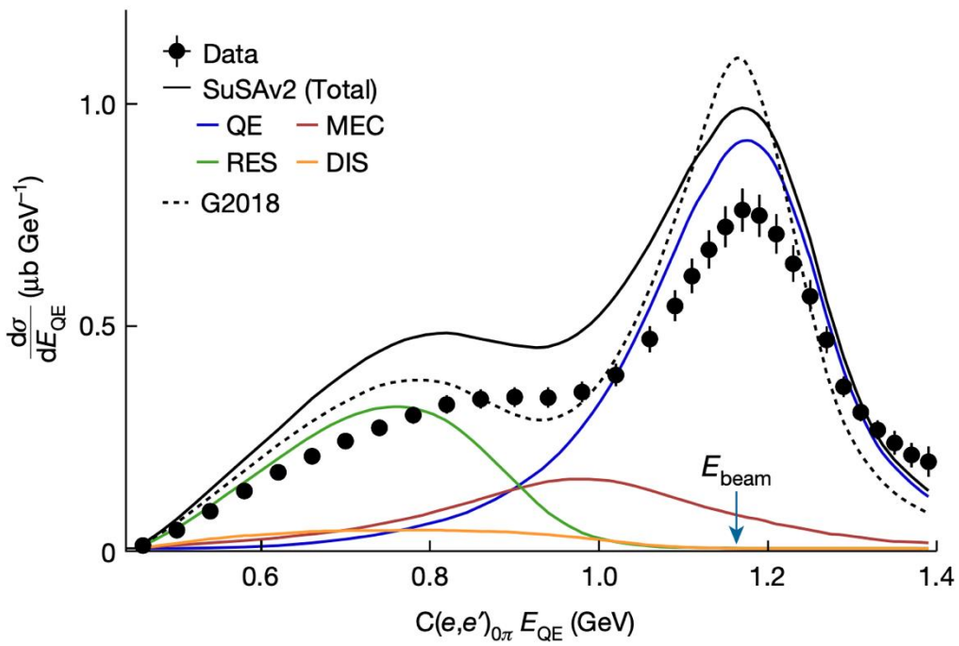
Acero et al [NOvA] Eur. Phys. J. C 80 (2020)



QCD is hard —- event generators must be tuned to agree with experimental data

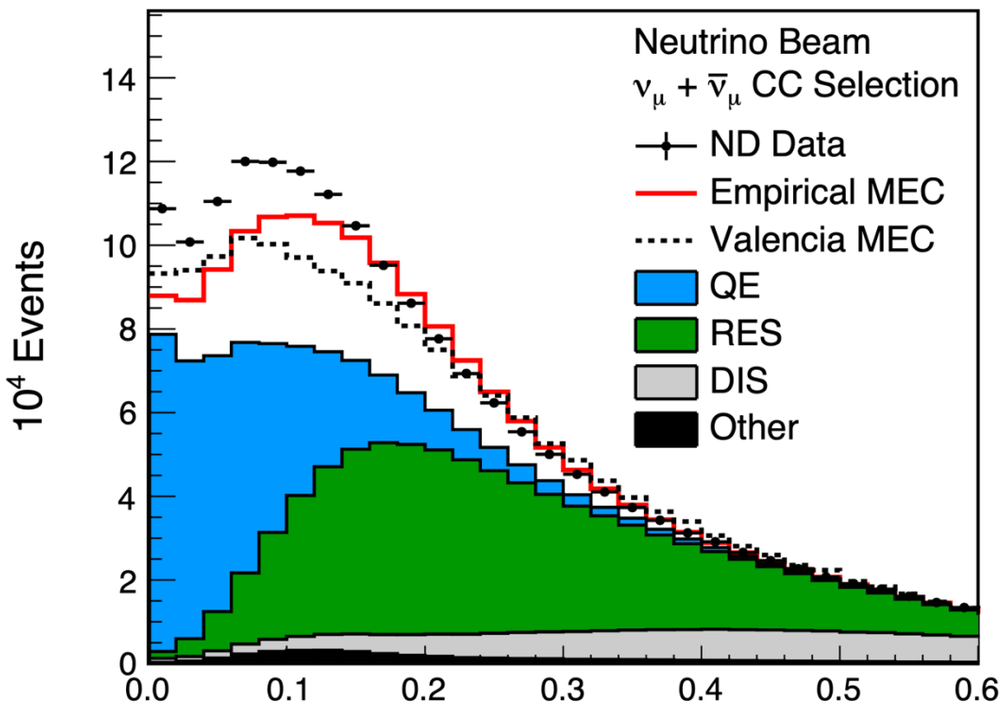
- Interplay of different reaction mechanisms means there's no unique tuning prescription
- Models tuned to agree with one dataset do not necessarily agree with others

Khachatryan et al [Clas and e4v] Nature 599 (2021)



Tuning is hard

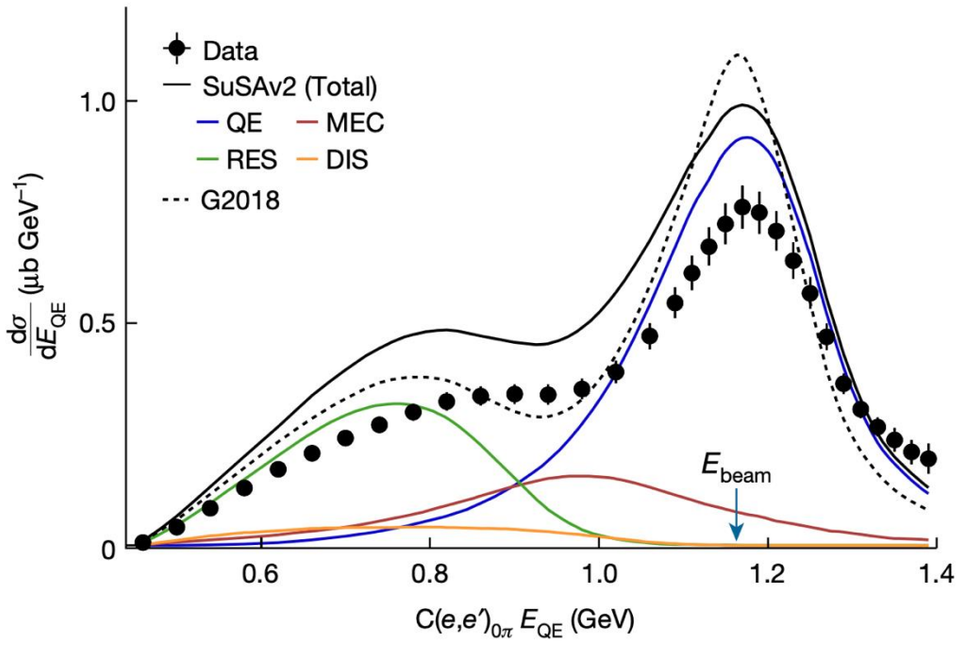
Acero et al [NOvA] Eur. Phys. J. C 80 (2020)

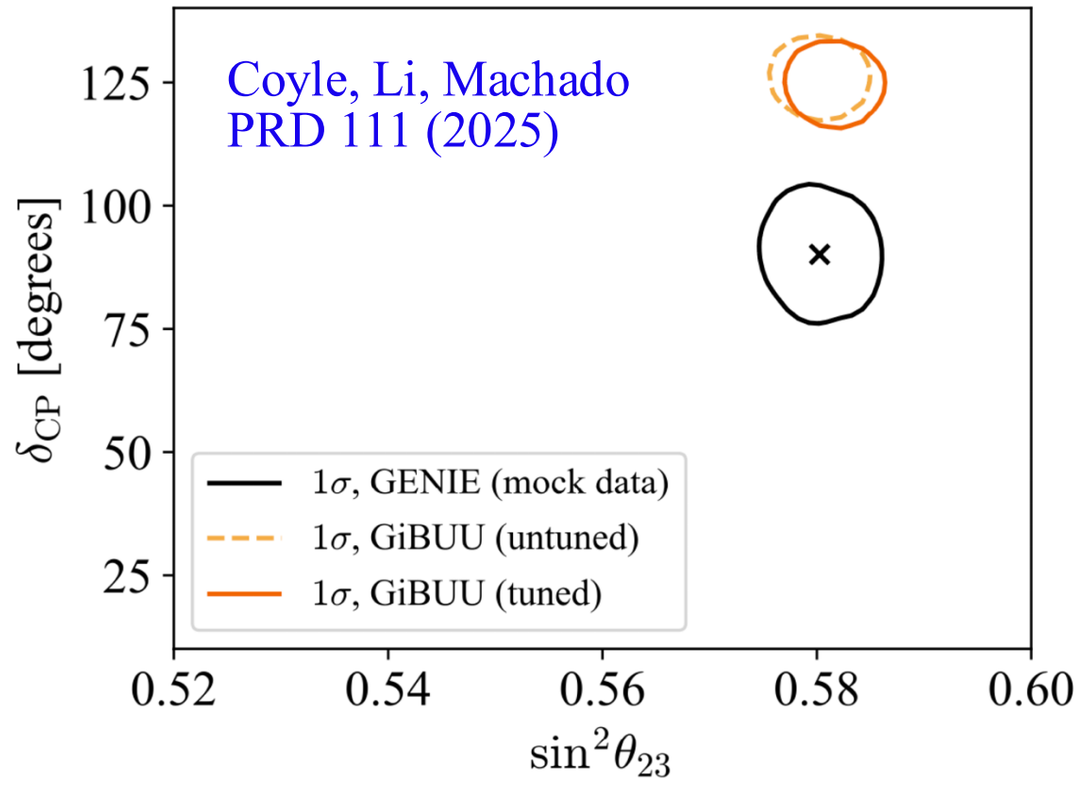


QCD is hard —- event generators must be tuned to agree with experimental data

- Interplay of different reaction mechanisms means there's no unique tuning prescription
- Models tuned to agree with one dataset do not necessarily agree with others
- Standard generator tuning does not always lead to unbiased oscillation analyses

Khachatryan et al [Clas and e4v] Nature 599 (2021)





ML for νA

Previous works have demonstrated that ML approaches can accurately model high-energy electron and neutrino scattering (e.g. NNPDF)

arXiv:hep-ph/0204232 (hep-ph)

[Submitted on 19 Apr 2002 (v1), last revised 31 Jul 2002 (this version, v3)]

Neural Network Parametrization of Deep-Inelastic Structure Functions

Stefano Forte, Lluis Garrido, Jose I. Latorre, Andrea Piccione

We construct a parametrization of deep-inelastic structure functions which retains information on experimental errors and correlations, and which does not introduce any theoretical bias while interpolating between existing data points. We generate a Monte Carlo sample of

arXiv:2406.06292 (nucl-th)

[Submitted on 10 Jun 2024]

Modeling inclusive electron-nucleus scattering with Bayesian artificial neural networks

Joanna E. Sobczyk, Noemi Rocco, Alessandro Lovato

We introduce a Bayesian protocol based on artificial neural networks that is suitable for modeling inclusive electron-nucleus scattering on a variety of nuclear targets with quantified uncertainties. Unlike previous applications in the field, which directly parameterize the cross sections, our approach employs artificial neural networks to represent the longitudinal and transverse response functions. In contrast to cross sections, which depend on the incoming energy, scattering angle, and energy transfer, the response functions are determined solely by the energy and momentum transfer to the system, allowing the angular component to be treated analytically. We

arXiv:2302.08527 (hep-ph)

[Submitted on 16 Feb 2023 (v1), last revised 5 Jun 2023 (this version, v2)]

Neutrino Structure Functions from GeV to EeV Energies

Alessandro Candido, Alfonso Garcia, Giacomo Magni, Tanjona Rabemananjara, Juan Rojo, Roy Stegeman

Our question: can a purely data-driven cross section model learned from near-detector enable an accurate neutrino oscillation analysis?

arXiv > hep-ph > arXiv:2412.16303

Search... Help | A

High Energy Physics - Phenomenology

[Submitted on 20 Dec 2024]

Machine Learning Neutrino-Nucleus Cross Sections

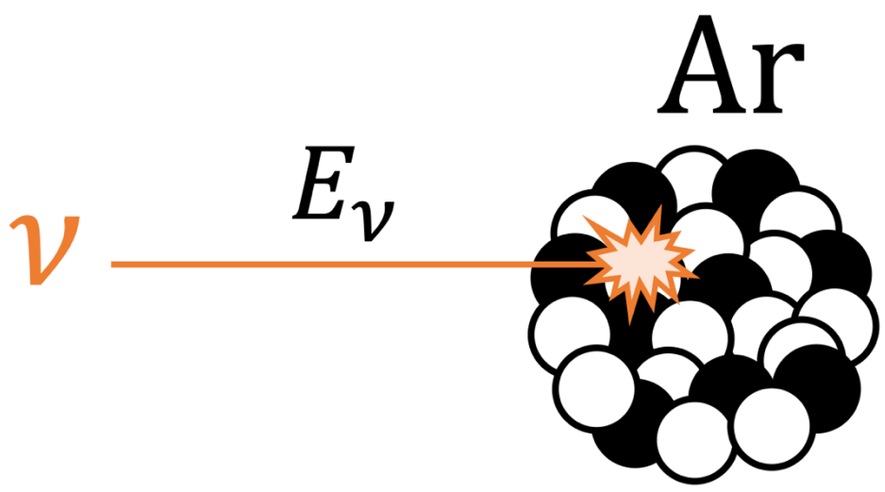
Daniel C. Hackett, Joshua Isaacson, Shirley Weishi Li, Karla Tame-Narvaez, Michael L. Wagman

