Fermilab

Machine Learning Neutrino-Nucleus Cross Sections

Dan Hackett

Probing the frontiers of nuclear physics with AI at the EIC (II)
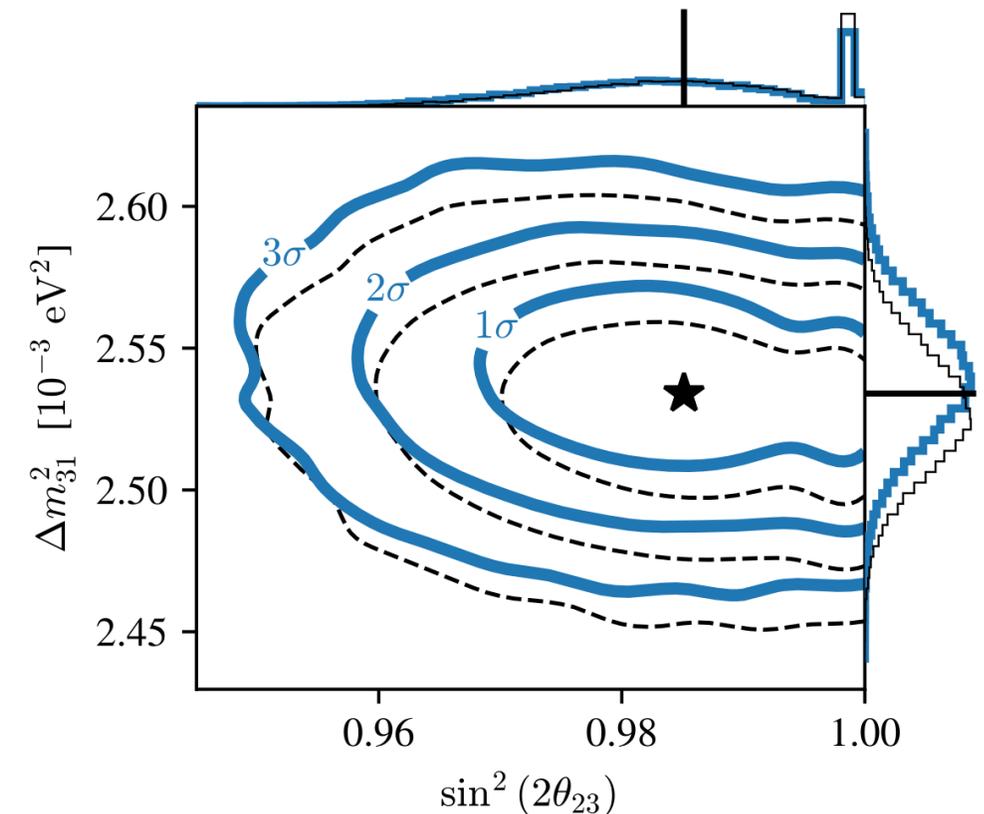CFNS, Stony Brook
March 19, 2025

Summary

Extracting oscillation parameters at DUNE requires model of $\nu - \mathrm{Ar}$ differential cross section

Q: Can we ML a cross-section model?

...from DUNE ND data

...well enough to do an oscillation analysis

A: Yes

Closure test **passes!**

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 events, finding that the modeled cross section achieves results consistent with what could be obtained if the true cross section were known exactly. This proof-of-principle study highlights the potential of future neutrino near-detector datasets and data-driven cross-section models.

Outline

Motivation
> DUNE & its nuclear theory challenges

Nuclear structure and $\nu - \mathrm{Ar}$ scattering
> Structure function (SF) decomposition
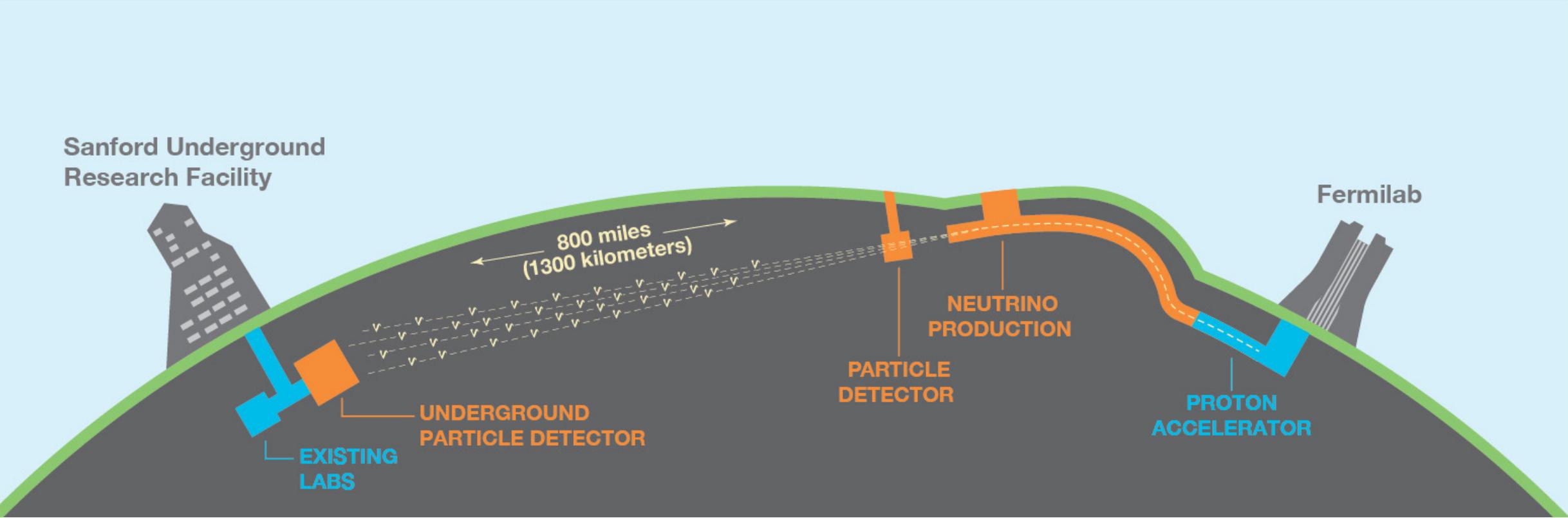
General approach
> Decompose into SFs
>
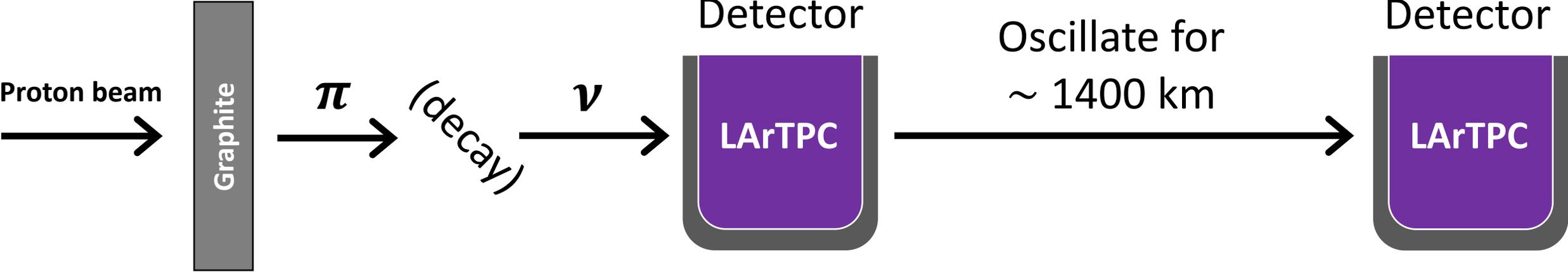> Parametrize SFs as NN

Closure test
> Learn cross-section on toy model of DUNE physics
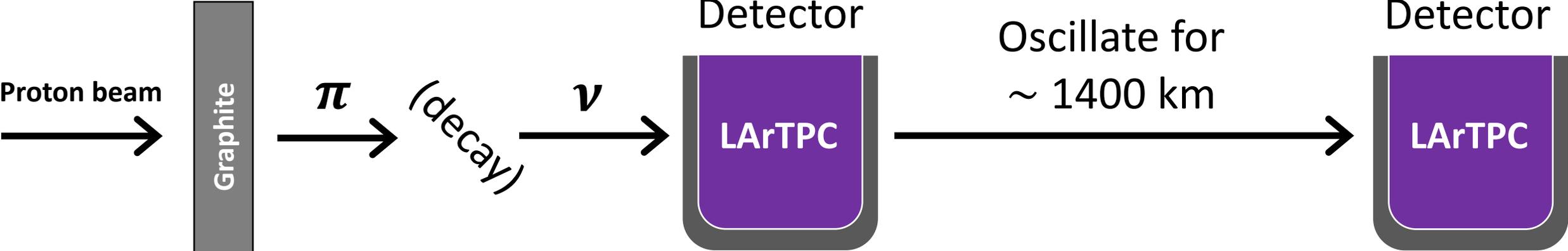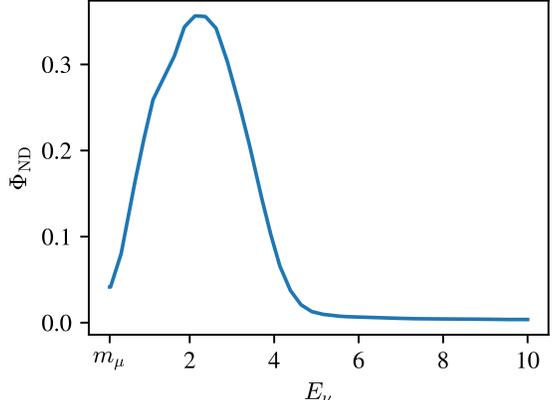
Motivation: DUNE

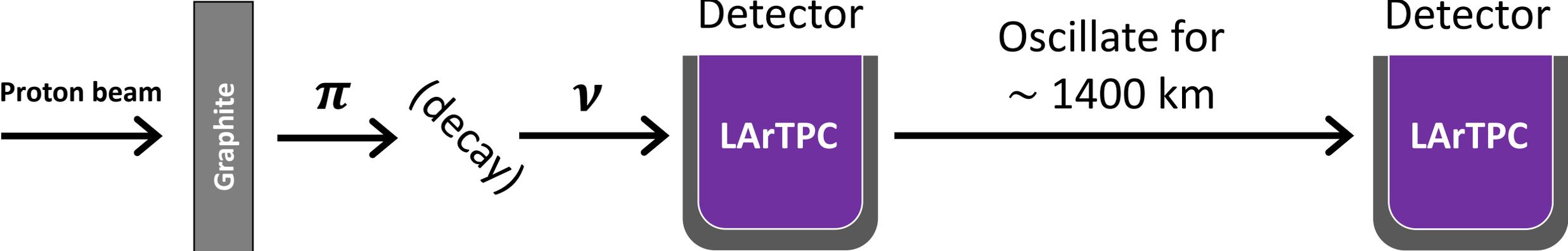Need to improve nuclear modeling!

DUNE, very schematically

Proton beam → Graphite → π → (decay) → ν → Near Detector LArTPC → Oscillate for ~ 1400 km → Far Detector LArTPC

DUNE, very schematically

Proton beam → Graphite → $\boldsymbol{\pi}$ → (decay) → $\boldsymbol{\nu}$ → Near Detector (LArTPC)

Oscillate for ∼ 1400 km

Far Detector (LArTPC)

Well-understood
→ accurately characterized
neutrino flux at ND

DUNE, very schematically

Proton beam → Graphite → $\boldsymbol{\pi}$ → (decay) → $\boldsymbol{\nu}$ → **Near Detector** LArTPC

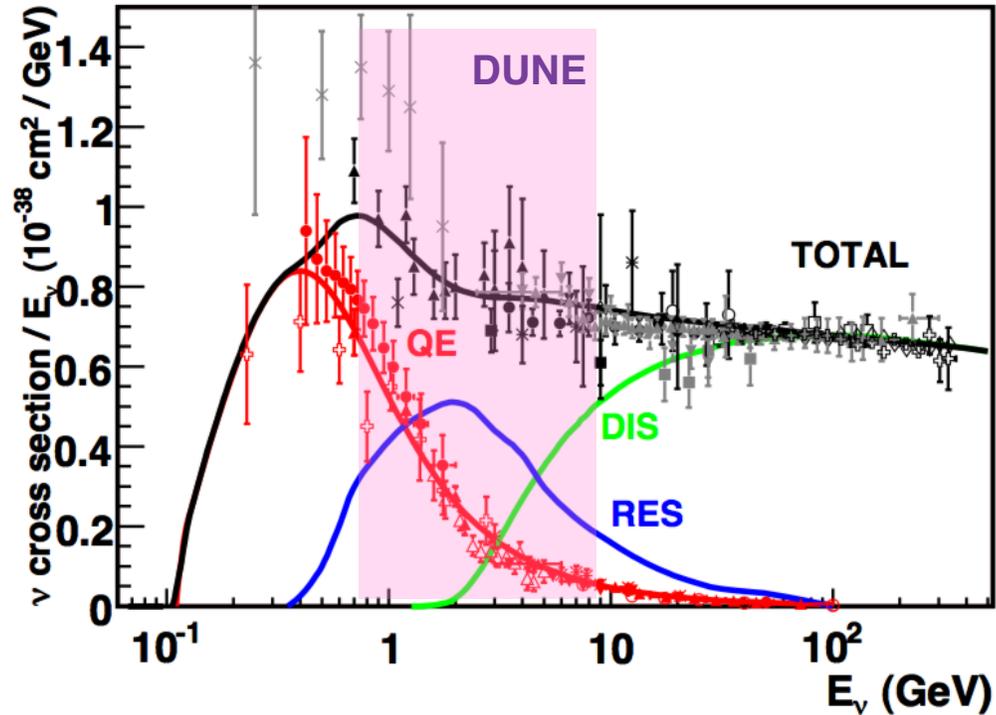Oscillate for ~ 1400 km → **Far Detector** LArTPC

Well-understood → accurately characterized neutrino flux at ND

QCD is hard

[1305.7513]

- Different mechanisms relevant for different E_ν
- Different theory frameworks for each
- Must add ad-hoc parameters to stitch together
  in event generators → "Generator tuning"
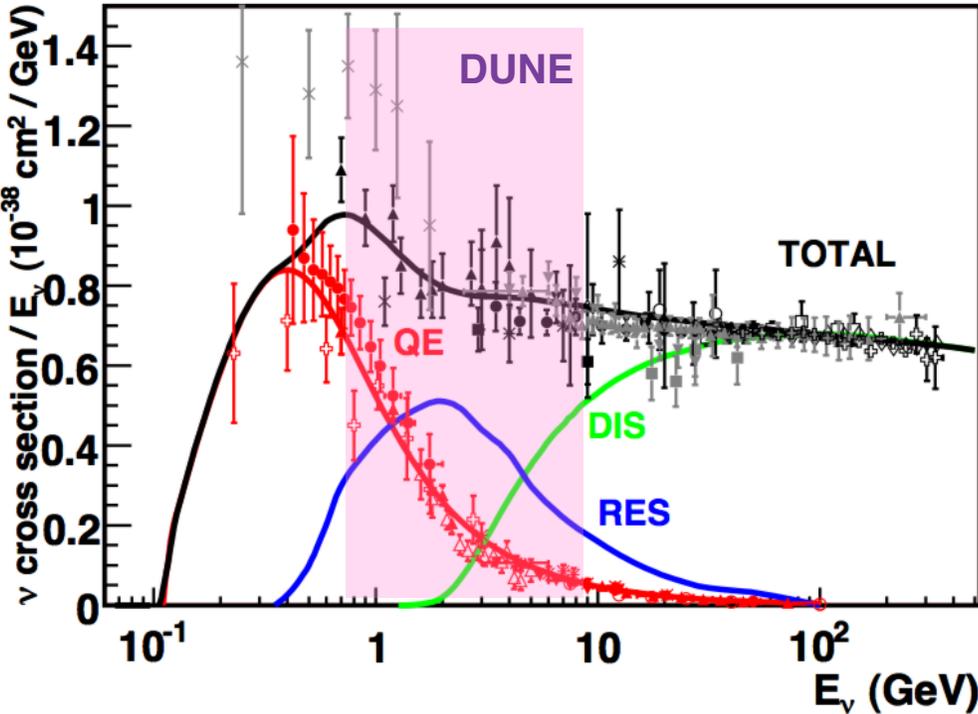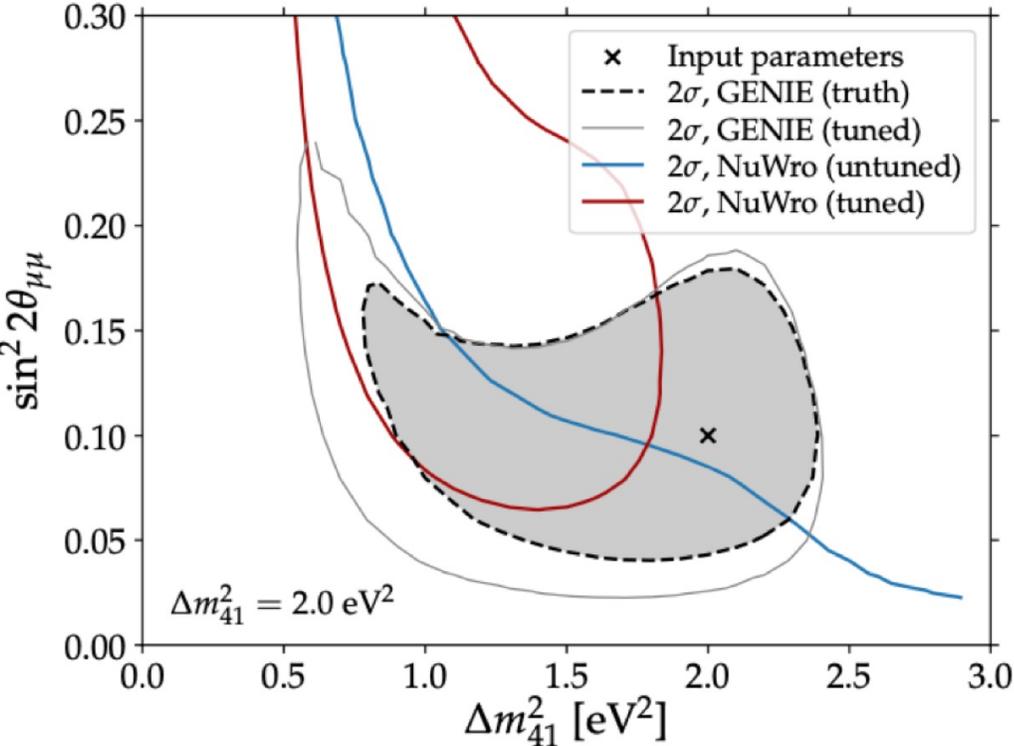
QCD is hard

[1305.7513]

[Coyle Li Machado 2210.03753]

- Different mechanisms relevant for different E_ν
- Different theory frameworks for each
- Must add ad-hoc parameters to stitch together in event generators → "Generator tuning"

- Generators encode a cross-section model
- When present generators are tuned on ND data, doesn't reliably generalize from ND to FD kinematics

Structure functions
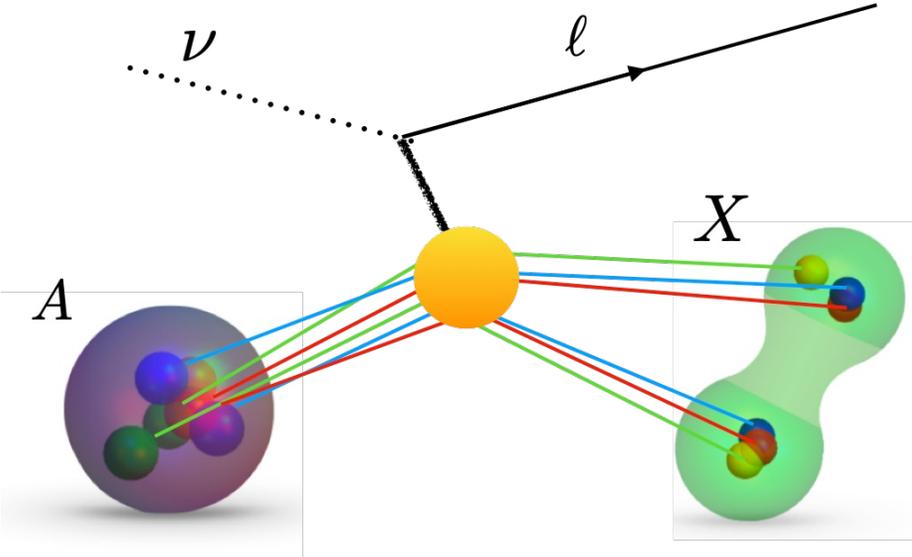
Cross section factorizes as