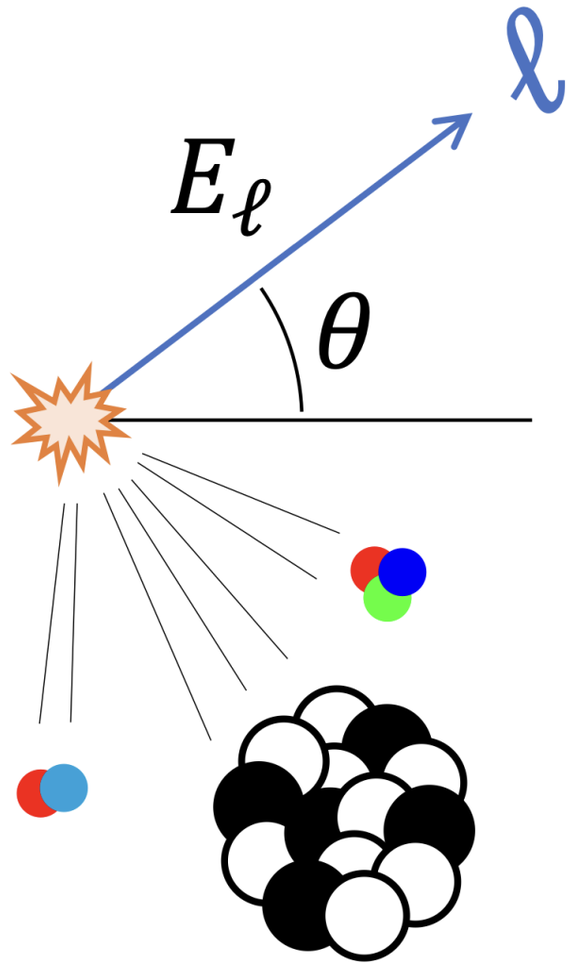
Neutrino-nucleus scattering cross sections are critical theoretical inputs for long-baseline neutrino oscillation experiments. However, robust modeling of these cross sections remains challenging. For a simple but physically motivated toy model of the DUNE experiment, we demonstrate that an accurate neural-network model of the cross section -- leveraging Standard Model symmetries -- can be learned from near-detector data. We then perform a neutrino oscillation analysis with simulated far-detector

νA scattering

DUNE will have precise LArTPC data for a variety of exclusive channels

- Incorporating ν_μ, $\overline{\nu}_\mu$, ν_e, and $\overline{\nu}_e$ data critical for mass hierarchy, CP violation

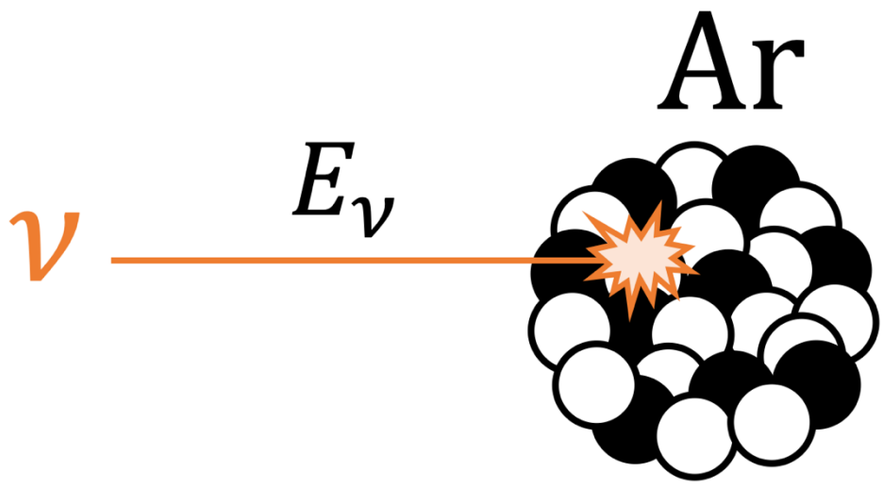




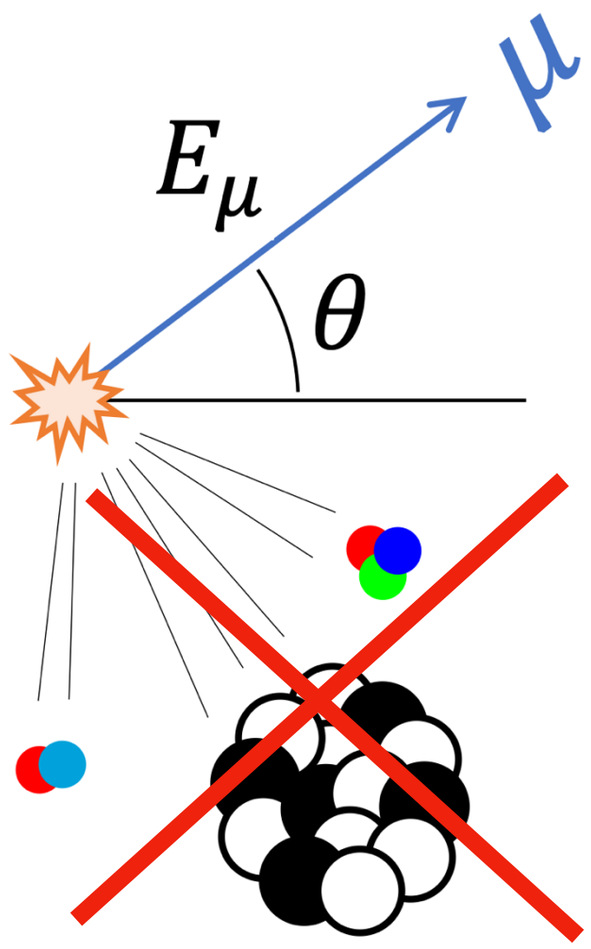
νA scattering

Theoretically simple starting point: inclusive ν_μ disappearance

- Adding other lepton channels in progress, exclusive hadrons more challenging
- Fully inclusive analysis may already be useful for Super K, Hyper-Kamiokande, …



$$\nu_\mu + A \rightarrow \mu^+ + X$$



Closure test

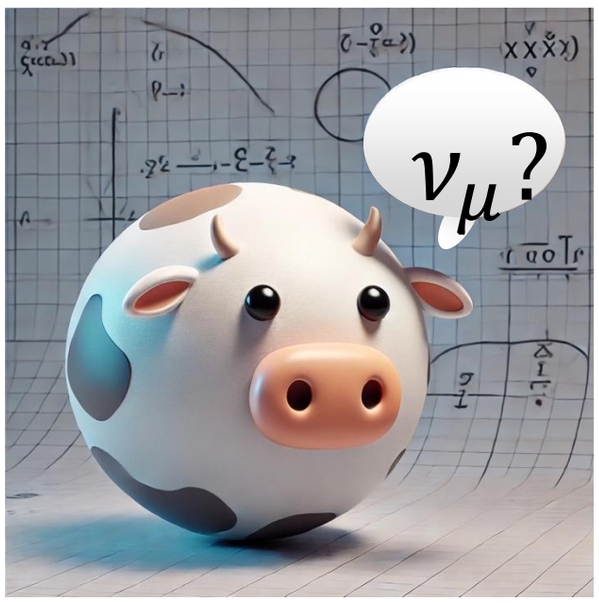
Can a learned xsec model reproduce “right answers” in a test where truth is known?

Toy model

Simplified model acting as “ground truth” for testing accuracy of learned model

- Explicit analytic xsec function, neutrino fluxes, oscillation parameters

- We won’t have access to this much information in reality



Learned model

Neural model trained to fit data

- Flexible parameterization of xsec trained to fit ND data
- Could be trained on real data

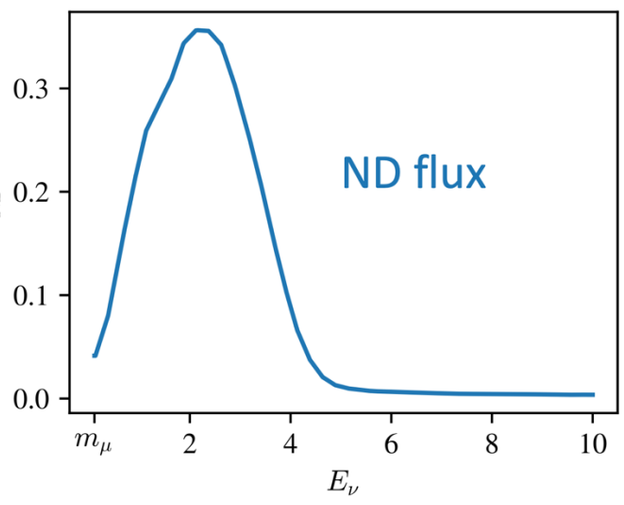
- Closure test: can learned model accurately approximate toy model?

Near detector data

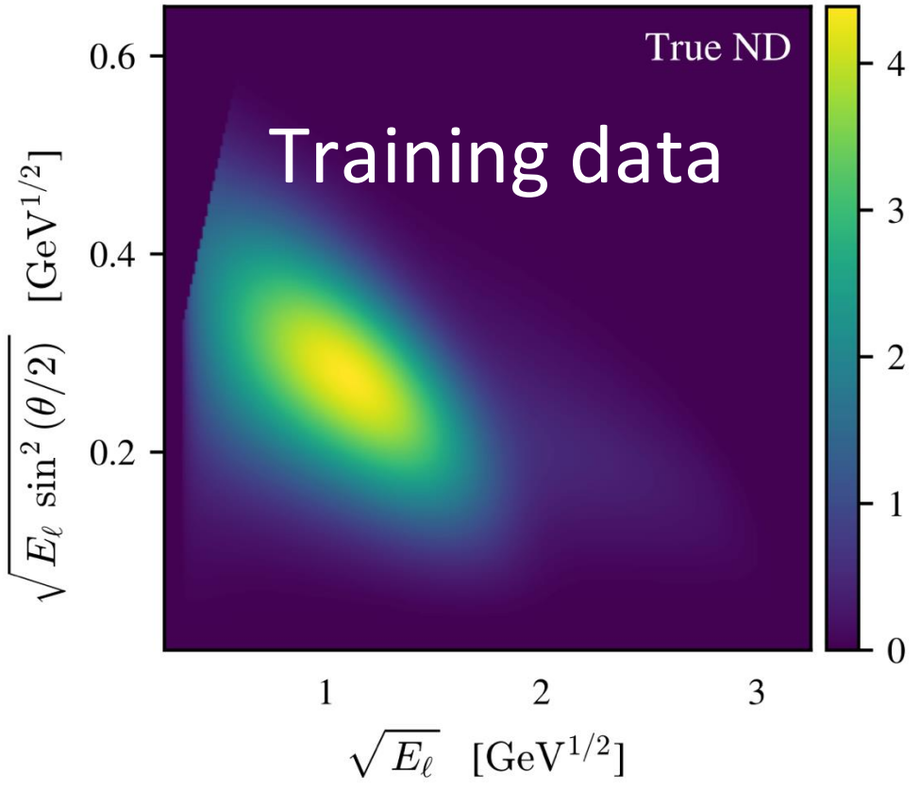
Event distribution = $\int dE_\nu$ Cross section x Neutrino flux

$$P_{ND} \propto \int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta}(E_\nu, E_l, \cos\theta)\Phi_{ND}(E_\nu)$$

$$\int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta} \times$$



=



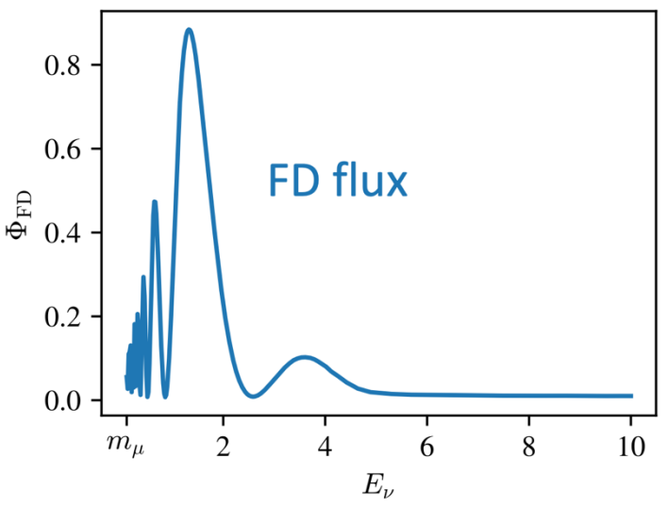
Suppose infinite ND stats for simplicity

Far detector data

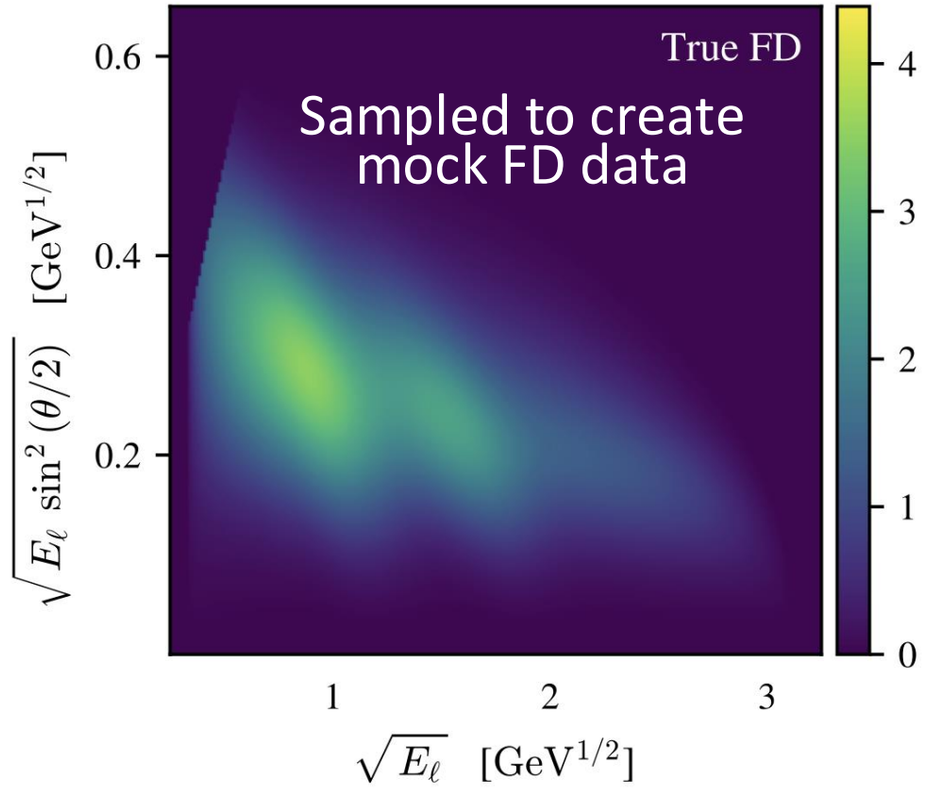
Event distribution = $\int dE_\nu$ Cross section x Neutrino flux

$$P_{FD} \propto \int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta}(E_\nu, E_l, \cos\theta)\Phi_{FD}(E_\nu)$$

$$\int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta} \times$$



=



Take 6200 events from FD marginal

Approx. 3.5 years

How <u>not</u> to learn νA

Cross sections are positive definite, can be interpreted as probability distributions

→ Model xsec as normalizing flow?

How <u>not</u> to learn νA

Cross sections are positive definite, can be interpreted as probability distributions

→ Model xsec as normalizing flow?