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

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

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

Structure functions

Cross section factorizes as

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

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

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

In more detail:

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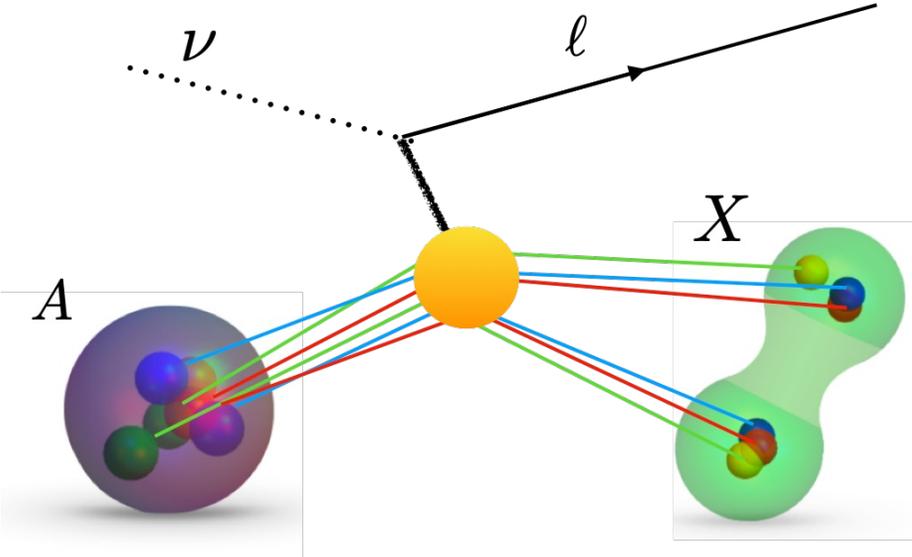
Structure functions

Cross section factorizes as

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

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

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

In more detail:

$$W_{\mu\nu} = W_1 g_{\mu\nu} + W_2 \frac{p_\mu p_\nu}{p^2} \pm W_3 \left(\frac{ip^\rho p^\sigma}{2p \cdot q}\right) \epsilon_{\mu\nu\rho\sigma} + W_4 \frac{q_\mu q_\nu}{q^2} - W_5 \frac{p_\mu q_\nu + q_\mu p_\nu}{p \cdot q}$$
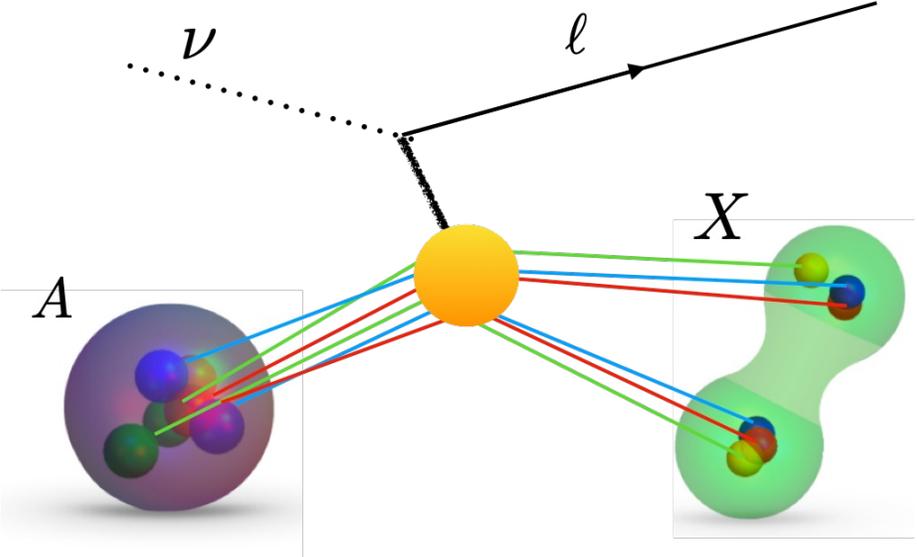
Usual DIS structure functions
$W_i = xF_i$ for $i \in \{1,3,4,5\}$ and $W_2 = \frac{2xM_A^2}{Q^2} F_2$

Parity-violating
Only for ν scattering

"Albright-Jarlskog functions"
Only for ν scattering
Suppressed by m_ℓ^2/Q^2

Structure functions

⇒ cross section decomposes into five SFs

$$\frac{d^2\sigma}{dE_\ell d\cos\theta}(E_\nu) = L^{\mu\nu}W_{\mu\nu} = \sum_{i=1}^{5} K_i(E_\nu, E_\ell, \cos\theta)\, W_i(x, Q^2)$$

Structure functions

⇒ cross section decomposes into five SFs

$$\frac{d^2\sigma}{dE_\ell d\cos\theta}(E_\nu) = L^{\mu\nu} W_{\mu\nu} = \sum_{i=1}^{5} K_i(E_\nu, E_\ell, \cos\theta)\, W_i(x, Q^2)$$

Specifically:

$$\boxed{\tilde{y} \equiv y\left(1 + \frac{m_\ell^2}{Q^2}\right)}$$

$$\frac{d^2\sigma}{dE_\ell d\cos\theta}(E_\nu) = \frac{|V_{ud}|^2 G_F^2}{\pi} \sqrt{E_\ell^2 - m_\ell^2} \Bigg\{$$

$$\tilde{y} W_1(x, Q^2) + \frac{E_\nu}{M_A}\left(1 - y - \frac{Q^2 + m_\ell^2}{4E_\nu^2}\right) W_2(x, Q^2) + \left(1 - \frac{\tilde{y}}{2}\right) W_3(x, Q^2) - \left(\frac{m_\ell^2}{Q^2}\right)[2W_5(x, Q^2) - \tilde{y} W_4(x, Q^2)]\Bigg\}$$

Structure functions

$\Rightarrow$ cross section decomposes into five SFs

$$\underbrace{\frac{d^2\sigma}{dE_\ell d\cos\theta}(E_\nu)}_{\text{3d}} = \textcolor{blue}{L^{\mu\nu}}\textcolor{green}{W_{\mu\nu}} = \sum_{i=1}^{5} \underbrace{\textcolor{blue}{K_i(E_\nu, E_\ell, \cos\theta)}}_{\text{3d}} \underbrace{\textcolor{green}{W_i(x, Q^2)}}_{\text{2d}}$$

Note: 3d cross section, but SFs are 2d
$\Rightarrow$ Can learn 3d cross section from 2d data!

Specifically:

$$\frac{d^2\sigma}{dE_\ell d\cos\theta}(E_\nu) = \textcolor{blue}{\frac{|V_{ud}|^2 G_F^2}{\pi}}\sqrt{E_\ell^2 - m_\ell^2}\Bigg\{ \qquad\qquad \tilde{y} \equiv y\left(1 + \frac{m_\ell^2}{Q^2}\right)$$

$$\textcolor{blue}{\tilde{y}}\textcolor{green}{W_1(x, Q^2)} + \textcolor{blue}{\frac{E_\nu}{M_A}\left(1 - y - \frac{Q^2 + m_\ell^2}{4E_\nu^2}\right)}\textcolor{green}{W_2(x, Q^2)} + \textcolor{blue}{\left(1 - \frac{\tilde{y}}{2}\right)}\textcolor{green}{W_3(x, Q^2)} - \textcolor{blue}{\left(\frac{m_\ell^2}{Q^2}\right)}[2\textcolor{green}{W_5(x, Q^2)} - \textcolor{blue}{\tilde{y}}\textcolor{green}{W_4(x, Q^2)}]\Bigg\}$$

General Approach

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

Model cross section

Known kinematic coefficients

General Approach

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

Model cross section

Known kinematic coefficients

Related work:

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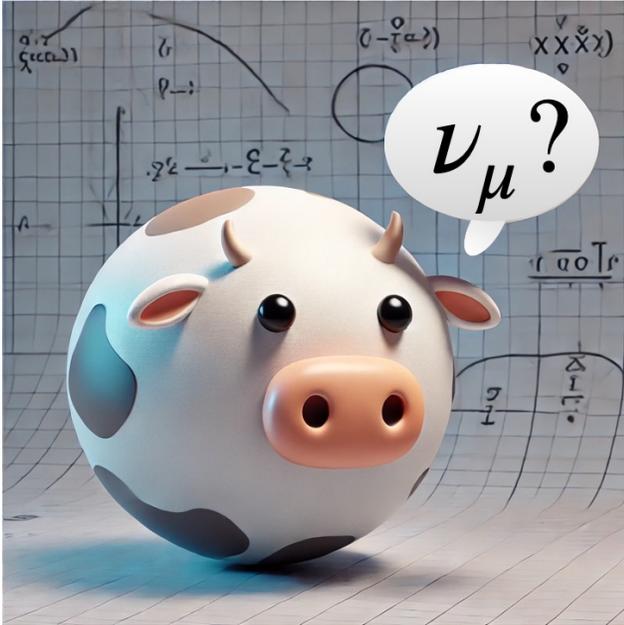
arXiv:hep-ph/0204232 (hep-ph)

[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: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

~ similar strategy to NNPDF, but with less nuclear modeling / theory inputs

Closure test(?)

Before applying a method to an unknown system, first check if it works on a known system
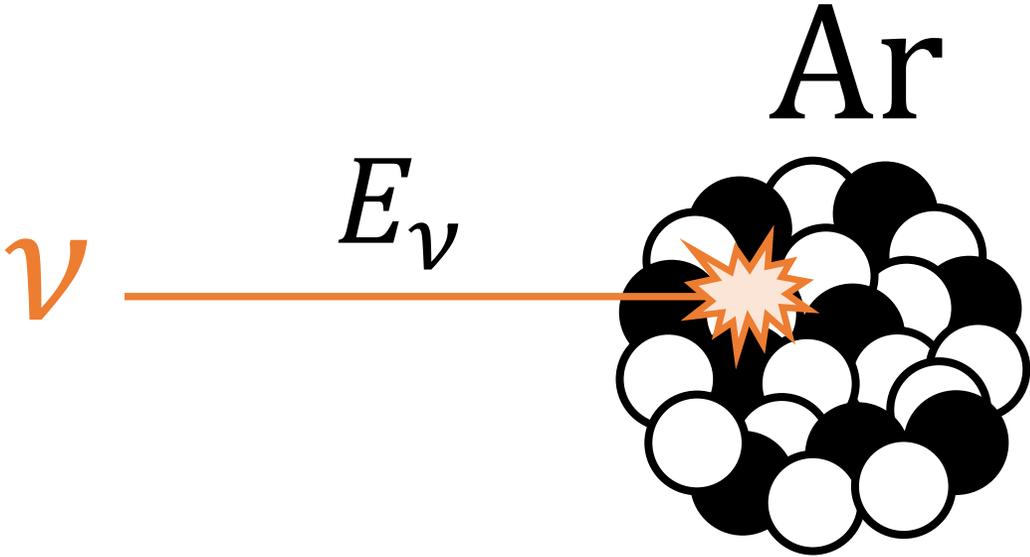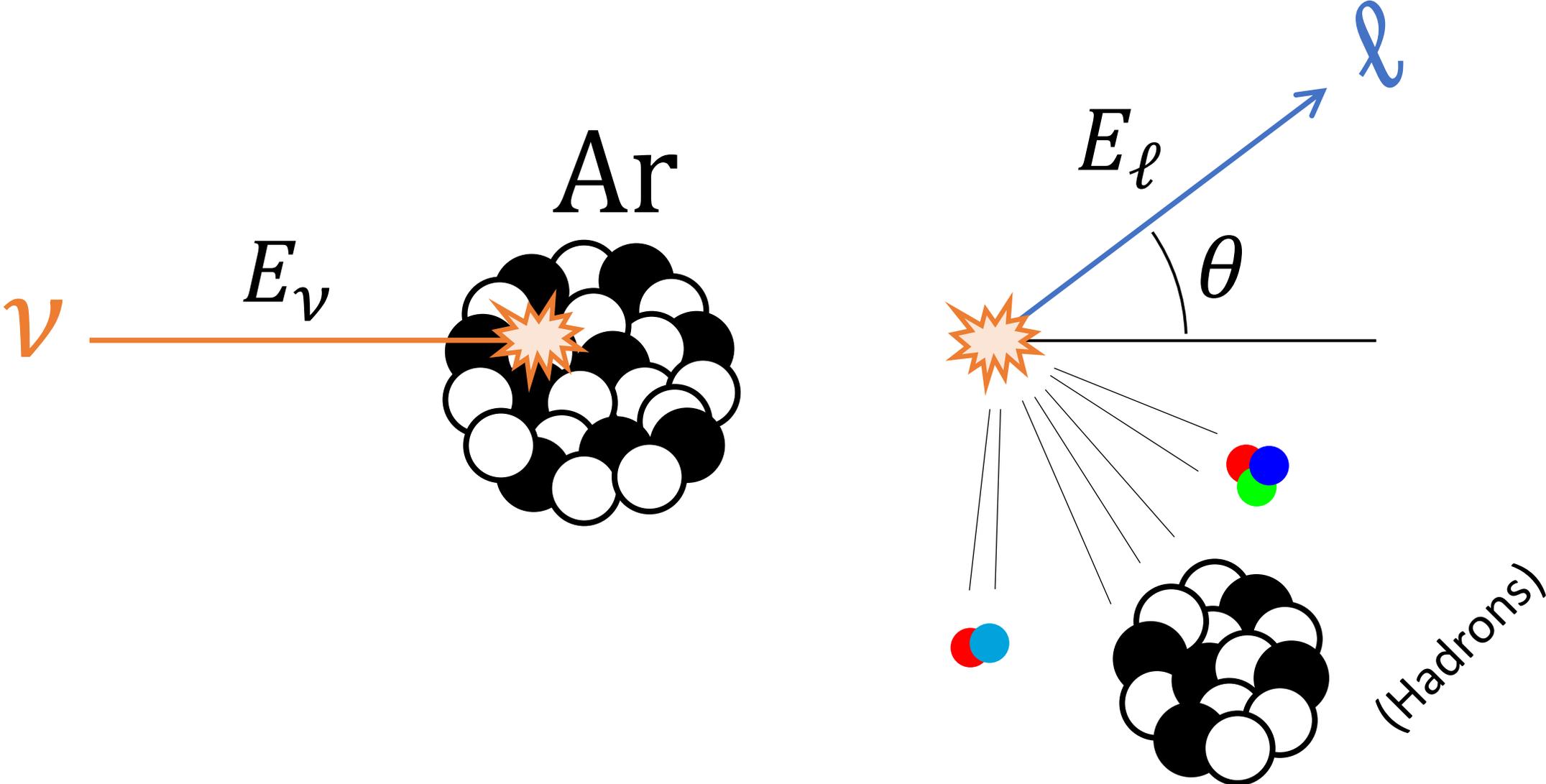

"Toy model" for DUNE physics

- *Known* cross section, ND flux, oscillation parameters
- Analytically tractable
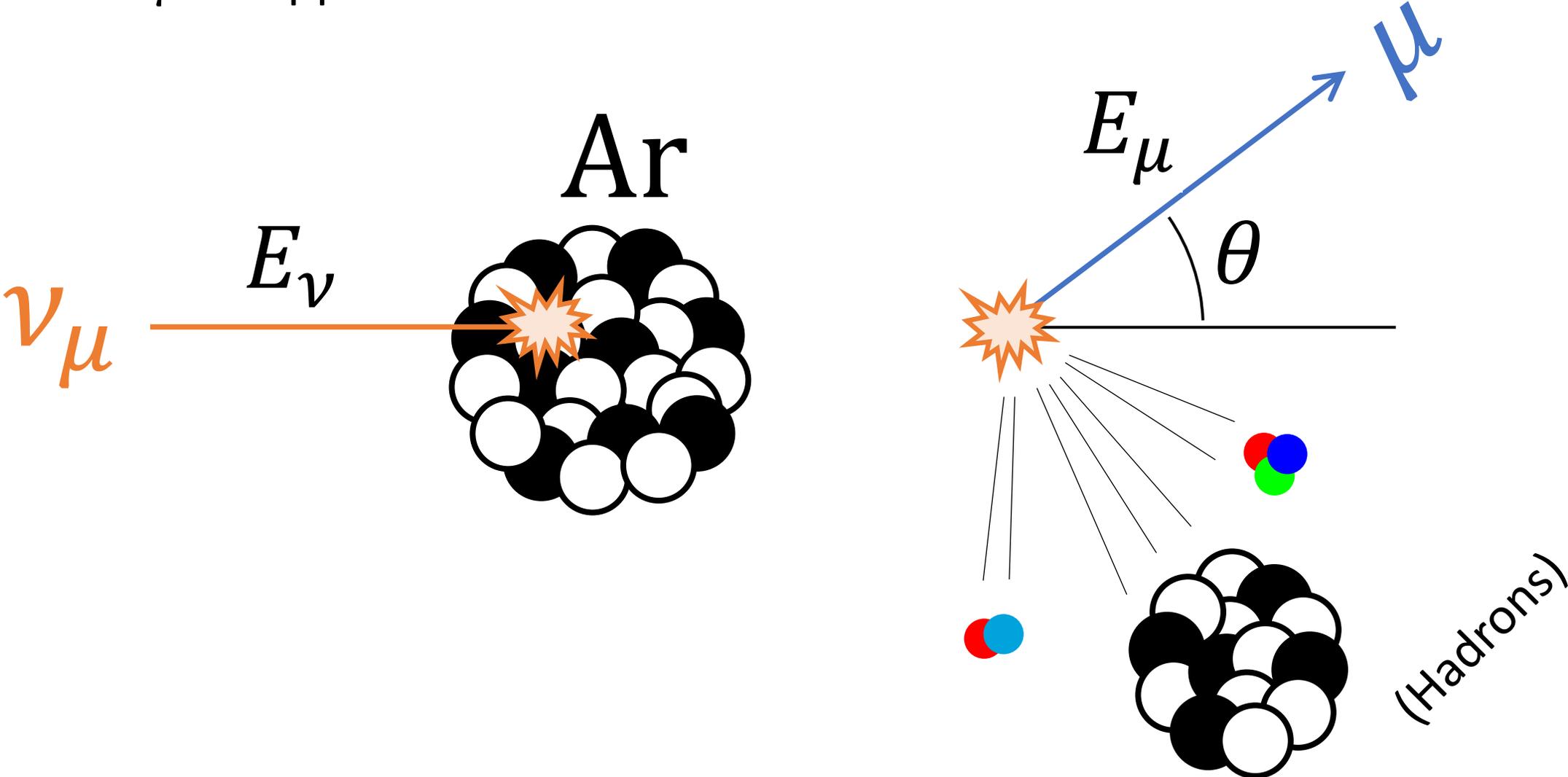- Sampleable

(Don't have this much info in reality)

Closure test(?)

Before applying a method to an unknown system, first check if it works on a known system

Generate fake data, forget we already know everything

"Toy model" for DUNE physics

- *Known* cross section, ND flux, oscillation parameters
- Analytically tractable
- Sampleable

(Don't have this much info in reality)

Closure test(?)

Before applying a method to an unknown system, first check if it works on a known system

"Toy model" for DUNE physics
- *Known* cross section, ND flux, oscillation parameters
- Analytically tractable
- Sampleable

(Don't have this much info in reality)

Generate fake data, forget we already know everything

Apply method

Closure test(?)

Before applying a method to an unknown system, first check if it works on a known system

Generate fake data, forget we already know everything

Apply method

"Toy model" for DUNE physics
- *Known* cross section, ND flux, oscillation parameters
- Analytically tractable
- Sampleable

(Don't have this much info in reality)

Accurate inference of
- 3D cross-section?
- oscillation parameters?

(Can check against known values from toy model!)

Neutrino-Argon scattering

Ar

ν $\xrightarrow{\quad E_\nu \quad}$

Neutrino-Argon scattering

Neutrino-Argon scattering

Consider: μ disappearance channel

Neutrino-Argon scattering

Consider: μ disappearance channel

Inclusive

Observed: 2d data

E_ν

Ar

E_μ

θ

μ

ν_μ

Hidden (marginalized)

Ignore (for now)

(Hadrons)

Full event info is 3d,
but only observe 2d!