Measured kinematic variables defined as change-of-variables from a simple (i.e. normal) base distribution

$E_l = g(\phi_1, \phi_2, \phi_3)$

$\cos\theta = h(\phi_1, \phi_2, \phi_3)$

$E_\nu = f(\phi_1, \phi_2, \phi_3)$

Parameterize with expressive differentiable function, e.g. neural net

How <u>not</u> to learn νA

Cross sections are positive definite, can be interpreted as probability distributions

→ Model xsec as normalizing flow?

Measured kinematic variables defined as change-of-variables from a simple (i.e. normal) base distribution

$$E_l = g(\phi_1, \phi_2, \phi_3)$$
$$\cos\theta = h(\phi_1, \phi_2, \phi_3)$$
$$E_\nu = f(\phi_1, \phi_2, \phi_3)$$

Parameterize with expressive differentiable function, e.g. neural net

Cross section defined through Jacobian

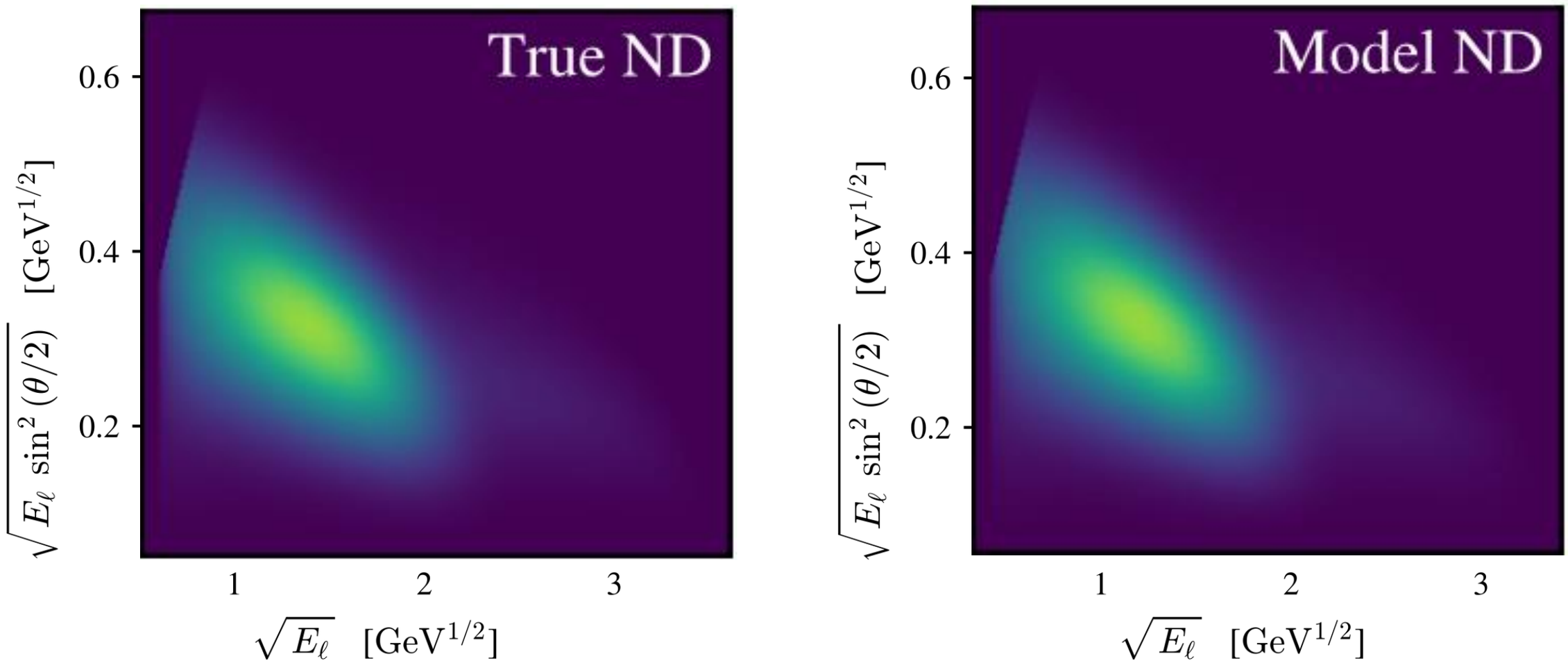
Base distribution (fixed)

$$\frac{d^2\sigma}{dE_l d\cos\theta}(E_\nu, E_l, \cos\theta) = \left| det \frac{\partial(E_\nu, E_l, \cos\theta)}{\partial(\phi_1, \phi_2, \phi_3)} \right| \tilde{B}(\phi_1, \phi_2, \phi_3)$$

Optimize change-of-variables functions to fit xsec training data

How <u>not</u> to learn νA — training

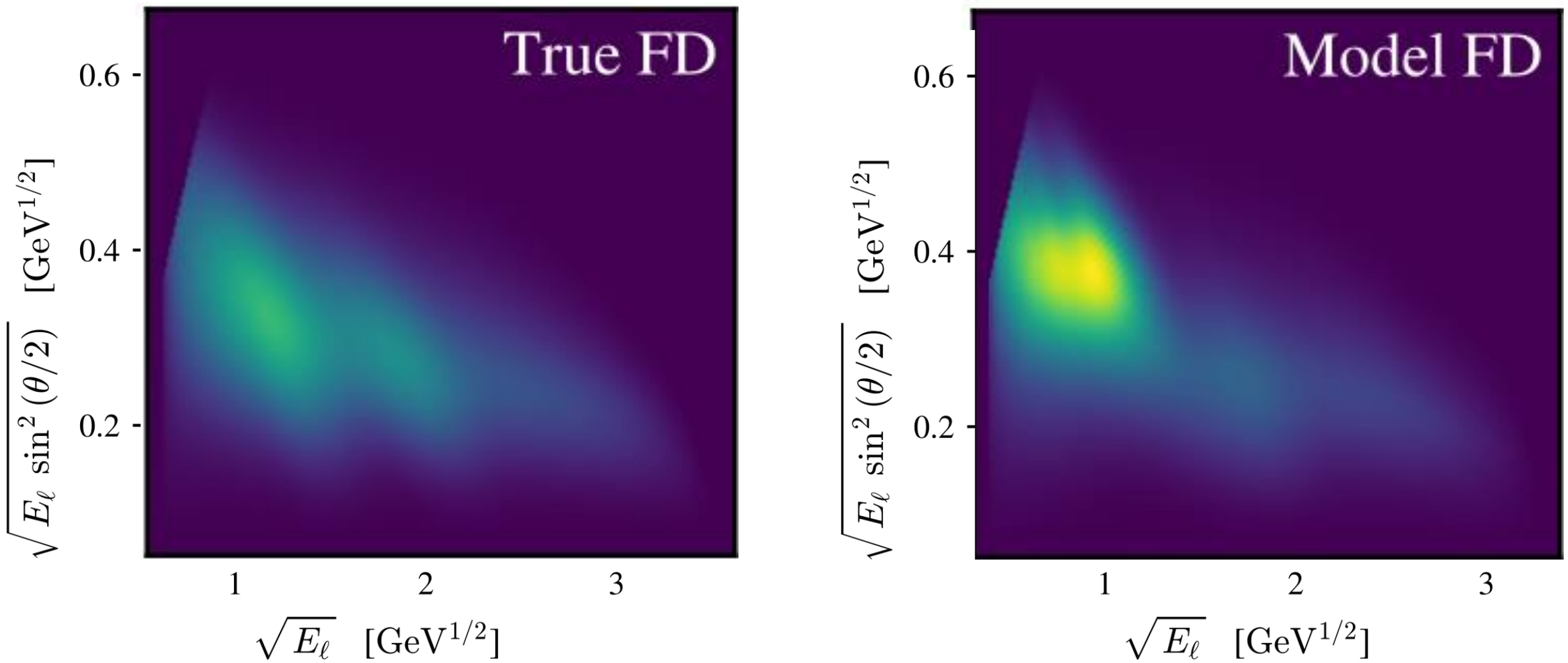
Normalizing flow xsec model can accurately reproduce ND distribution (training data)



$$P_{ND} \propto \int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta}(E_\nu, E_l, \cos\theta)\Phi_{ND}(E_\nu)$$

How <u>not</u> to learn νA — prediction

Normalizing flow xsec model **<u>cannot accurately generalize</u>** to FD distribution

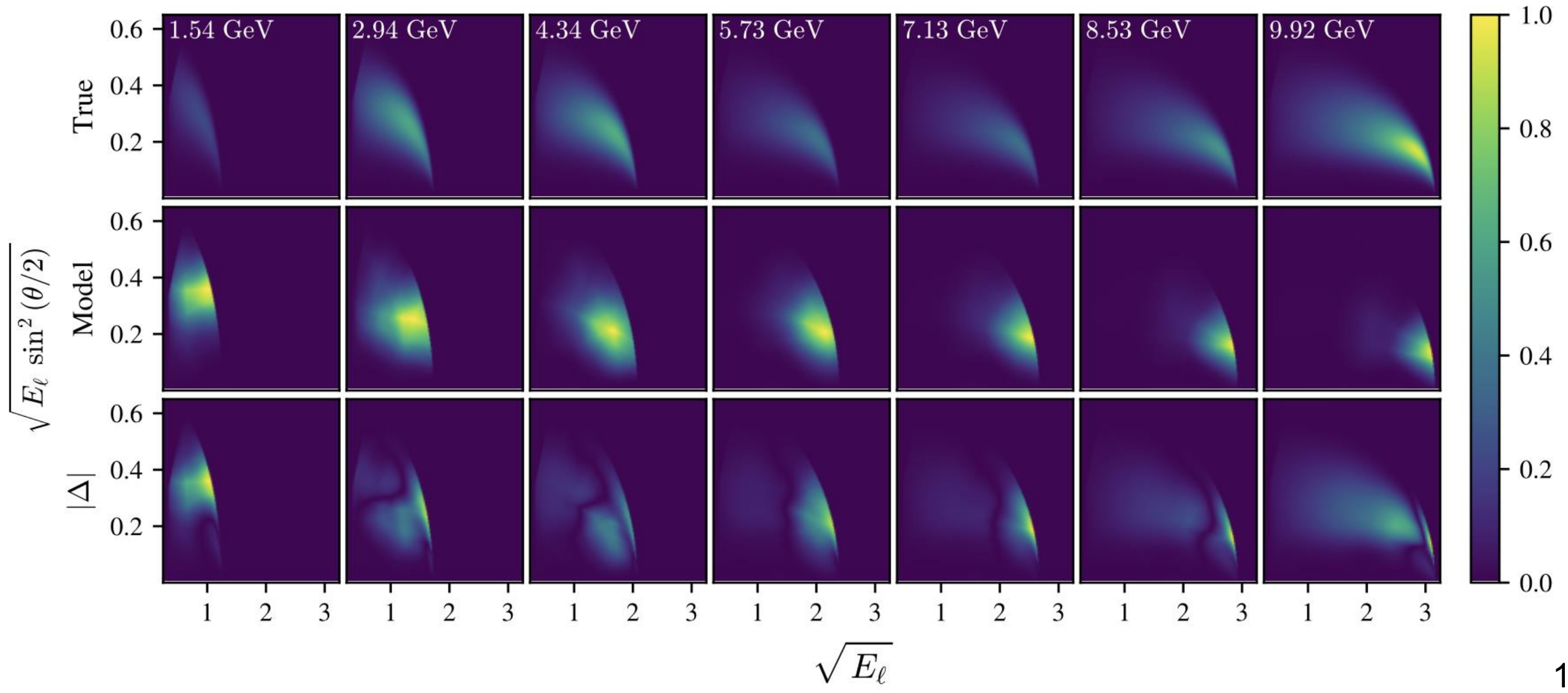


$$P_{FD} \propto \int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta}(E_\nu, E_l, \cos\theta)\Phi_{FD}(E_\nu)$$

How <u>not</u> to learn νA — diagnosis

Training data forms one 2D slice of a 3D distribution

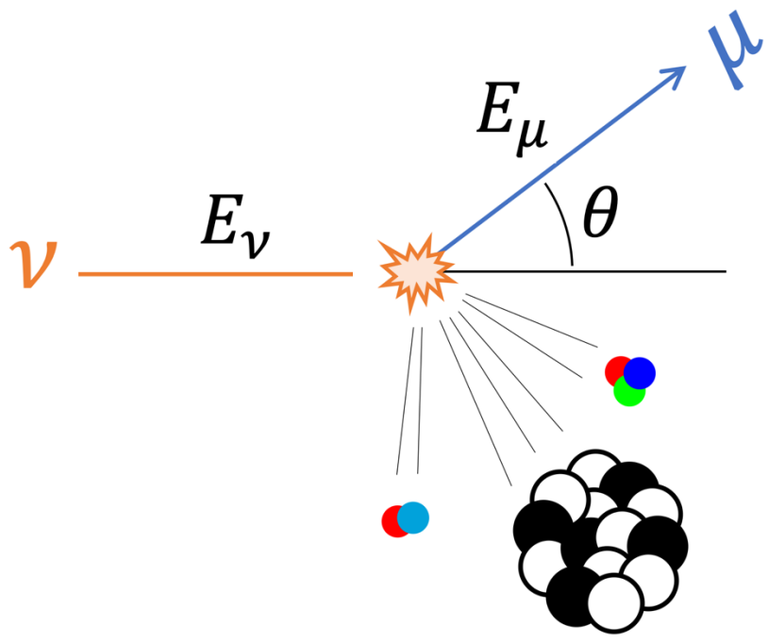
- Accurately learning 2D slice doesn't imply accurate predictions for other slices
- In particular, cross section on slices of fixed E_ν is **<u>not</u>** accurately predicted



Nuclear structure functions

Cross-section factorizes into product of lepton tensor and hadron tensor (tree-level *W*-boson exchange)