Toy model of DUNE physics

Ingredient 1: ND flux

DUNE projection
- Linearly interpolated
- 0 outside $m_\mu \leq E_\nu \leq 10$ GeV

Toy model of DUNE physics

Ingredient 1: ND flux

ND ν_μ flux

DUNE projection
- Linearly interpolated
- 0 outside $m_\mu \leq E_\nu \leq 10$ GeV

Ingredient 2: Oscillation parameters

	Normal Ordering
	bfp $\pm 1\sigma$
$\sin^2 \theta_{12}$	$0.307^{+0.012}_{-0.011}$
$\theta_{12}/°$	$33.68^{+0.73}_{-0.70}$
$\sin^2 \theta_{23}$	$0.561^{+0.012}_{-0.015}$
$\theta_{23}/°$	$48.5^{+0.7}_{-0.9}$
$\sin^2 \theta_{13}$	$0.02195^{+0.00054}_{-0.00058}$
$\theta_{13}/°$	$8.52^{+0.11}_{-0.11}$
$\delta_{\mathrm{CP}}/°$	177^{+19}_{-20}
$\dfrac{\Delta m^2_{21}}{10^{-5}~\mathrm{eV}^2}$	$7.49^{+0.19}_{-0.19}$
$\dfrac{\Delta m^2_{3\ell}}{10^{-3}~\mathrm{eV}^2}$	$+2.534^{+0.025}_{-0.023}$

IC19 without SK atmospheric data

Take CV of **NuFit-6.0** NO best fit as "true" oscillation parameters
[2410.05380]

Toy model of DUNE physics

Ingredient 1: ND flux

DUNE projection
- Linearly interpolated
- 0 outside $m_\mu \leq E_\nu \leq 10$ GeV

Ingredient 2: Oscillation parameters

Take CV of **NuFit-6.0** NO best fit as "true" oscillation parameters
[2410.05380]

$$\Phi_{\text{FD}}(E_\nu) = \Phi_{\text{ND}}(E_\nu)\, P_{\mu\mu}(E_\nu)$$

[NuFast 2405.02400]

Toy model of DUNE physics

Ingredient 3: Cross section

$$\frac{d^2\sigma}{dE_\ell d\cos\theta} = \sum_i K_i \, W_i$$

$$W_2 = \frac{4x^2 M_A^2}{AQ^2}(\bar{u} + d + \bar{c} + s)$$

$$W_3 = 2x(d - \bar{u} + s - \bar{c})$$

$$2xW_1 = 2xW_5 = \frac{AQ^2}{2M_A^2} W_2$$

$$W_4 = 0$$

$\bar{u}, d, \bar{c}, s \sim$ CT18NNLO PDFs

[1912.10053]

Upshot:

DIS/pQCD at all q^2

Ar $\sim$ proton at LO in quark-parton model

Toy model of DUNE physics

Ingredient 3: Cross section

$$\frac{d^2\sigma}{dE_\ell d\cos\theta} = \sum_i K_i \, W_i$$

$$W_2 = \frac{4x^2 M_A^2}{AQ^2}(\bar{u} + d + \bar{c} + s)$$

$$W_3 = 2x(d - \bar{u} + s - \bar{c})$$

$$2xW_1 = 2xW_5 = \frac{AQ^2}{2M_A^2} W_2$$

$$W_4 = 0$$

$\bar{u}, d, \bar{c}, s \sim$ CT18NNLO PDFs

[1912.10053]

Upshot:

DIS/pQCD at all q^2

Ar $\sim$ proton at LO in quark-parton model

Cross section $\times$ flux $=$ event distribution

Marginalize over $E_\nu \rightarrow$ 2d

$$P(E_\ell, \cos\theta) = \frac{\int dE_\nu \, \frac{d^2\sigma}{dE_\ell \, d\cos\theta} \, \Phi}{\int dE_\nu \, dE_\ell \, d\cos\theta \, \frac{d^2\sigma}{dE_\ell \, d\cos\theta} \, \Phi}$$

Setup: ND Inference
(i.e. learning the cross section)

Assume known

Setup: ND Inference
(i.e. learning the cross section)

Assume known

Φ_{ND}

ND flux

E_ν

$P_{\mathrm{ND}}(E_\ell, \cos\theta) \propto \int dE_\nu \, \dfrac{d^2\sigma}{dE_\ell \, d\cos\theta} \, \Phi_{\mathrm{ND}}$

True ND

$\sqrt{E_\ell \, \sin^2(\theta/2)}$ [GeV$^{1/2}$]

$\sqrt{E_\ell}$ [GeV$^{1/2}$]

Setup: ND Inference
(i.e. learning the cross section)

Assume ~ infinite ND stats
→ 2d ND event distribution known

Assume known

ND flux

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

True ND

Training data!

Setup: FD Inference
(i.e. oscillation analysis)

Must infer (params of)!

FD flux

Φ_{FD}

E_ν

m_μ

Setup: FD Inference
(i.e. oscillation analysis)

Must infer (params of)!

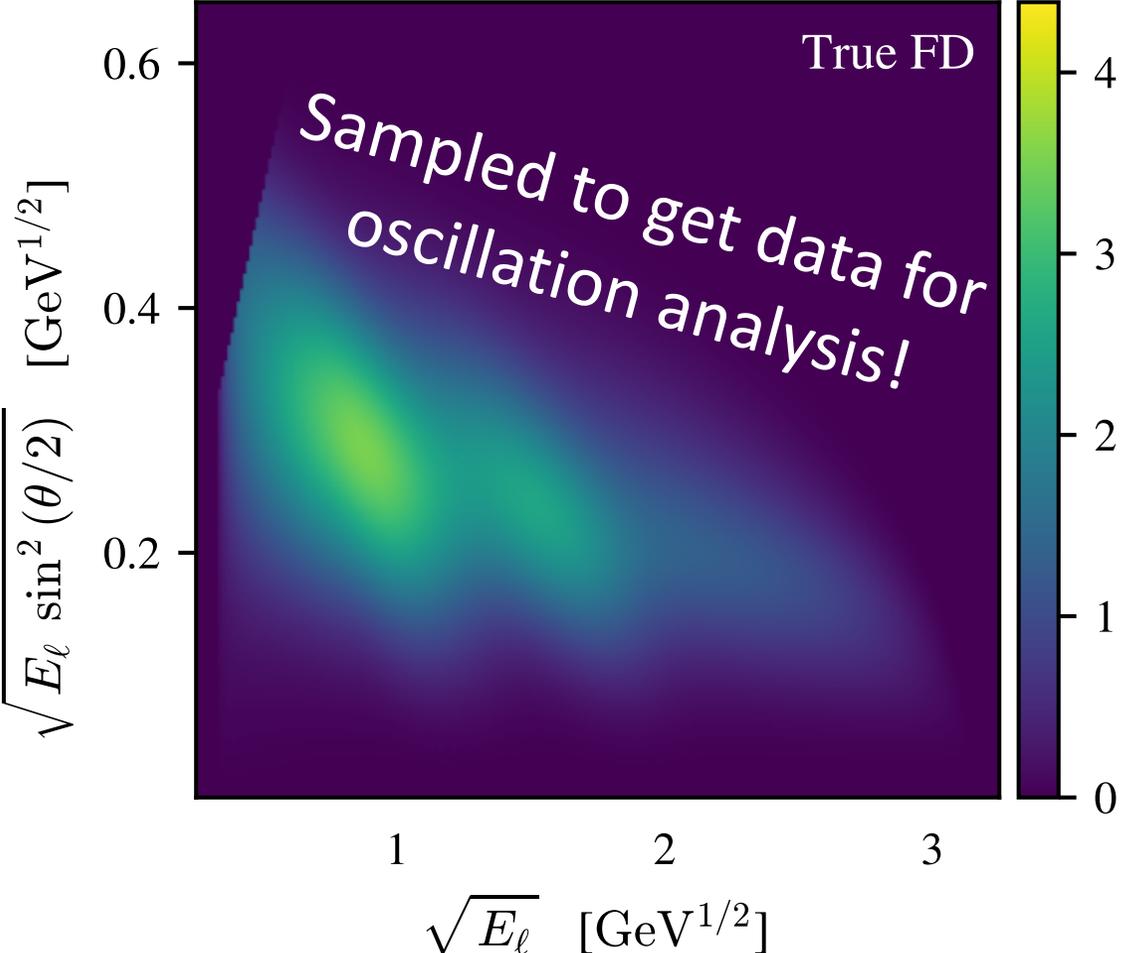$$P_{\mathrm{FD}}(E_\ell, \cos\theta) \propto \int dE_\nu \, \frac{d^2\sigma}{dE_\ell \, d\cos\theta} \, \Phi_{\mathrm{FD}}$$

Setup: FD Inference
(i.e. oscillation analysis)

Must infer (params of)!

Take 6200 events from FD marginal
(~ 3.5y of DUNE)

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

Complication: SF ambiguities

μ disappearance only $\rightarrow$ Only one m_ℓ

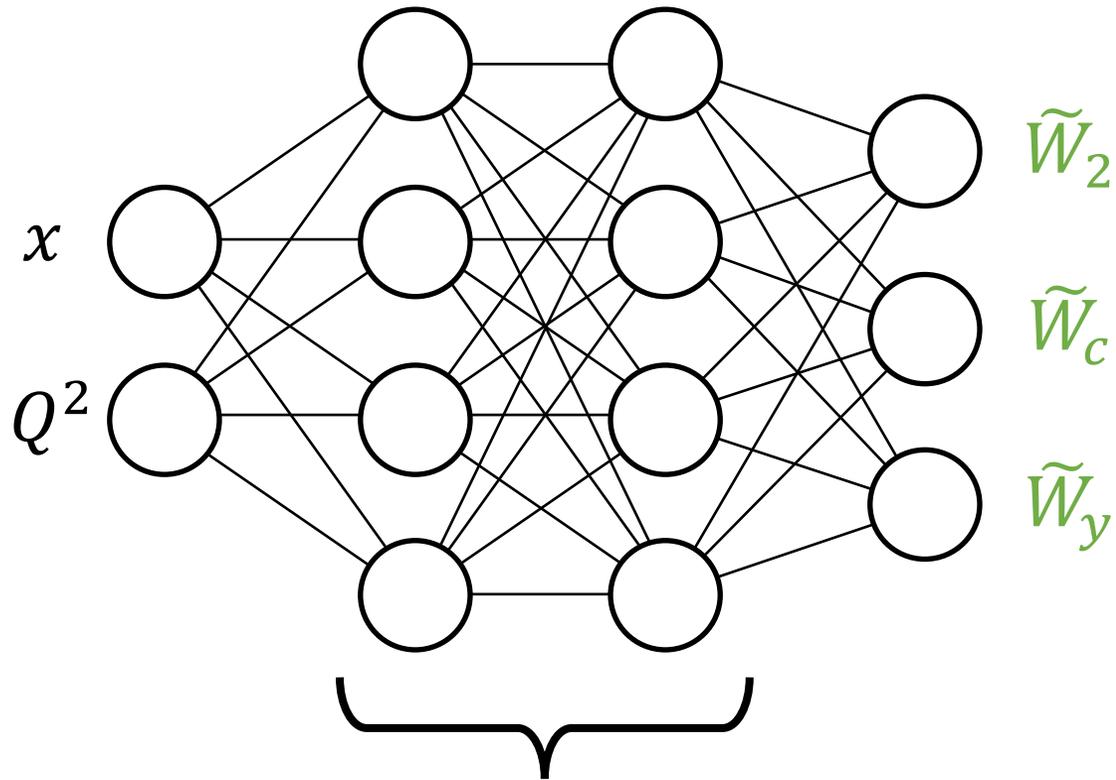$+$ Can always redefine Ws up to factors of x, Q^2

$\Rightarrow$ Only sensitive to 3 independent linear combinations of $W_1, \ldots, W_5$

$$\frac{d^2\sigma}{dE_\ell d\cos\theta}(E_\nu) = \frac{|V_{ud}|^2 G_F^2}{\pi}\sqrt{E_\ell^2 - m_\ell^2}\Bigg\{$$

$$\frac{E_\nu}{M_A}W_2(x, Q^2) + W_c(x, Q^2) + \tilde{y}W_y(x, Q^2)\Bigg\}$$
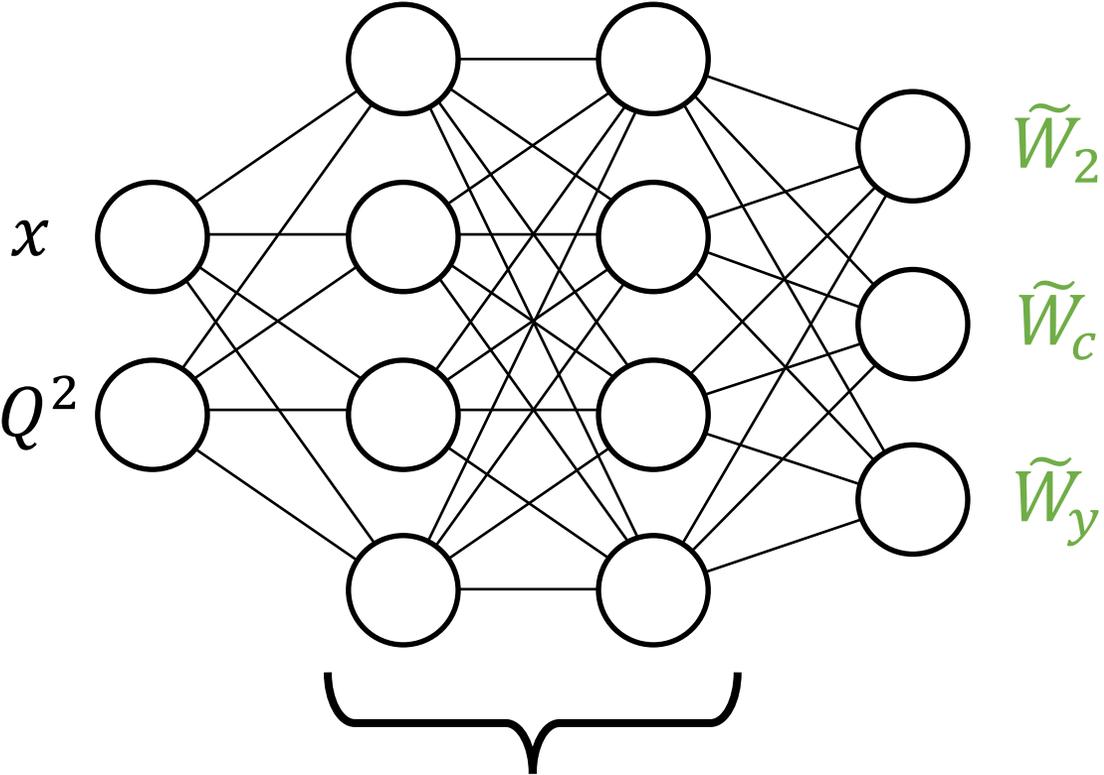
ML Setup

Architecture:

Parametrize SFs via an MLP

x

Q^2

$\widetilde{W}_2$

$\widetilde{W}_c$

$\widetilde{W}_y$

(Actually: 4 hidden layers of width 64)

ML Setup

Architecture:
Parametrize SFs via an MLP

x

Q^2

$\widetilde{W}_2$

$\widetilde{W}_c$

$\widetilde{W}_y$

(Actually: 4 hidden layers of width 64)

Training:
Model SFs $\times$ known K_i

$\rightarrow$ Model cross section

$$\widetilde{\frac{d^2\sigma}{dE_\mu \; d\cos\theta}} = \sum_i K_i \widetilde{W}_i$$

$\times$ known Φ_{ND}, marginalize E_ν

$\rightarrow$ Model ND event distribution

$$Q_{\mathrm{ND}} \propto \int dE_\nu \; \widetilde{\frac{d^2\sigma}{dE_\mu \; d\cos\theta}} \; \Phi_{\mathrm{ND}}$$

Tune so that $Q_{\mathrm{ND}} \approx P_{\mathrm{ND}}$ as closely as possible. MSE loss:

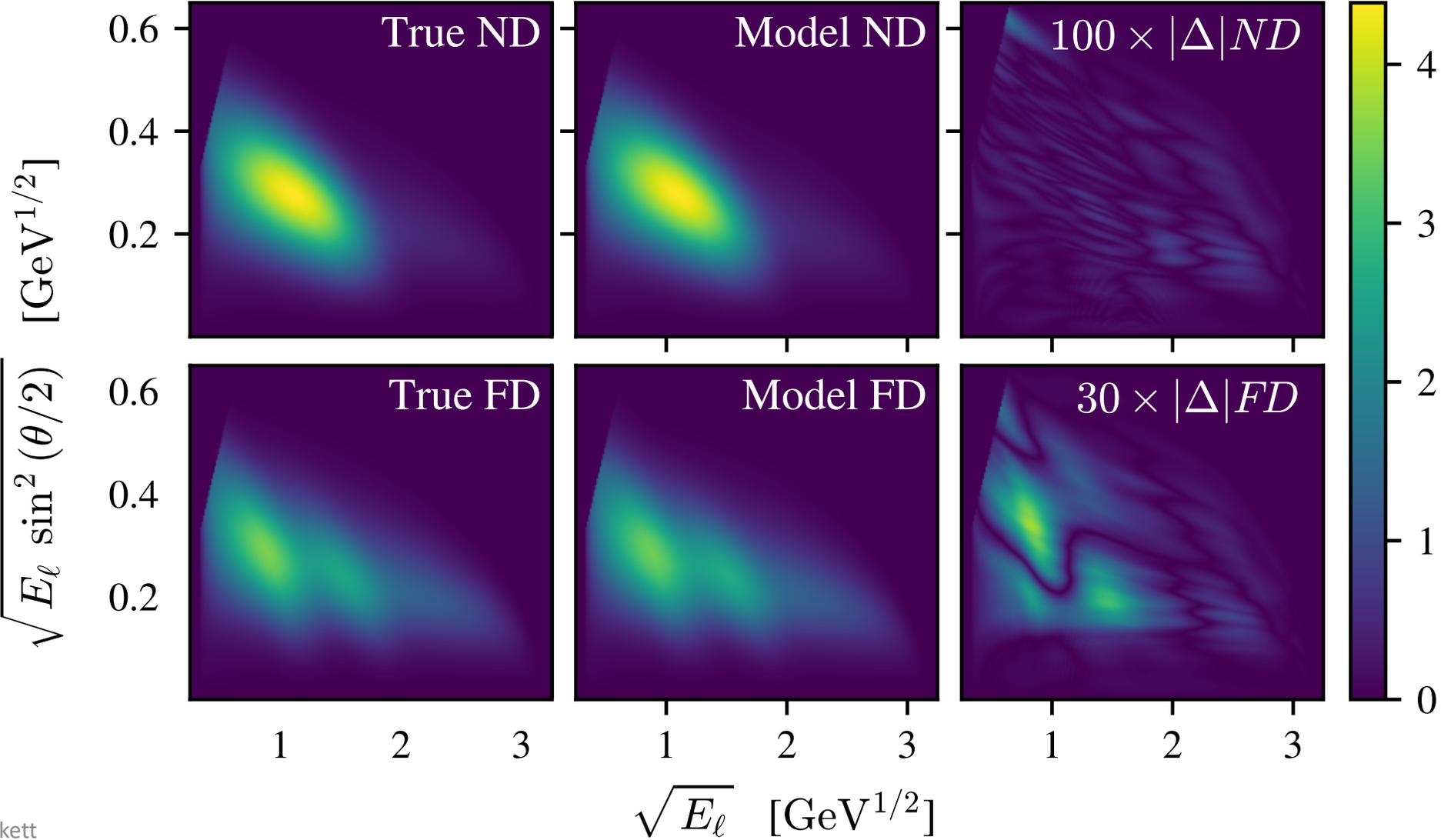$$\mathcal{L} = \int dE_\mu d\cos\theta [P_{\mathrm{ND}} - Q_{\mathrm{ND}}]^2$$