- Only assumes Standard Model symmetries, no nuclear theory



$$\frac{d^2\sigma}{dE_\ell d\cos\theta} = L^{\mu\nu} W_{\mu\nu}$$

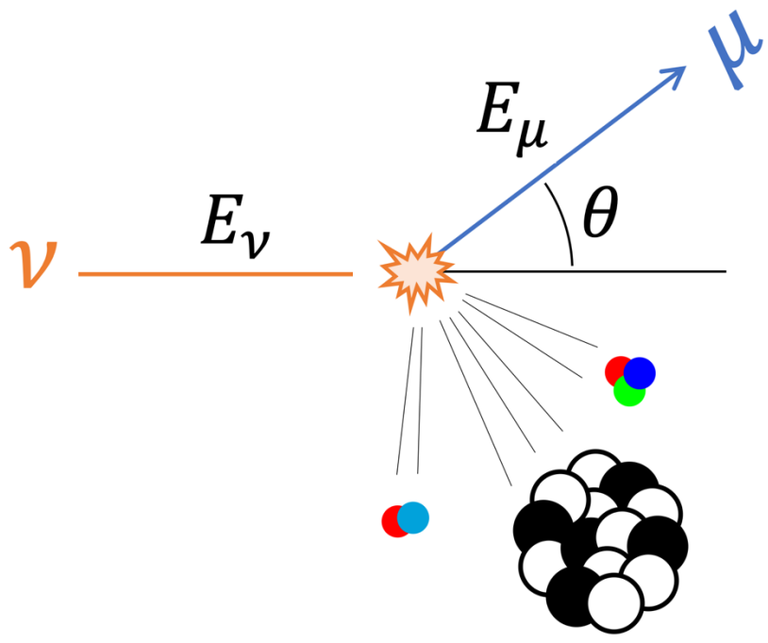
Lepton tensor
Known function of
E_ν, E_ℓ, $\cos\theta$

Hadron tensor
$W = W(x, Q^2)$
Non-perturbative
Encodes nuclear structure

Nuclear structure functions

Cross-section factorizes into product of lepton tensor and hadron tensor (tree-level *W*-boson exchange)

- Only assumes Standard Model symmetries, no nuclear theory



$$\frac{d^2\sigma}{dE_\ell d\cos\theta} = L^{\mu\nu} W_{\mu\nu}$$

Lepton tensor
Known function of
$E_\nu, E_\ell, \cos\theta$

Hadron tensor
$W = W(x, Q^2)$
Non-perturbative
Encodes nuclear structure

Parity violating, specific to ν scattering

$$W_{\mu\nu} = W_1 g_{\mu\nu} + W_2 \frac{p_\mu p_\nu}{p^2} \pm W_3 \left(\frac{i p^\rho p^\sigma}{2p \cdot q}\right) \epsilon_{\mu\nu\rho\sigma} + W_4 \frac{q_\mu q_\nu}{q^2} - W_5 \frac{p_\mu q_\nu + q_\mu p_\nu}{p \cdot q}$$

Usual DIS structure functions (but at low Q^2)

"Albright-Jarlskog functions" suppressed by m_l^2/Q^2 but non-negligible

Neural SF model

3D xsec is sum of unknown 2D functions times known 3D functions

$$\frac{\widetilde{d^2\sigma}}{dE_\ell d\cos\theta}(E_\nu) = \sum_{i=1}^{5} K_i(E_\nu, E_\ell, \cos\theta)\, \widetilde{W}_i(x, Q^2)$$

Learn from data

2D function

Known kinematic coefficients

3D function

Modeled cross section

3D function

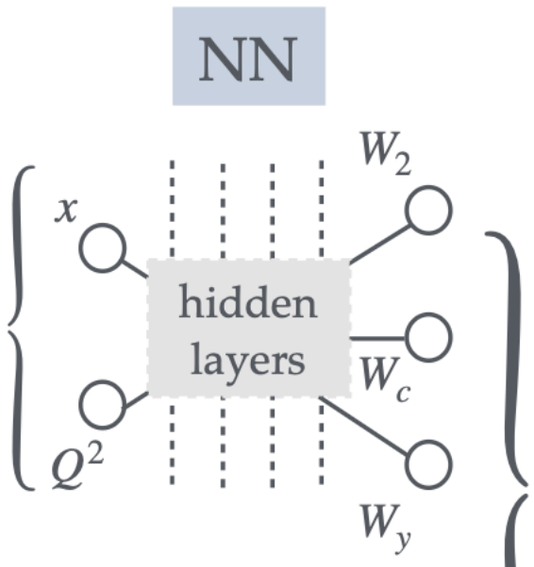
Well-posed problem to learn 2D structure functions from 2D training data

- If ND and FD distributions overlap in (x, Q^2) then changing E_ν flux doesn't require learned model to extrapolate
- Dependence on E_ν flux primarily captured by known kinematic coefficients

ML setup

When only considering one lepton channel (e.g. muon disappearance), only sensitive to 3 linear combinations of structure functions

$$\frac{d^2\sigma}{dE_\ell d\cos\theta}(E_\nu) = \frac{|V_{ud}|^2 G_F^2}{\pi}\sqrt{E_\ell^2 - m_\ell^2}\left\{ \frac{E_\nu}{M_A} W_2(x,Q^2) + W_c(x,Q^2) + \tilde{y} W_y(x,Q^2)\right\}$$



Mock ND data
$p(E_l, \cos\theta)$

from true cross section model $d\sigma(E_\nu, E_l, \cos\theta)$ with ND flux(E_ν)

Neural SF model: parameterize independent structure functions as fully connected neural nets

Training

Model cross section from trained NNs

Mock FD data
$p(E_l, \cos\theta)$

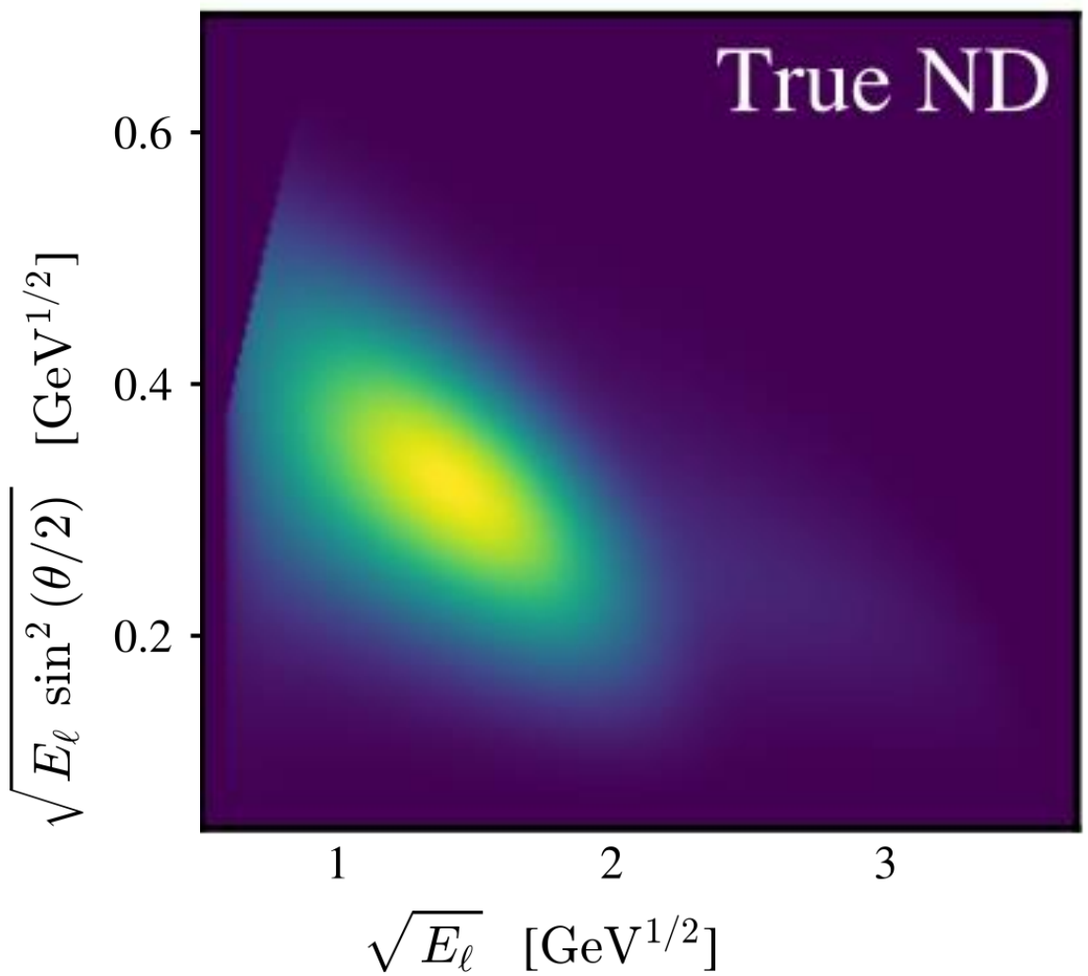
from true cross section model $d\sigma(E_\nu, E_l, \cos\theta)$ with FD flux(E_ν)

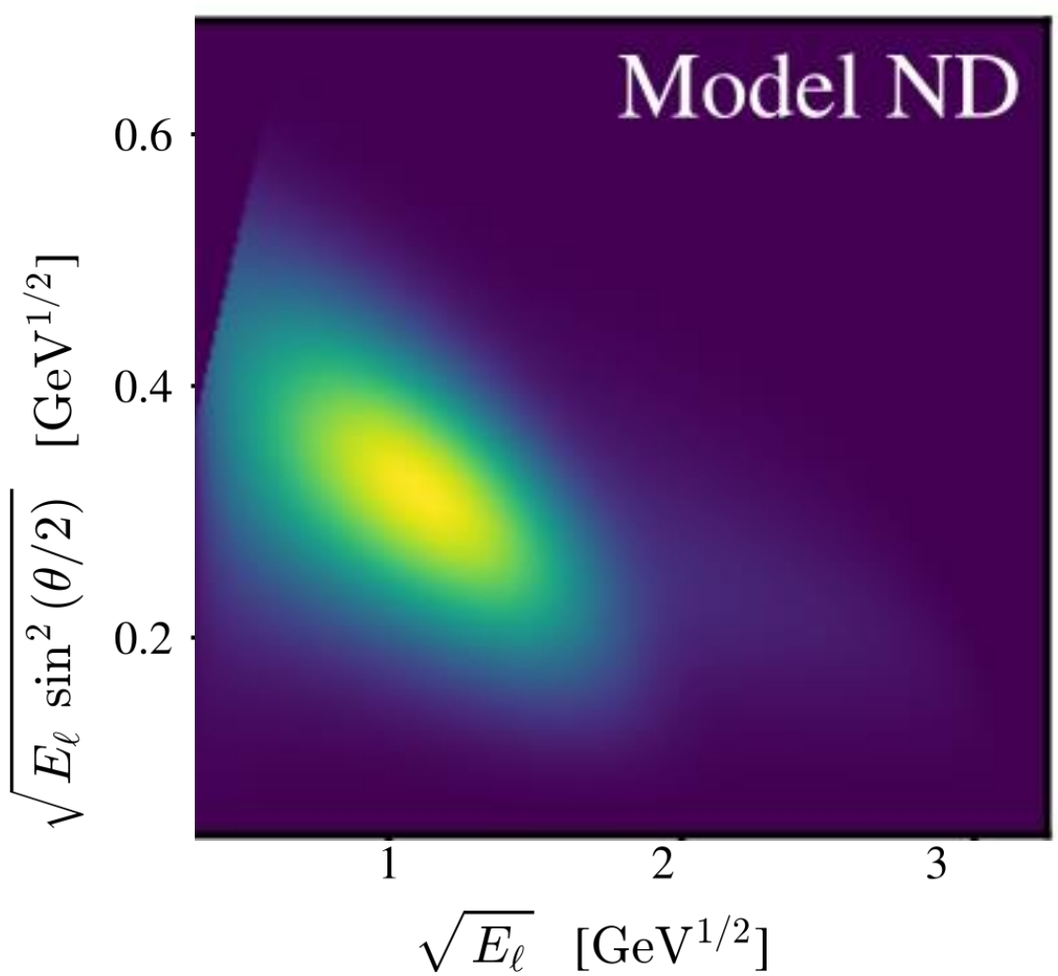
Maximum likelihood estimation

Oscillation results

How to learn νA — training

Neural SF xsec model can accurately reproduce ND distribution (training data)