Results: event distributions

Trained model not only approximates ND event distribution (expected), also generalizes to FD!
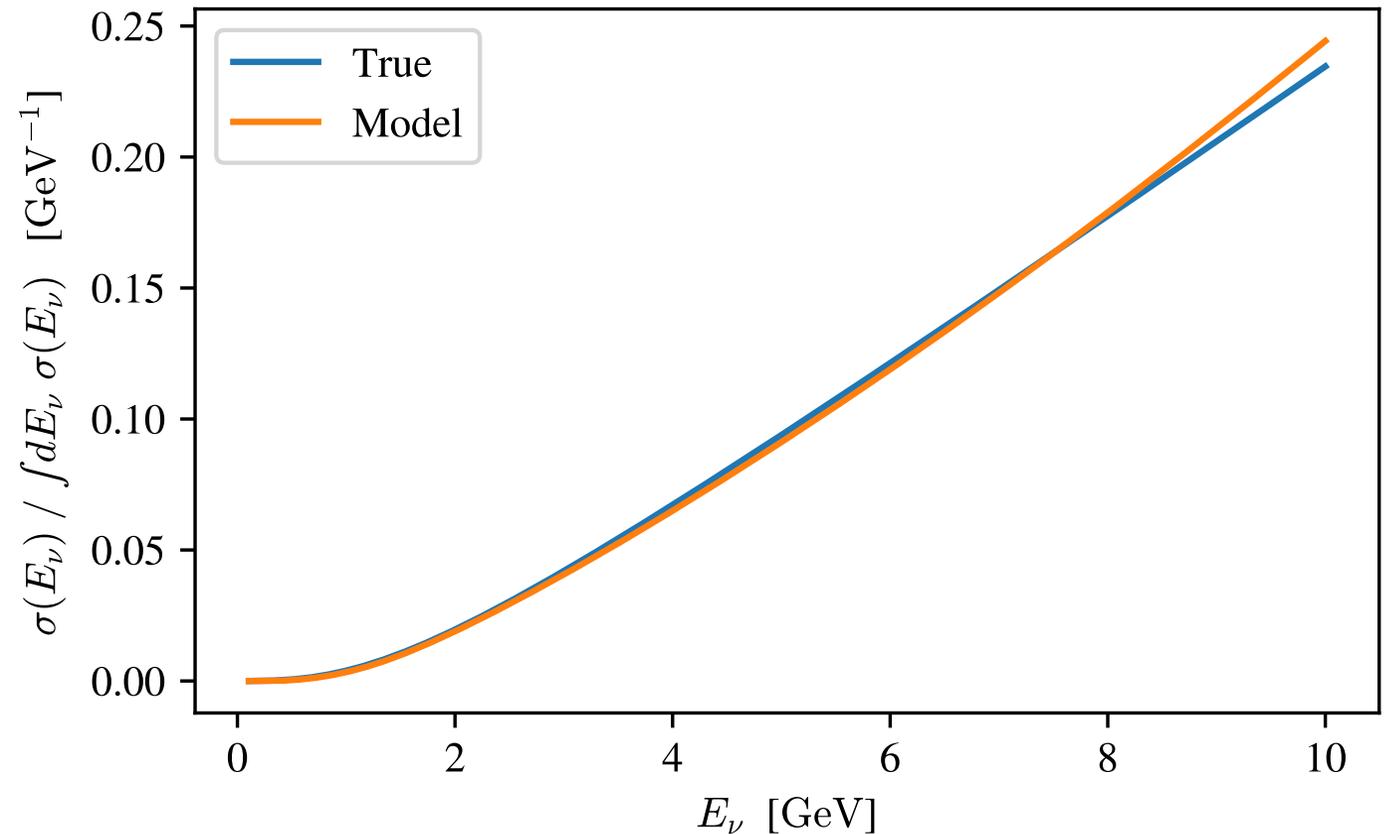
Results: event distributions

Trained model not only approximates ND event distribution (expected), also generalizes to FD!

Results: cross section

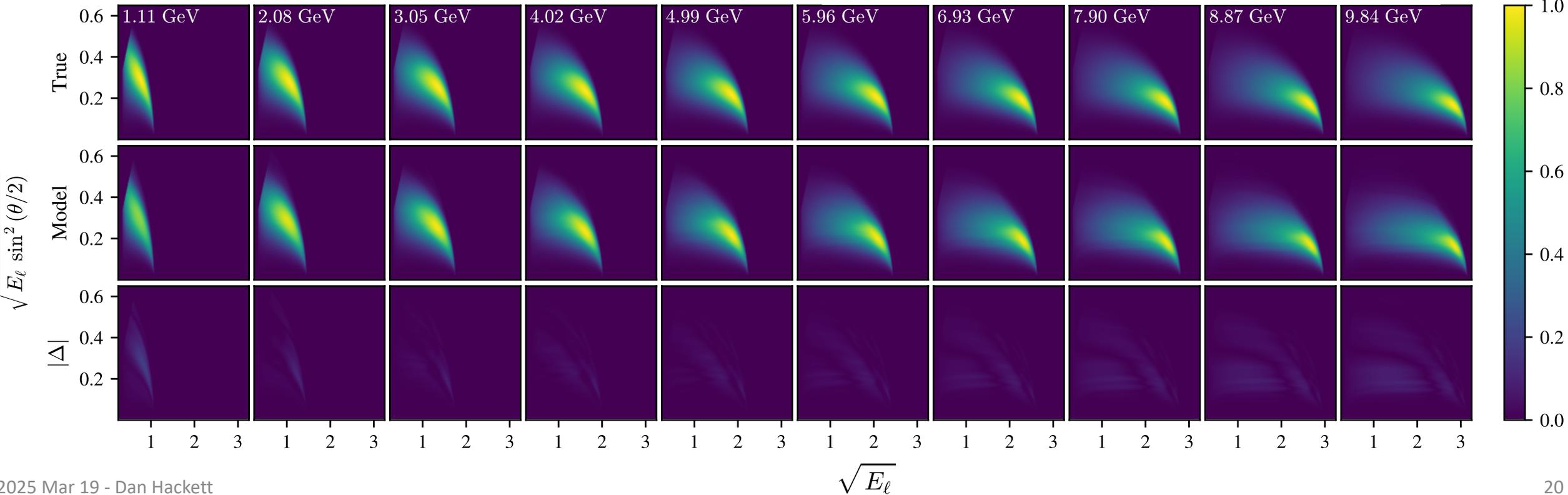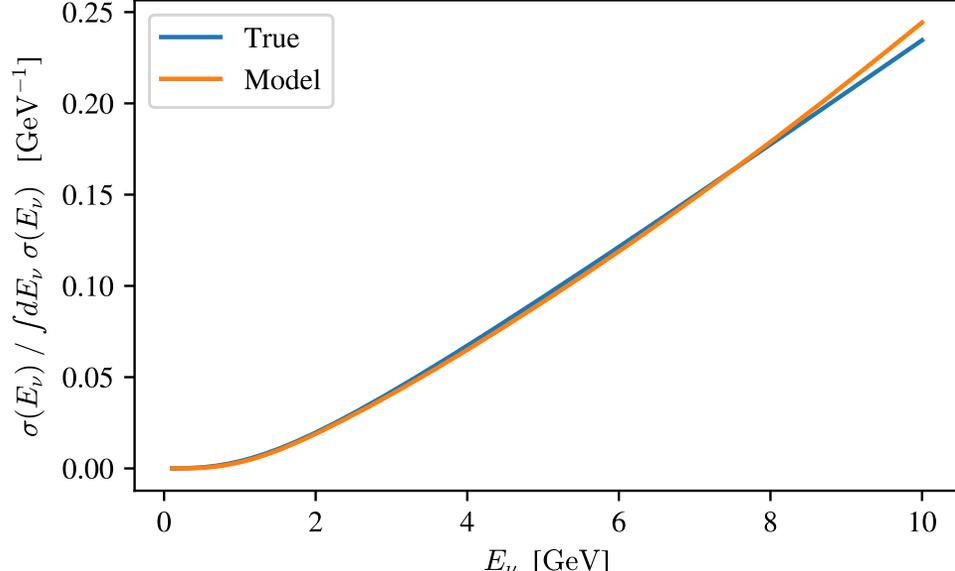FD generalization reflects good modeling of full 3d cross section
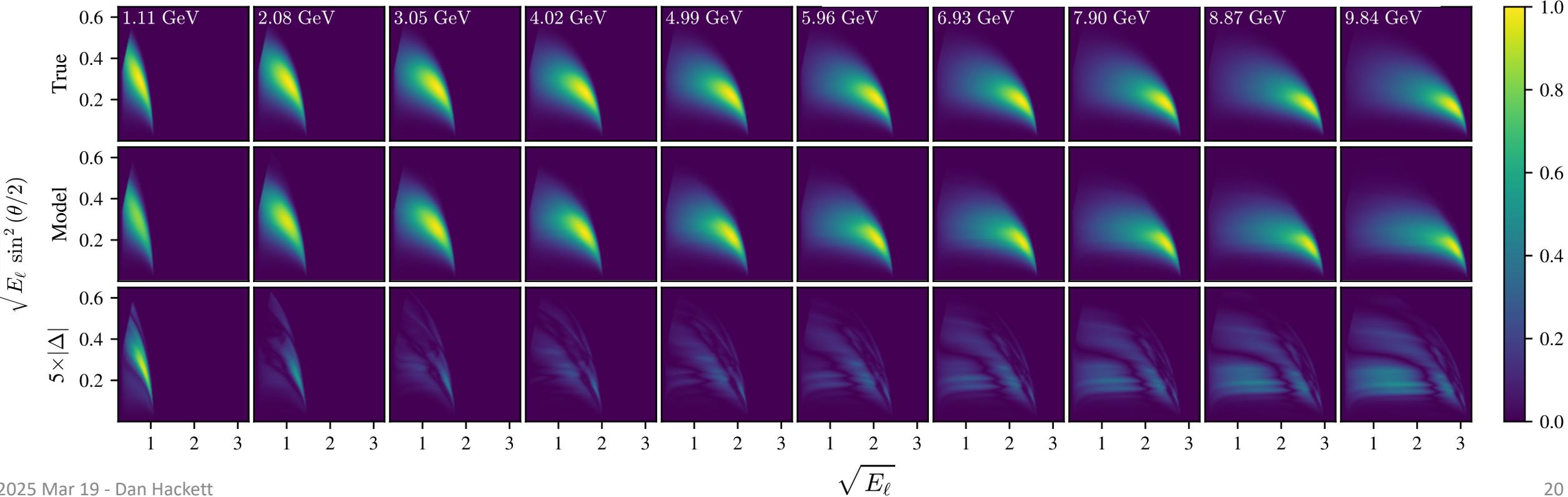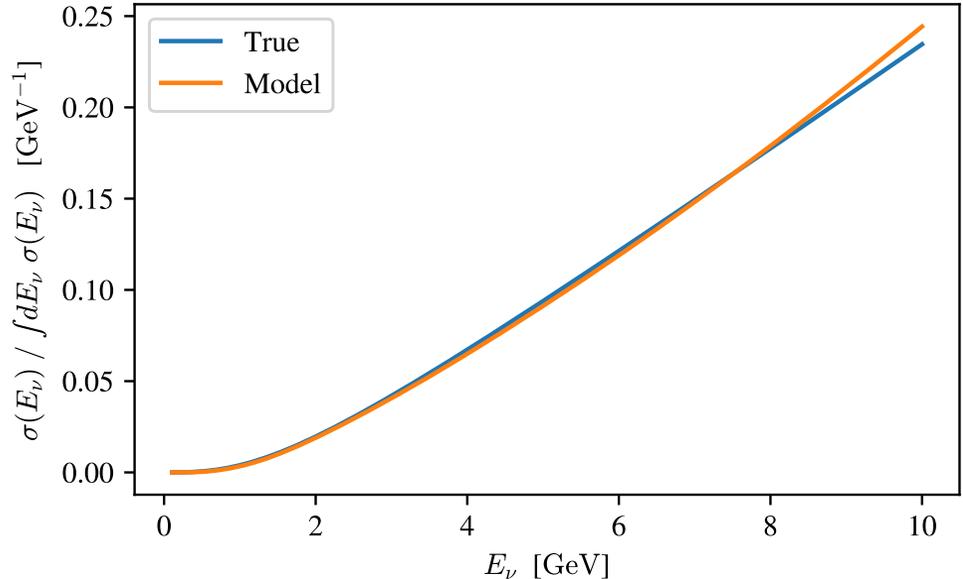
...even though only trained on 2d ND marginal!