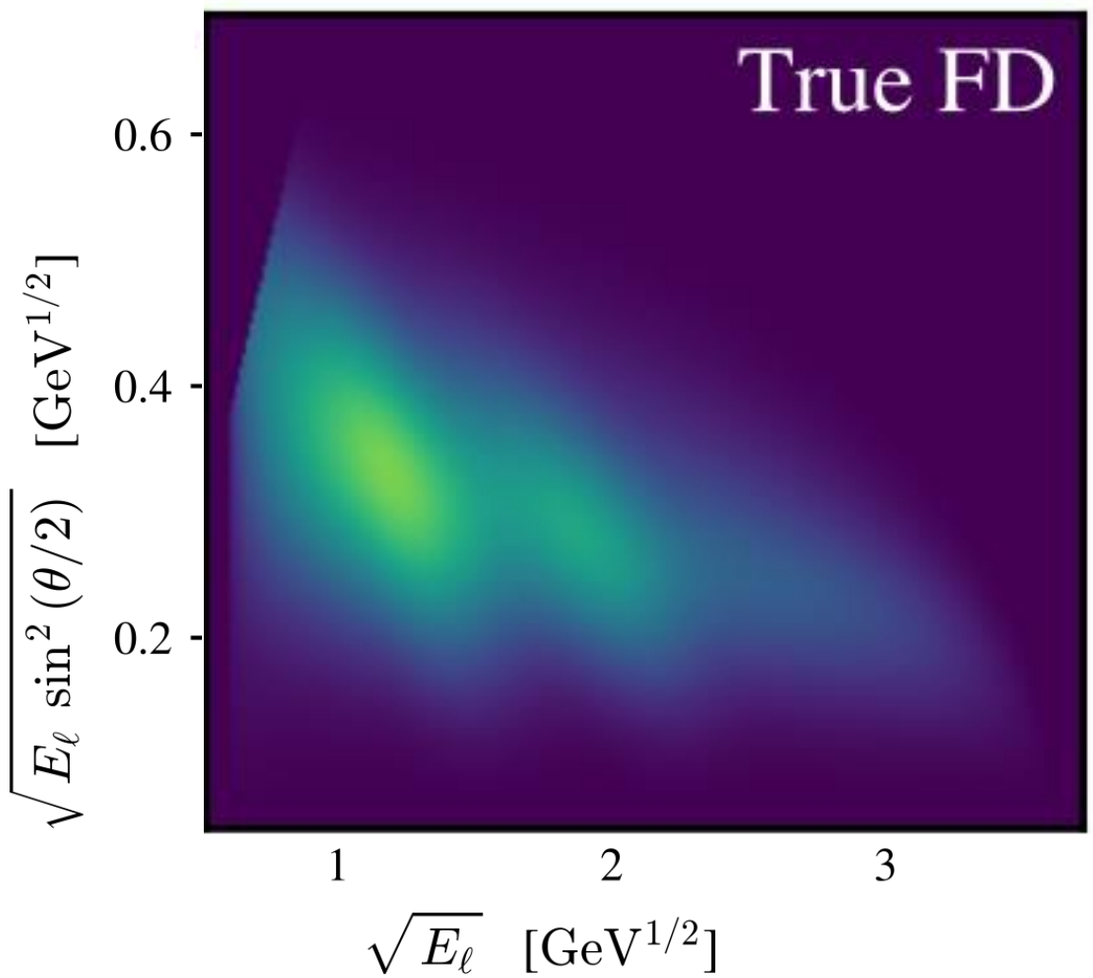


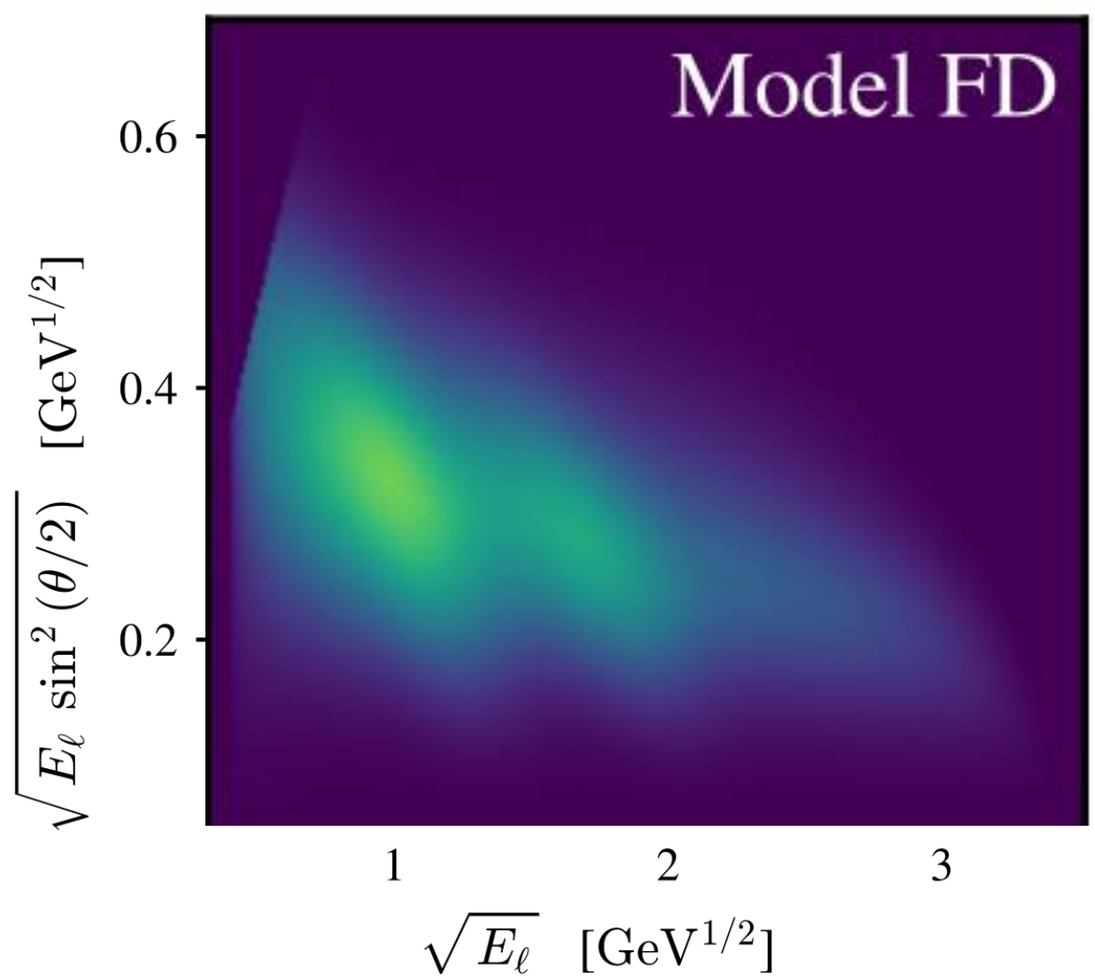


$$P_{ND} \propto \int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta}(E_\nu, E_l, \cos\theta)\Phi_{ND}(E_\nu)$$

How to learn νA — prediction

Neural SF xsec model **<u>can accurately generalize</u>** to FD distribution



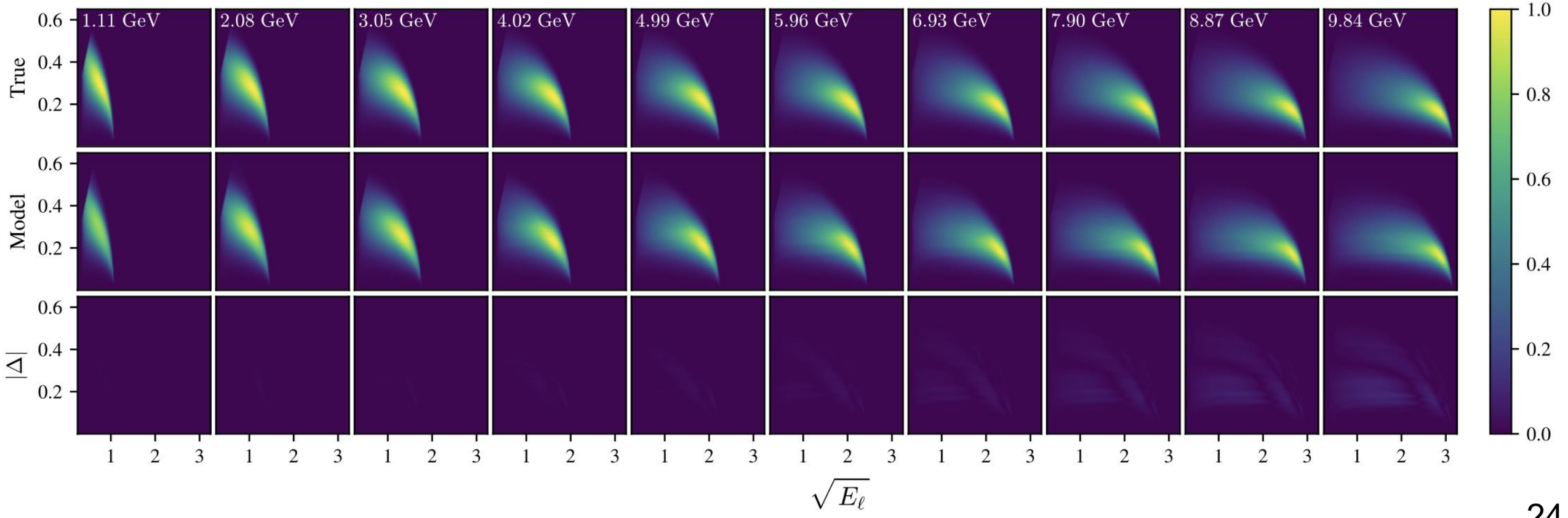


$$P_{FD} \propto \int dE_\nu \frac{d^2\sigma}{dE_l d\cos\theta}(E_\nu, E_l, \cos\theta)\Phi_{FD}(E_\nu)$$

3D xsec from 2D data + SF theory

Neural SF xsec model trained on ND data can accurately predict xsec at fixed E_ν over relevant range of kinematics

- 3D structure encoded by kinematic coefficients sufficient to enable 2D SFs learned using one E_ν slice to predict xsec for other E_ν slices



Does the model learn the true SFs?

Does the model learn the true SFs?

No

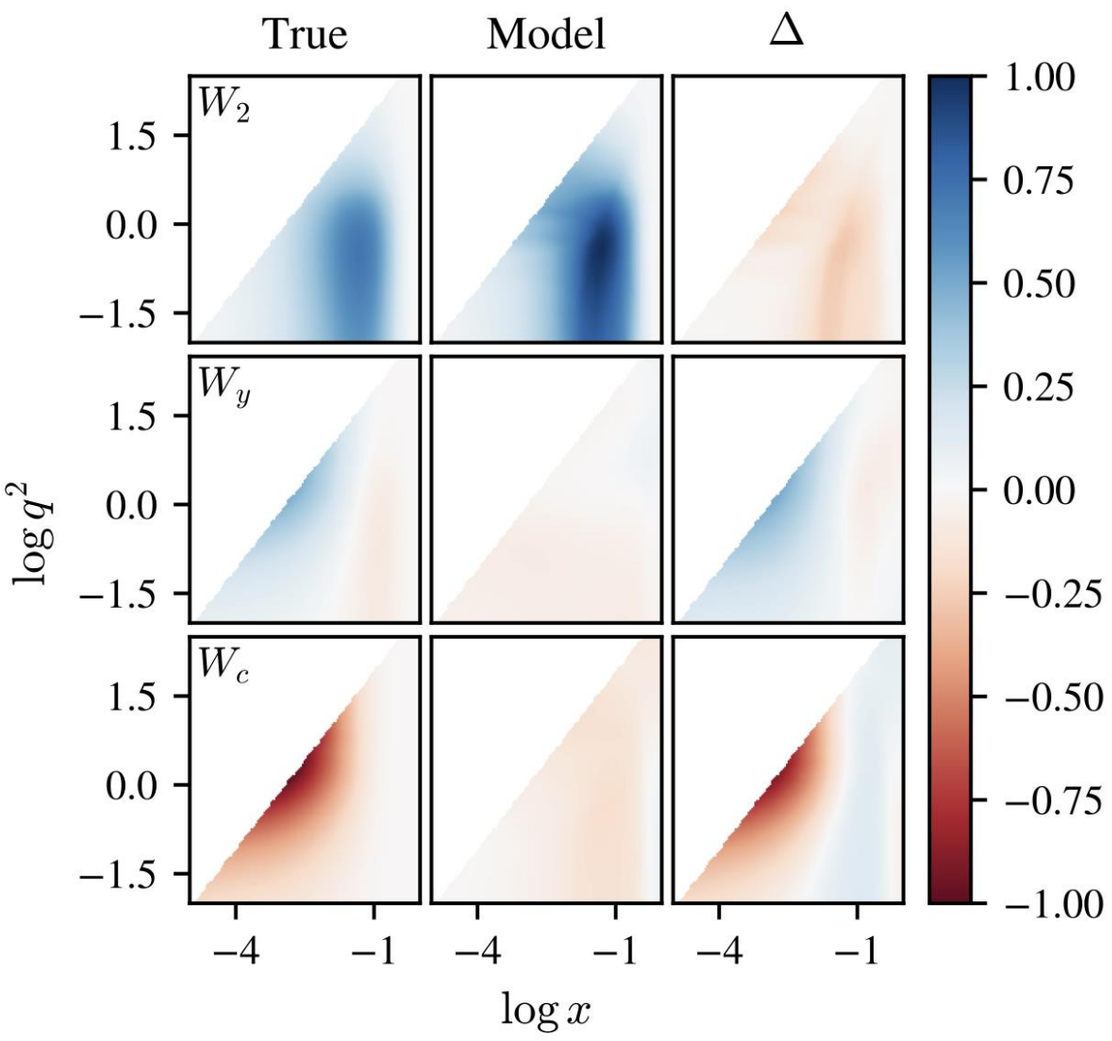
Does the model learn the true SFs?

No

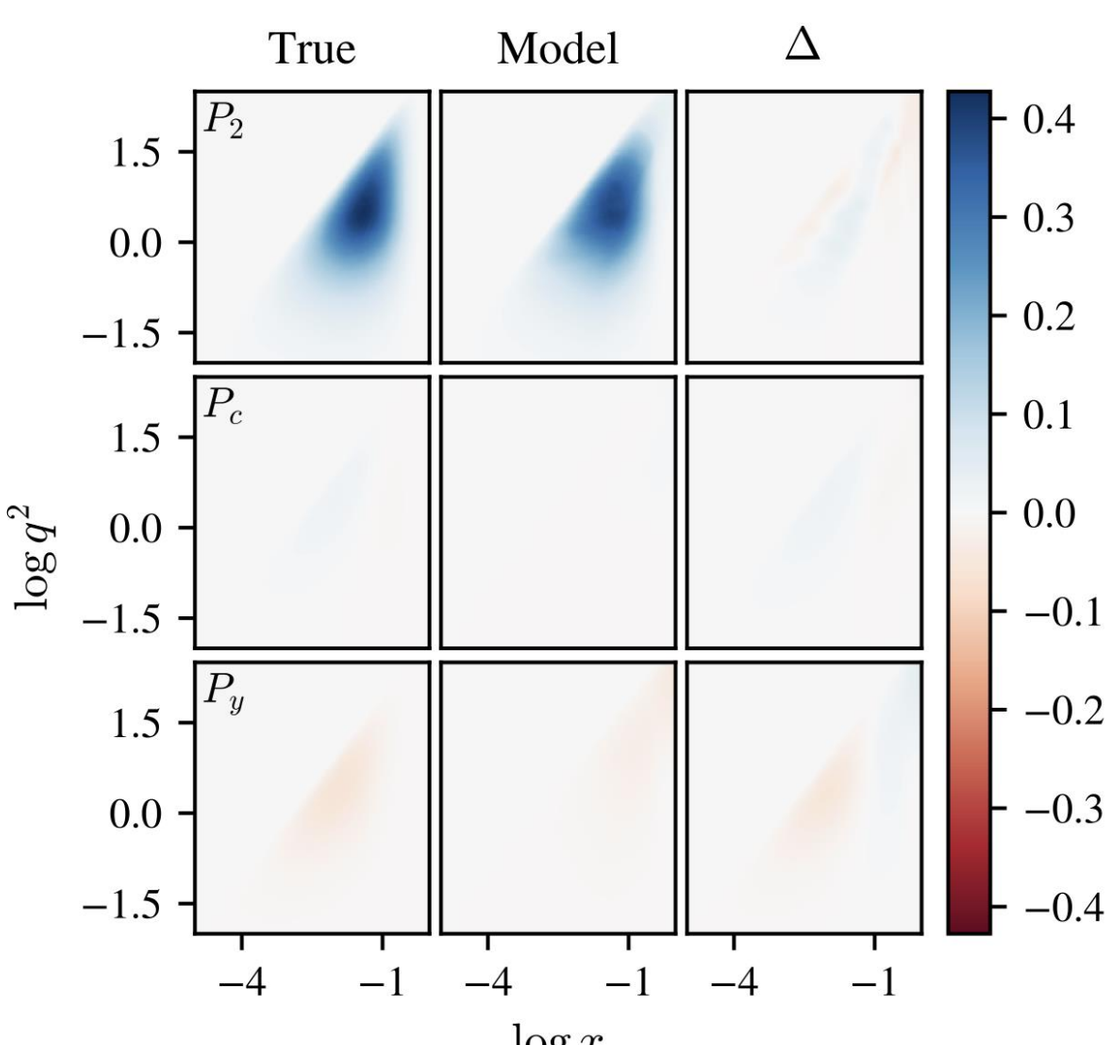
Weighting SFs by importance to ND kinematics shows neural model only learns the regions needed to predict ND

- Useful for oscillations if and only if same regions are important for ND and FD

Unweighted



Weighted

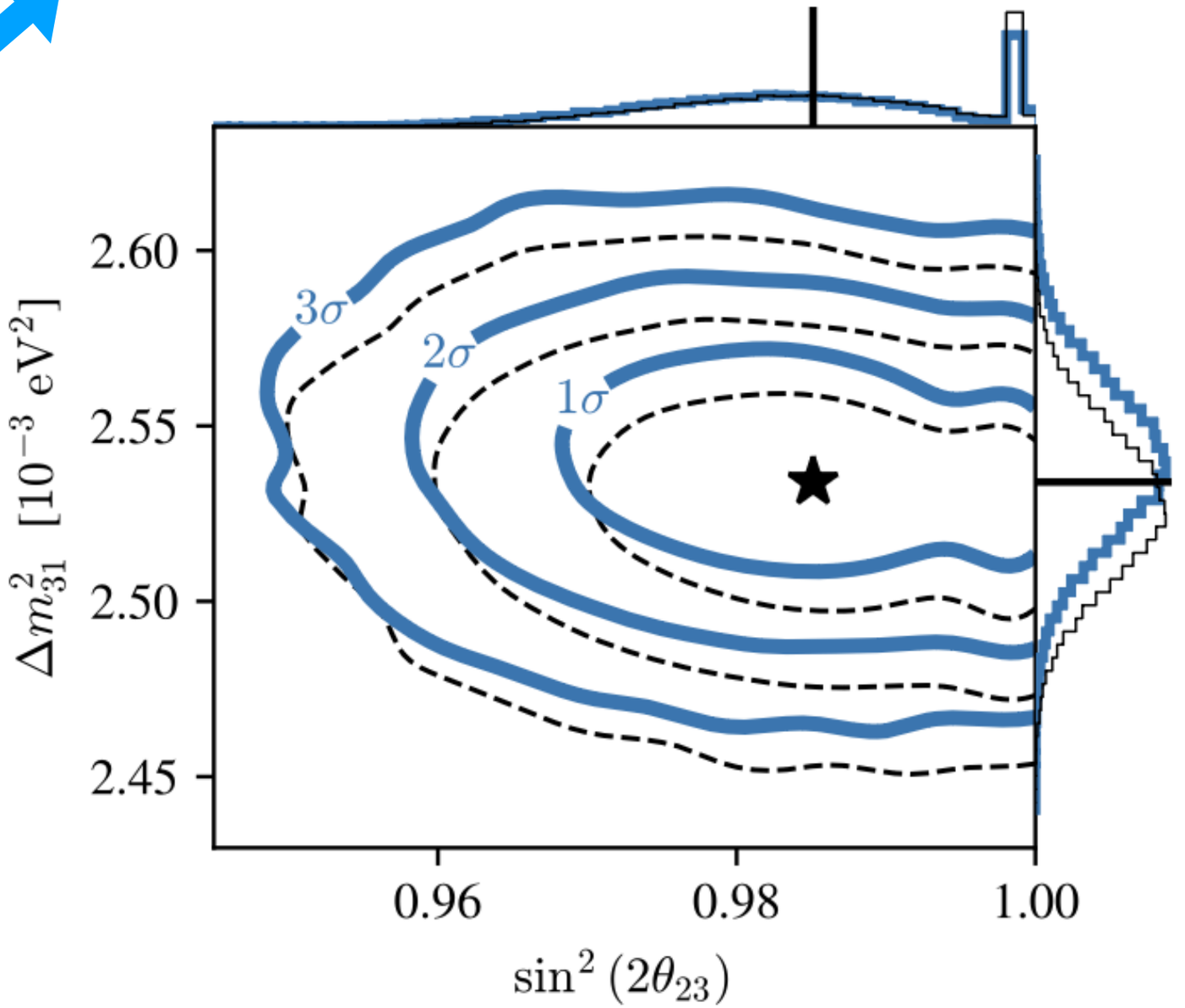


Oscillation analysis

Neural model trained on ND data held fixed for FD oscillation analsysis

$$\tilde{P}_{FD} \propto \int dE_\nu \frac{\tilde{d^2\sigma}}{dE_l d\cos\theta} \Phi_{ND}(E_\nu) P_{\mathrm{osc}}(E_\nu; \theta_{23}, \Delta m^2_{23}, \dots)$$

Agreement with FD data used to constrain oscillation parameters



Maximum likelihood estimation used for sample of 6,200 FD events ~ 3.5 years of DUNE

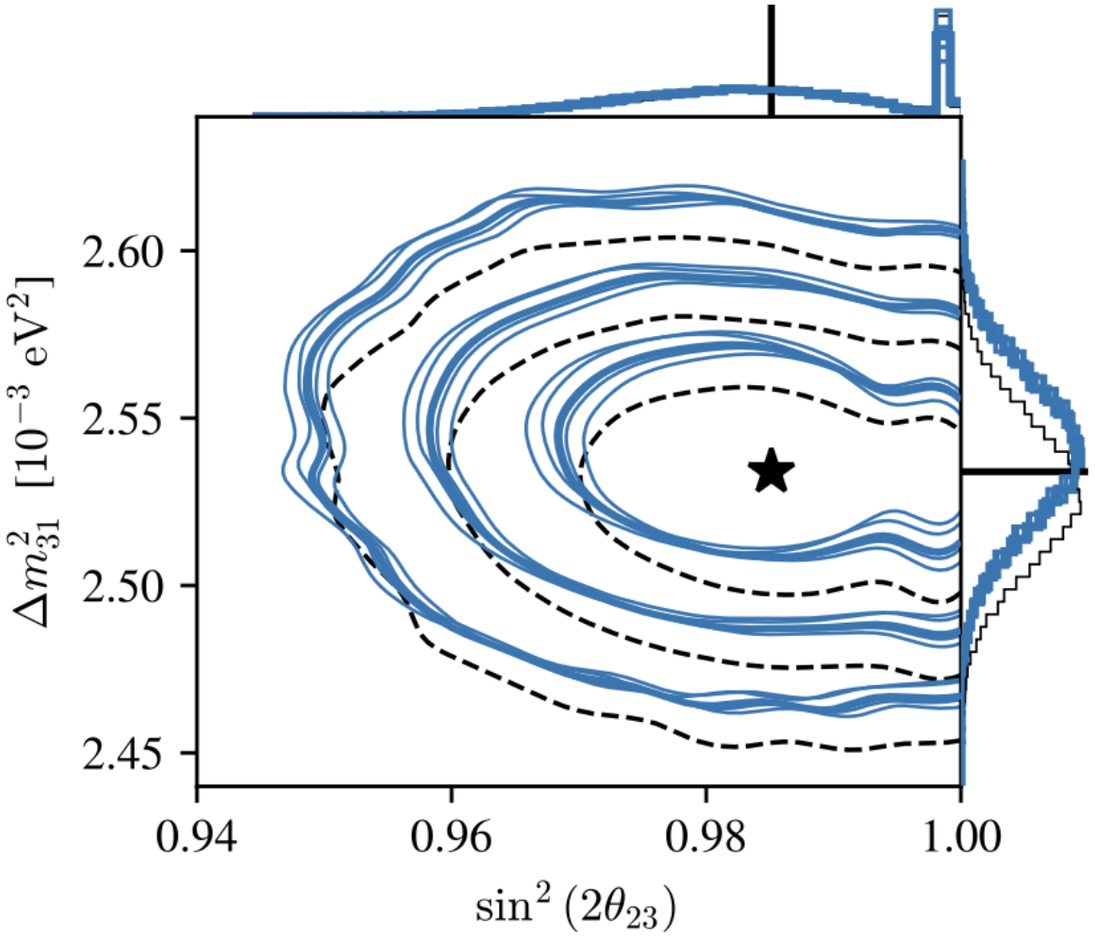
Oscillation constraints using learned xsec overlap with those obtained using true (toy model) xsec

Systematics 1: finite statistics

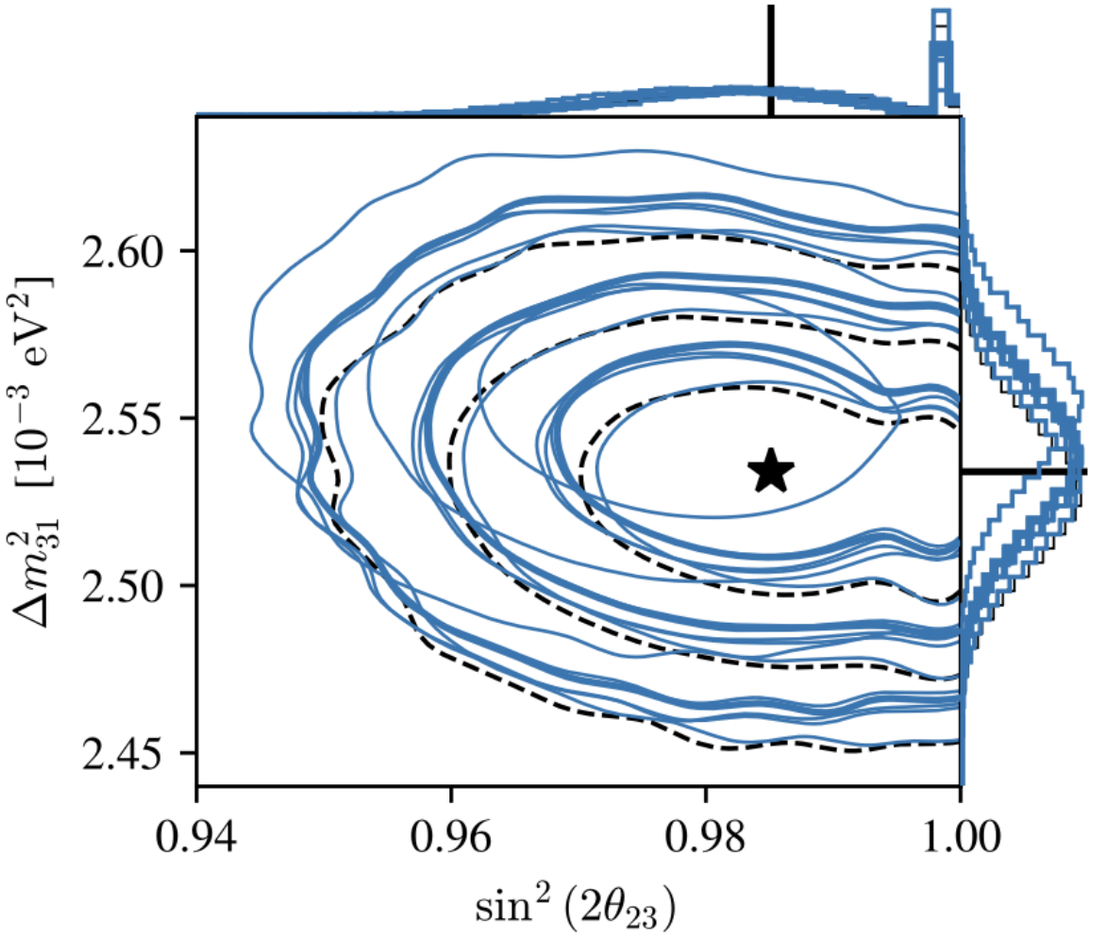
Previous results assumed infinite ND statistics, ignored experimental systematics

- ND training can be performed using histogram from finite sample of ND events
- Expected DUNE statistics look $\approx \infty$, an order of magnitude lower doesn't

30×10^6 ND events



3×10^6 ND events



Systematics 2: detector smearing

Detector effects "smear" events between nearby $(E_l, \cos\theta)$ bins

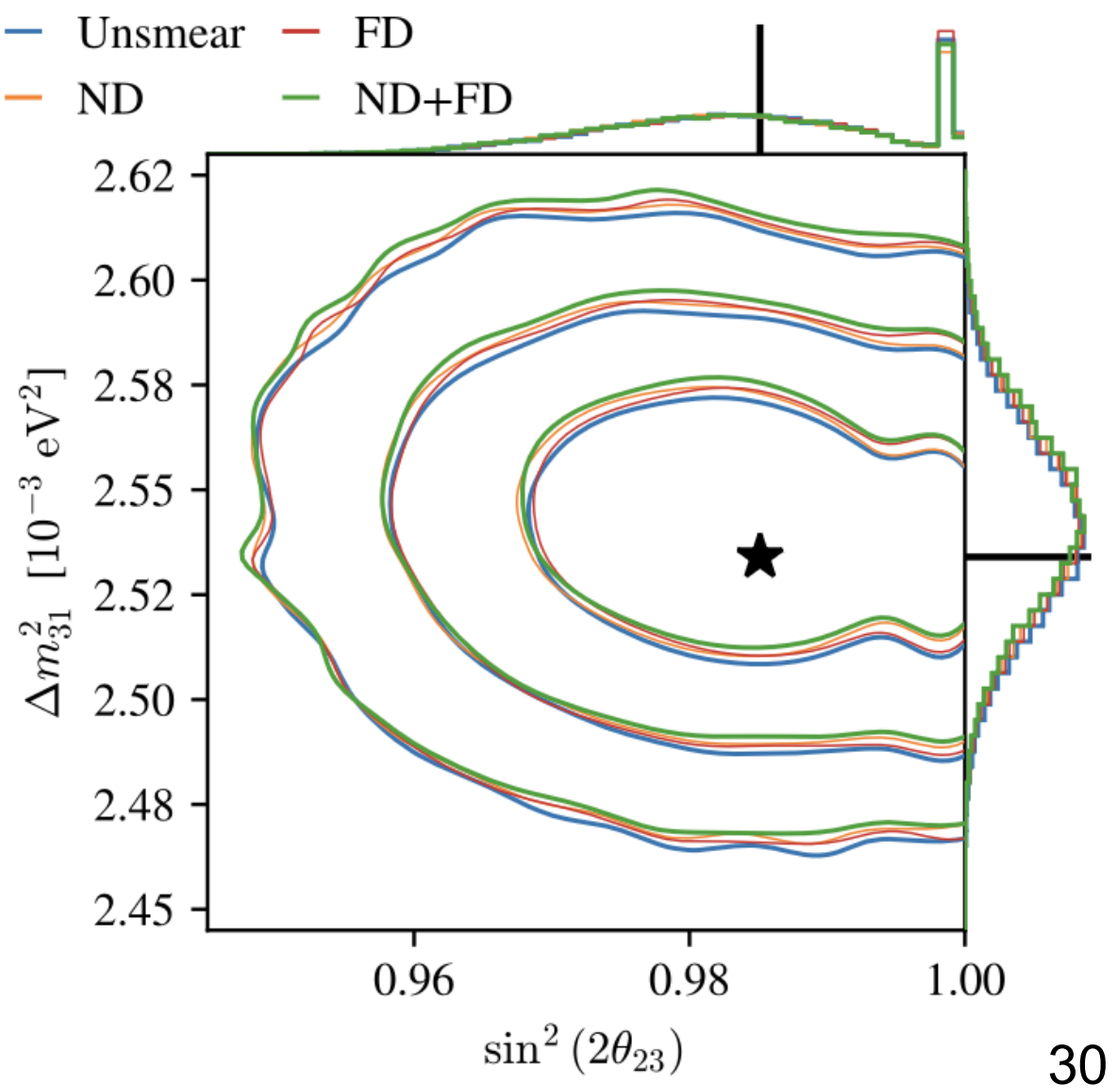
$$P(E_l, \cos\theta) = \int dE_l' d\cos\theta' S(E_l, \cos\theta | E_l', \cos\theta') P(E_l', \cos\theta')$$

- 4% expected E_l uncertainty incorporated through Gaussian smearing kernel

- 1° expected θ uncertainty negligible for histograms used in current training

Smearing added via "forward folding:" original event rate predictions convolved with smearing kernel and then compared to smeared data

- Small effect on final results



Systematics 3: neutrino flux shape

Flux normalization uncertainty cancels exactly when fitting event rates interpreted as (unit normalized) probability distributions

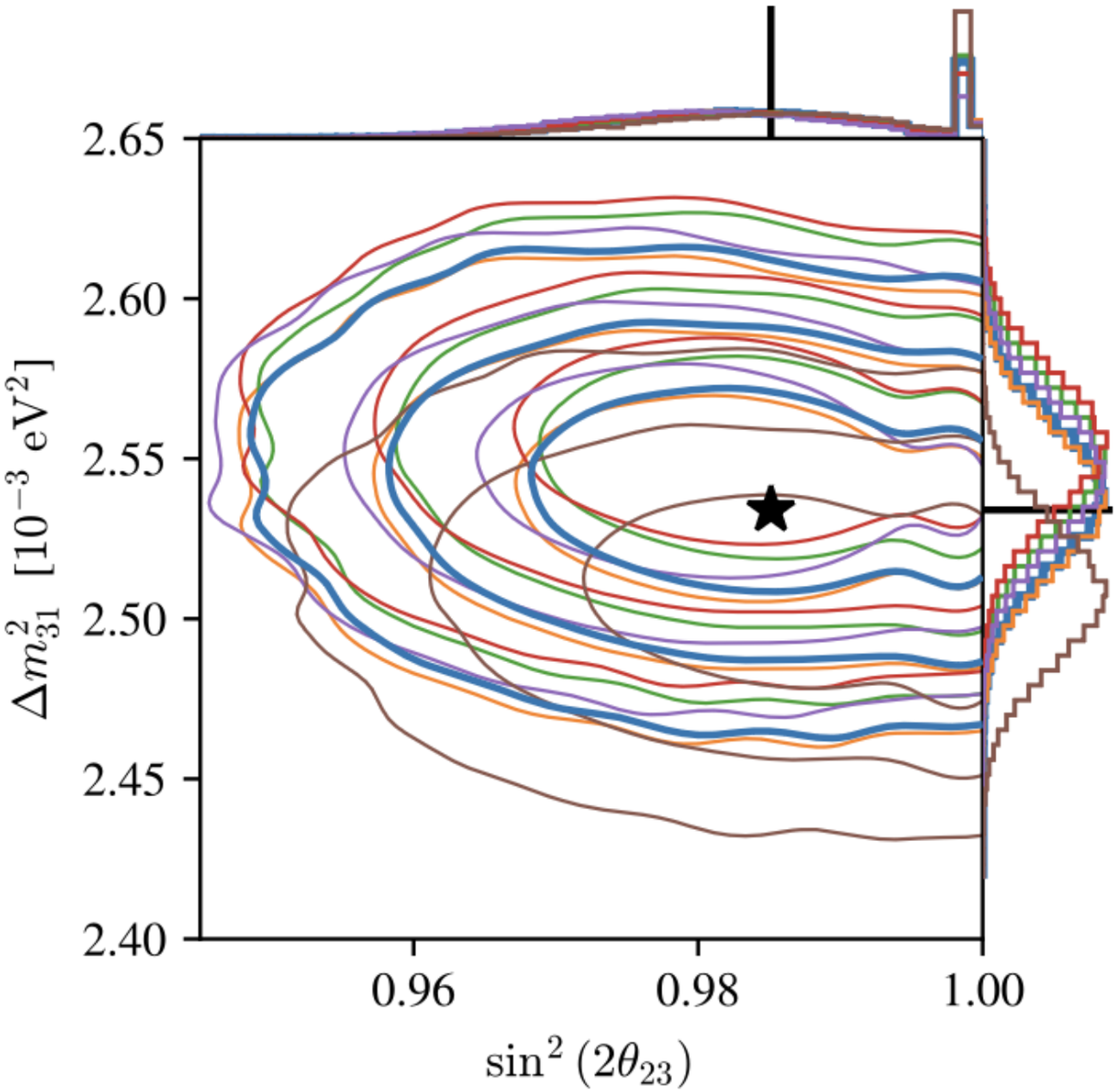
Flux shape uncertainties also arise at ~5% level

- Flux sampled with uncertainties added via DUNE collaboration flux uncertainty model

Abi et al [DUNE], Eur. Phys. J. C 80, 978 (2020)

Flux shape uncertainties can have significant effects on oscillation analysis **with both true and learned cross sections**

- Limitations arise from only using inclusive data rather than ML training procedure



Systematics 4: NN initialization

Variations due to random seed used in NN initialization larger than finite statistics or detector smearing effects, smaller than flux shape

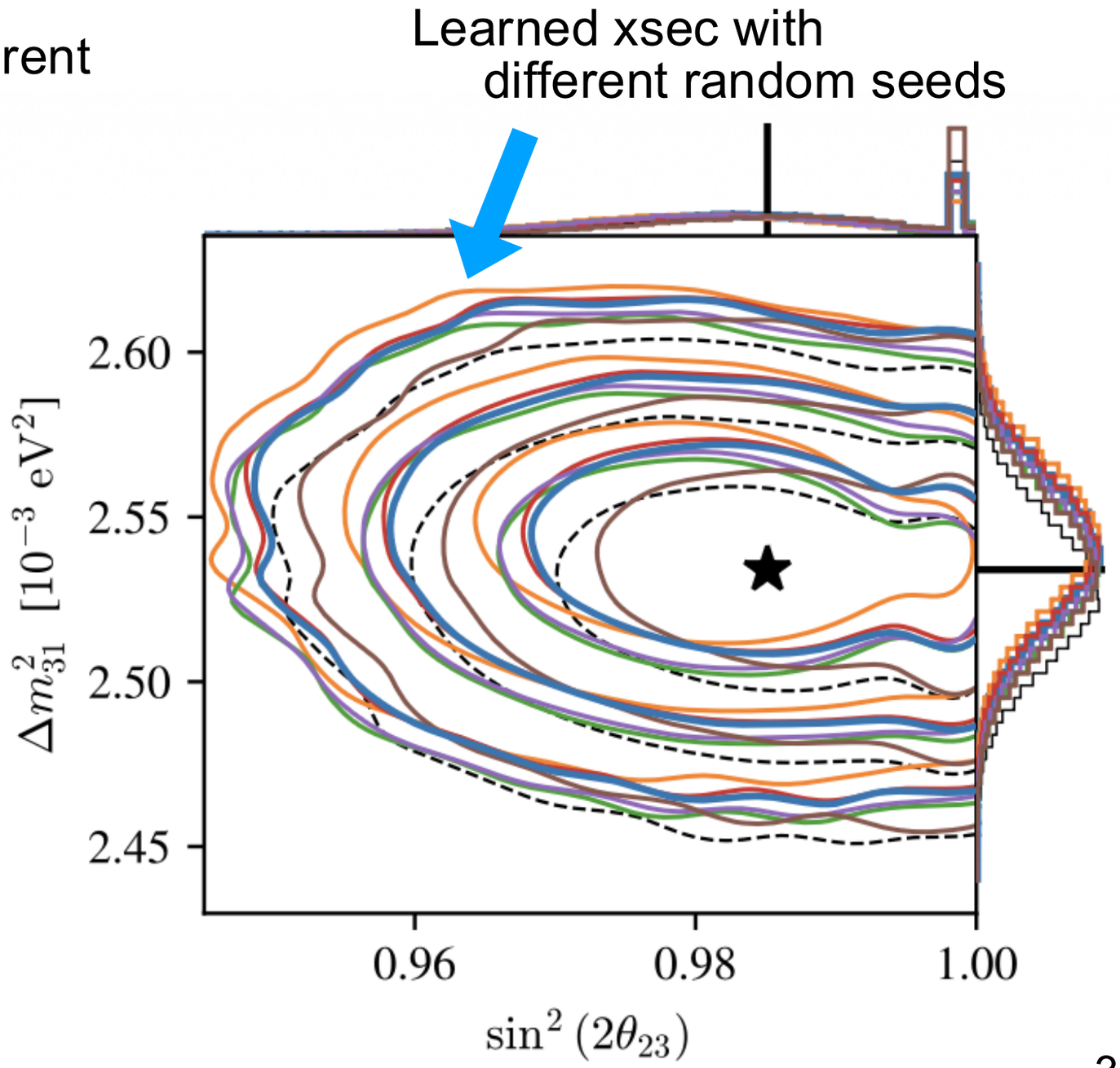
Ensemble of NNs trained using different

1) random seeds,
2) flux draws
3) ND stats draws

Used for basic systematic uncertainty quantification

- 1σ confidence intervals defined as region contained within 84% of individual 1σ CIs

- 2σ CIs from region within 95% of individual 2σ CIs



Oscillation analysis + systematics

Systematic uncertainty quantification from ensemble of NNs trained using different

1) random seeds
2) flux draws
3) ND stats draws

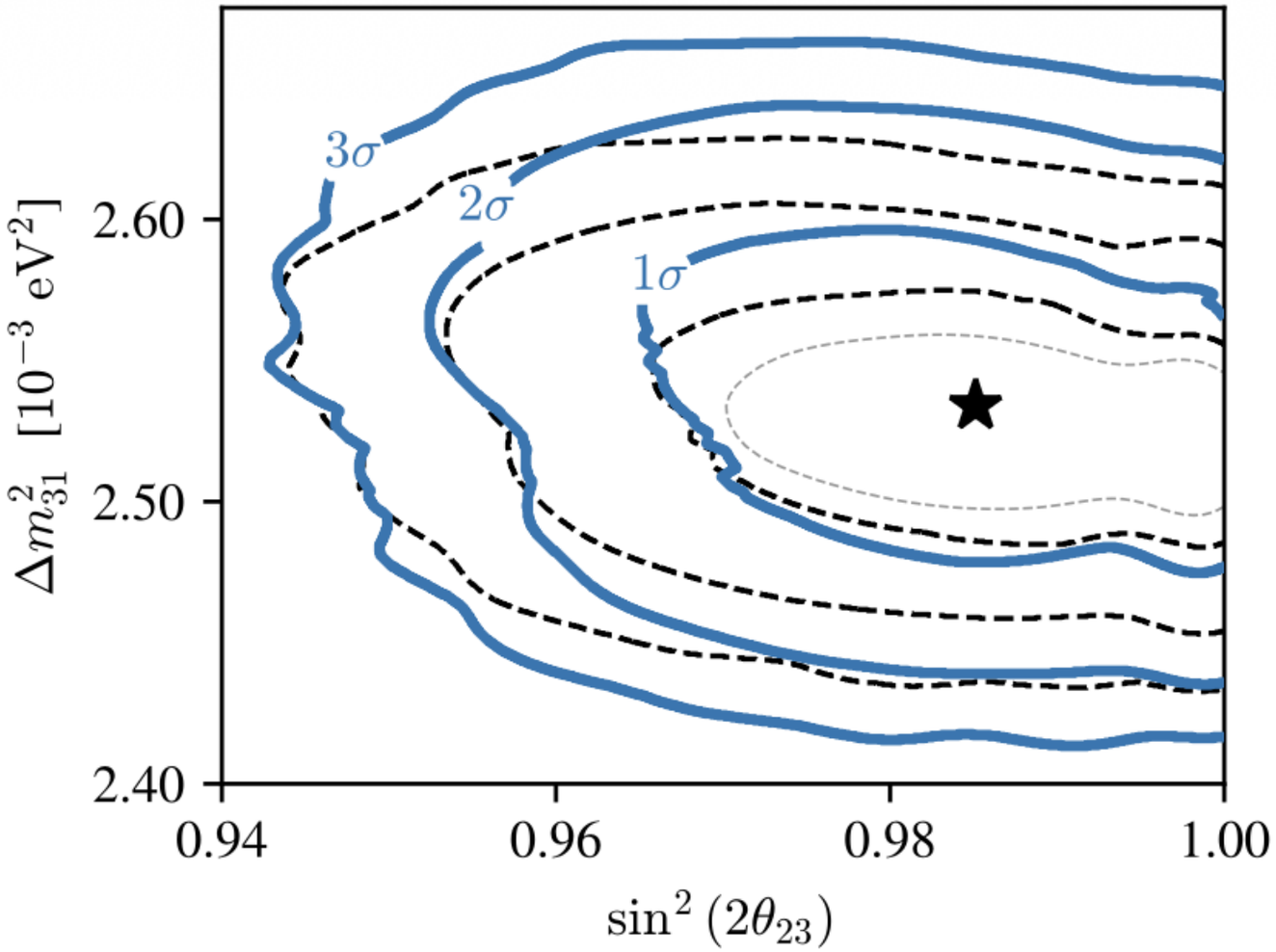
Final $n\sigma$ CIs defined as regions contained in fraction f_n of the $n\sigma$ CIs for each ensemble

$$f_n = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\mathrm{n}}{\sqrt{2}}\right)\right]$$

- Motivated by fact that this rule induces linear addition of stat + sys uncertainties

$$2\sigma \approx 2\sigma_{\mathrm{stat}} + 2\sigma_{\mathrm{sys}}$$

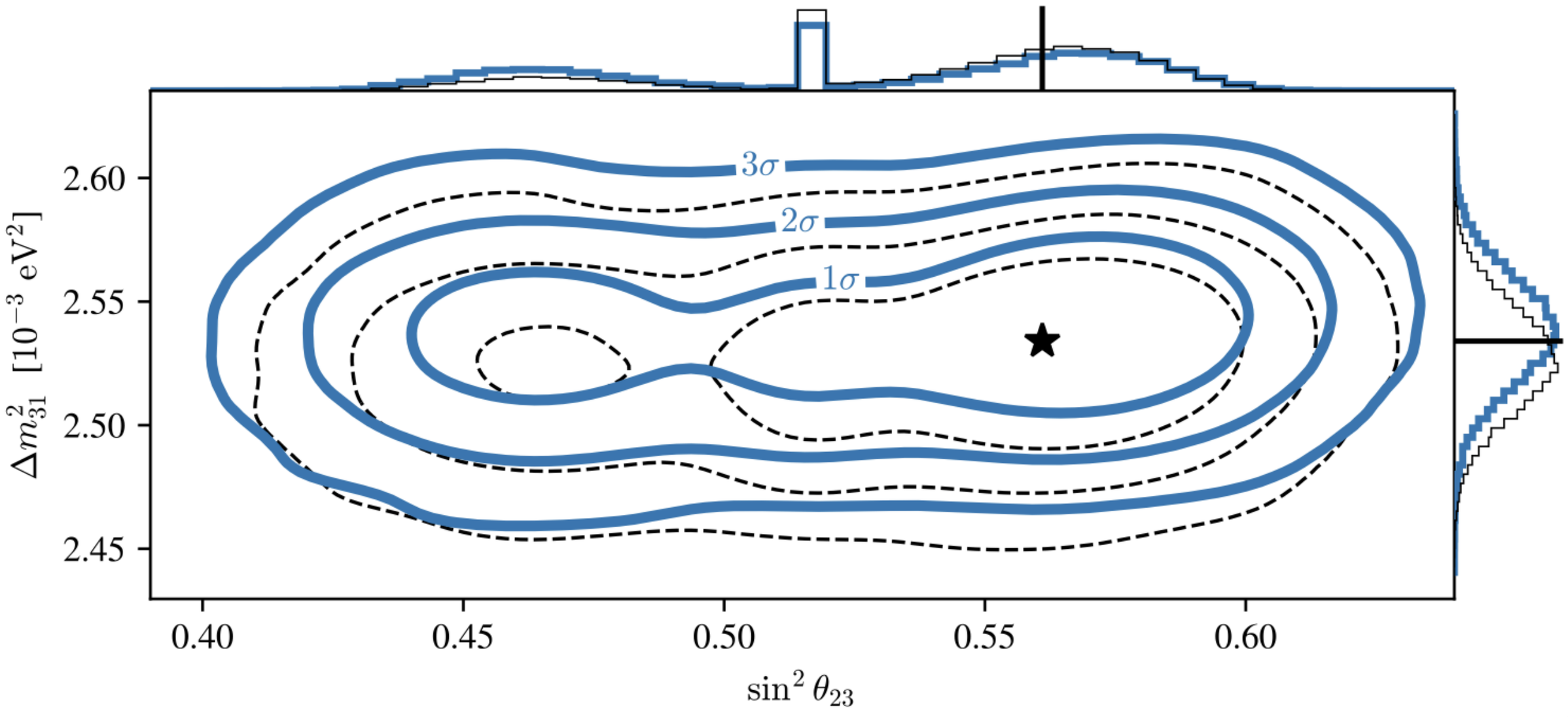
- Empirically well-calibrated: true (toy model) CIs fully contained by learned CIs with systematics



Next steps: octant degeneracy

Muon disappearance channel alone cannot determine octant of θ_{23}

- Generic problem not related to ML approach

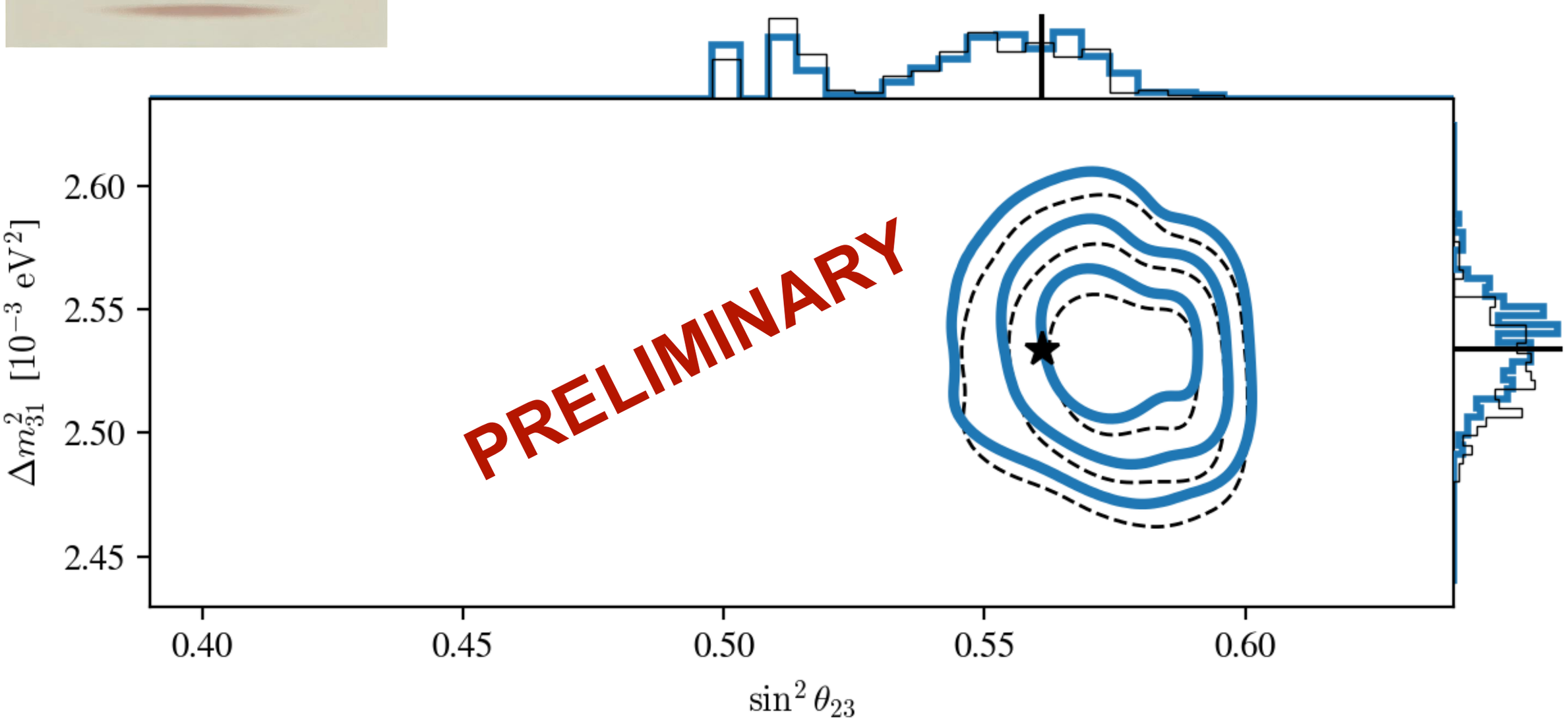


Next steps: more leptons

Electron appearance requires full set of 5 SFs

- Breaks octant degeneracy

Antineutrino samples for *CP* violation require independent set of 5 new SFs for non-isoscalar nuclei like argon



Thanks to my collaborators

Dan Hackett (Fermilab)

Josh Isaacson (MSU)

Shirley Li (UC Irvine)

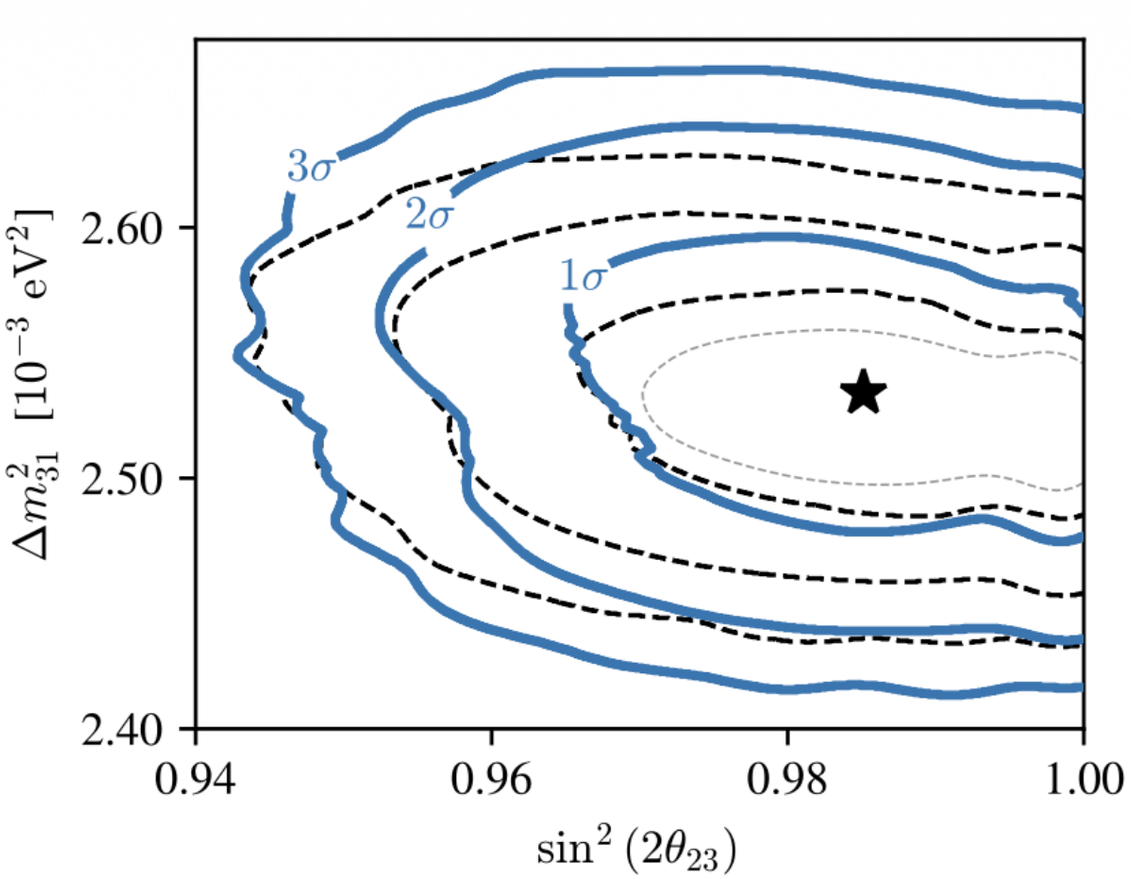
Karla Tame Narveaz (Fermilab)

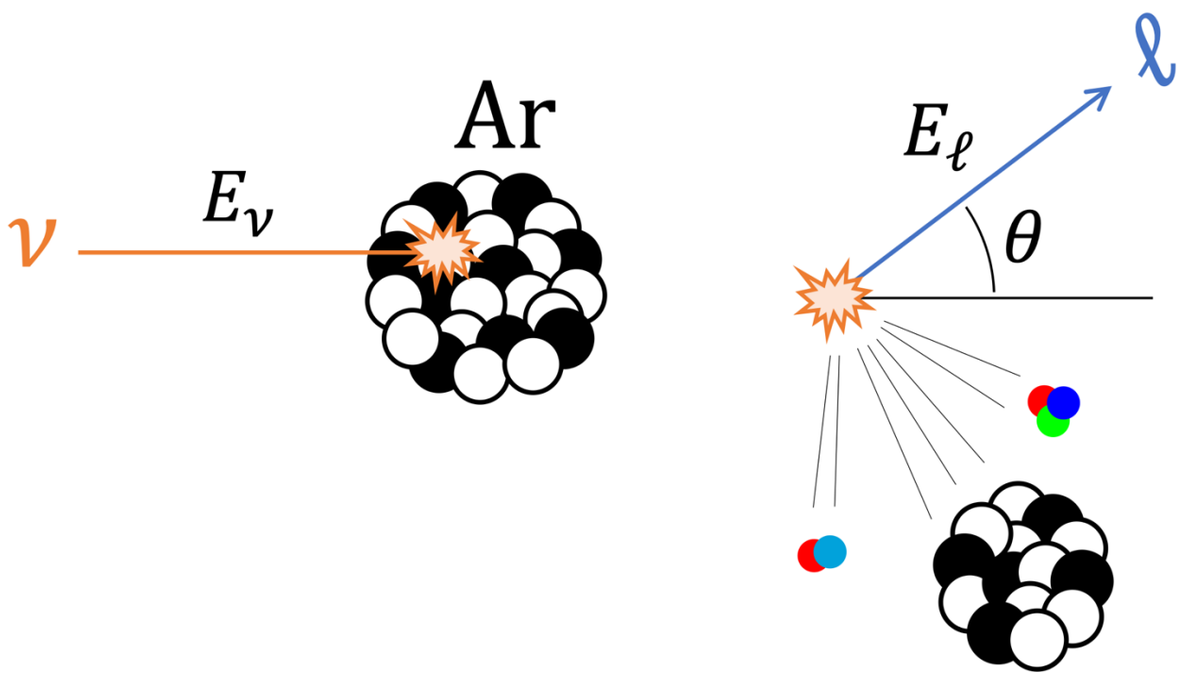
Outlook

Closure test demonstrates that neural SF models can (sometimes) accurately generalize from ND to FD data for DUNE kinematics

- What about a more realistic toy model including resonances?

- What about other experiments? Hyper-Kamiokande? MINERvA? NOvA? T2K? SBND?





- Can we include exclusive hadronic final states?

- Can we incorporate some nuclear theory priors?