Results: cross section

FD generalization reflects good modeling of full 3d cross section

...even though only trained on 2d ND marginal!

Results: cross section

FD generalization reflects good
modeling of full 3d cross section

...even though only trained on
2d ND marginal!
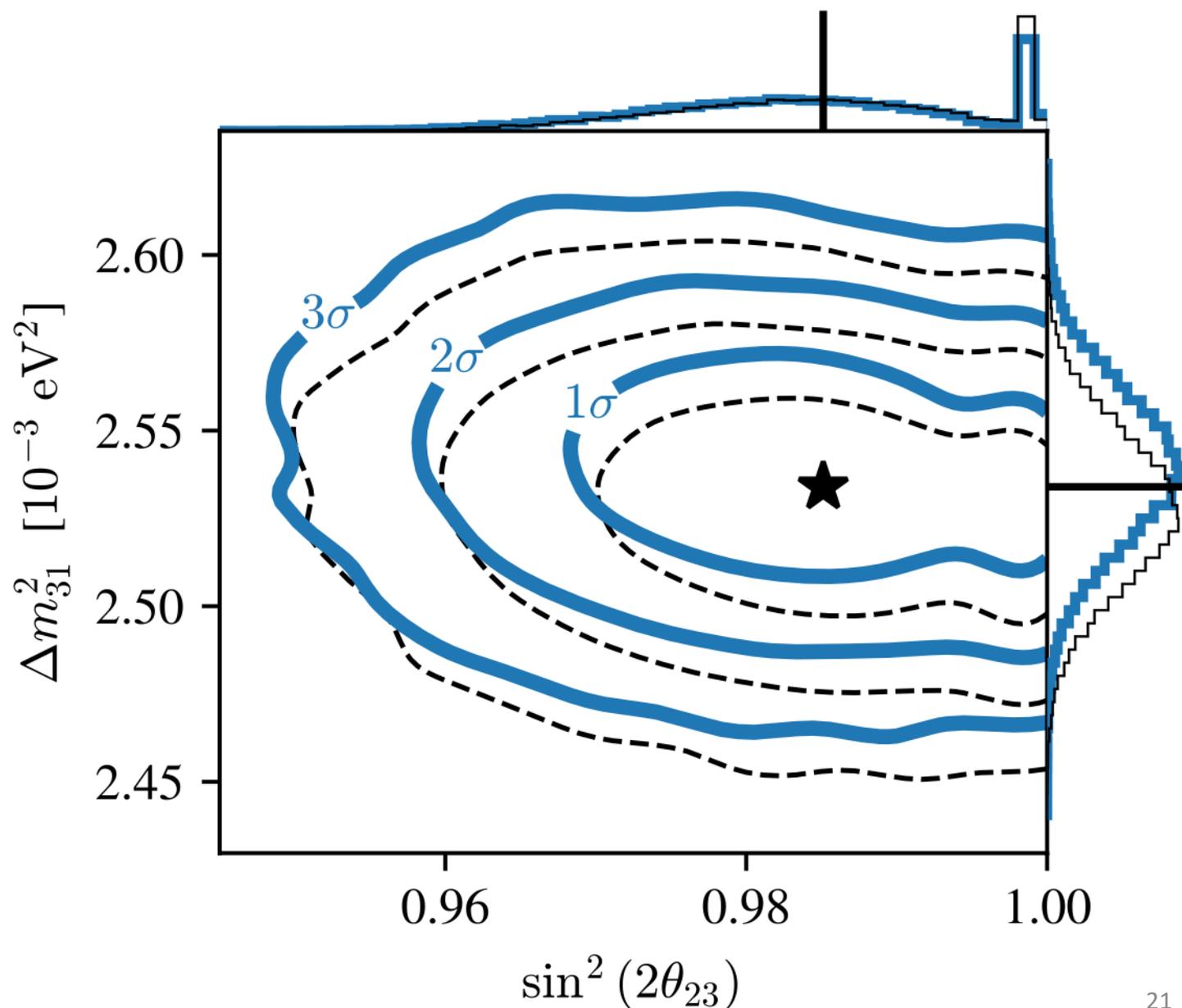
Results: oscillation parameters

Model FD event distribution:

$$\widetilde{\Phi}_{\text{FD}}(E_\nu; \omega) = \Phi_{\text{ND}}(E_\nu)\, \tilde{p}_{\mu\mu}(E_\nu; \omega)$$

$$\tilde{P}_{\text{FD}} \propto \int dE_\nu \frac{d^2\sigma}{dE_\mu\, d\cos\theta}\, \widetilde{\Phi}_{\text{FD}}$$
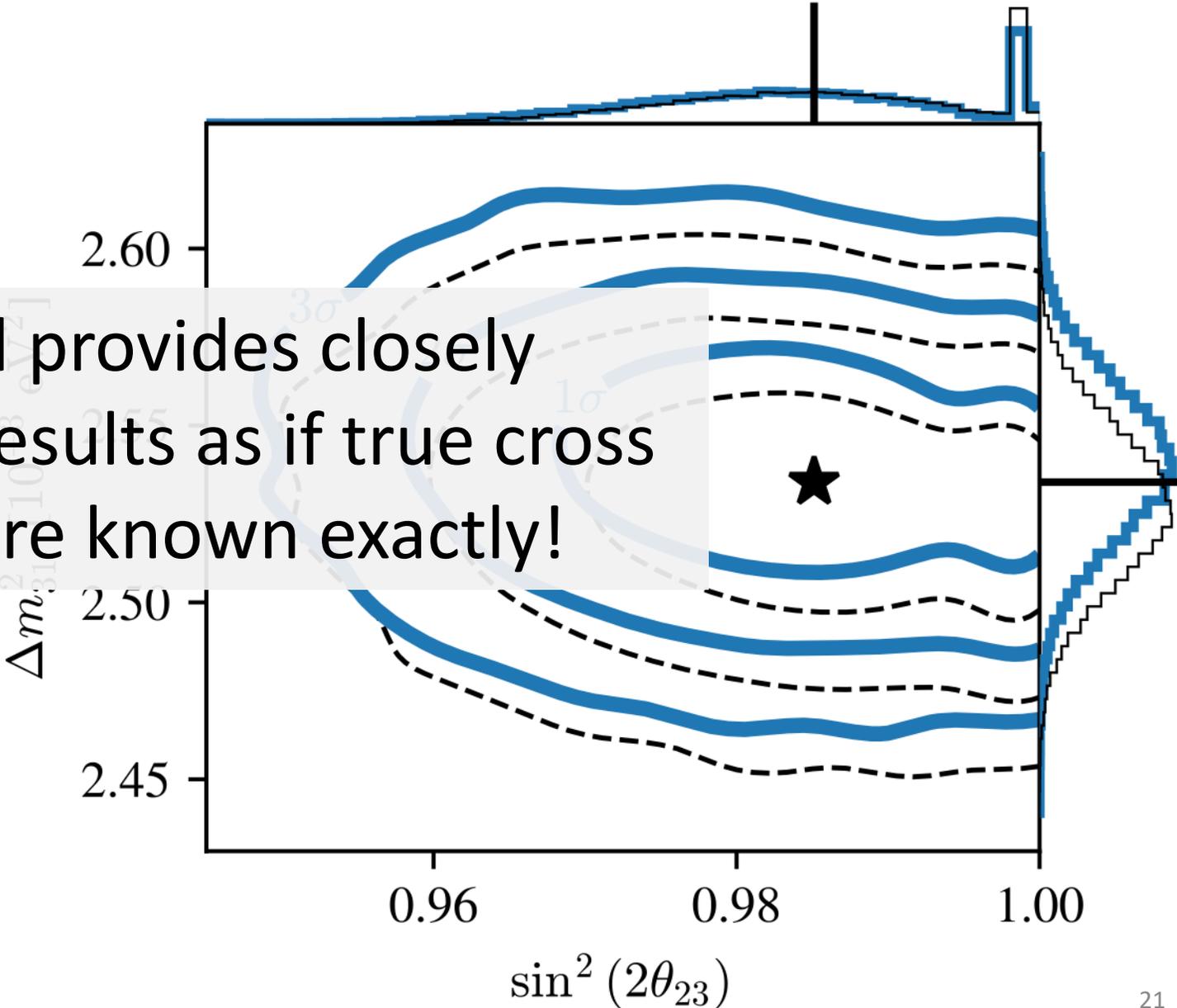
Maximum likelihood inference

$$\mathcal{L}(\omega) = \prod_{i=1}^{6200} \tilde{P}_{\text{FD}}(E_\ell^{(i)}, \cos\theta^{(i)}; \omega)$$

$$\omega^* = \max_{\omega} \mathcal{L}(\omega)$$

Bootstrap over FD events to construct confidence intervals

Results: oscillation parameters

Model FD event distribution:

$$\widetilde{\Phi}_{\mathrm{FD}}(E_\nu; \omega) = \Phi_{\mathrm{ND}}(E_\nu)\, \tilde{p}_{\mu\mu}(E_\nu; \omega)$$

$$\tilde{P}_{\mathrm{FD}} \propto \int dE_\nu \frac{d^2\sigma}{dE_\mu\, d\cos\theta}\, \widetilde{\Phi}_{\mathrm{FD}}$$

Maximum likelihood inference

$$\mathcal{L}(\omega) = \prod_{i=1}^{6200} \tilde{P}_{\mathrm{FD}}(E_\ell; \cos\theta; \omega)$$

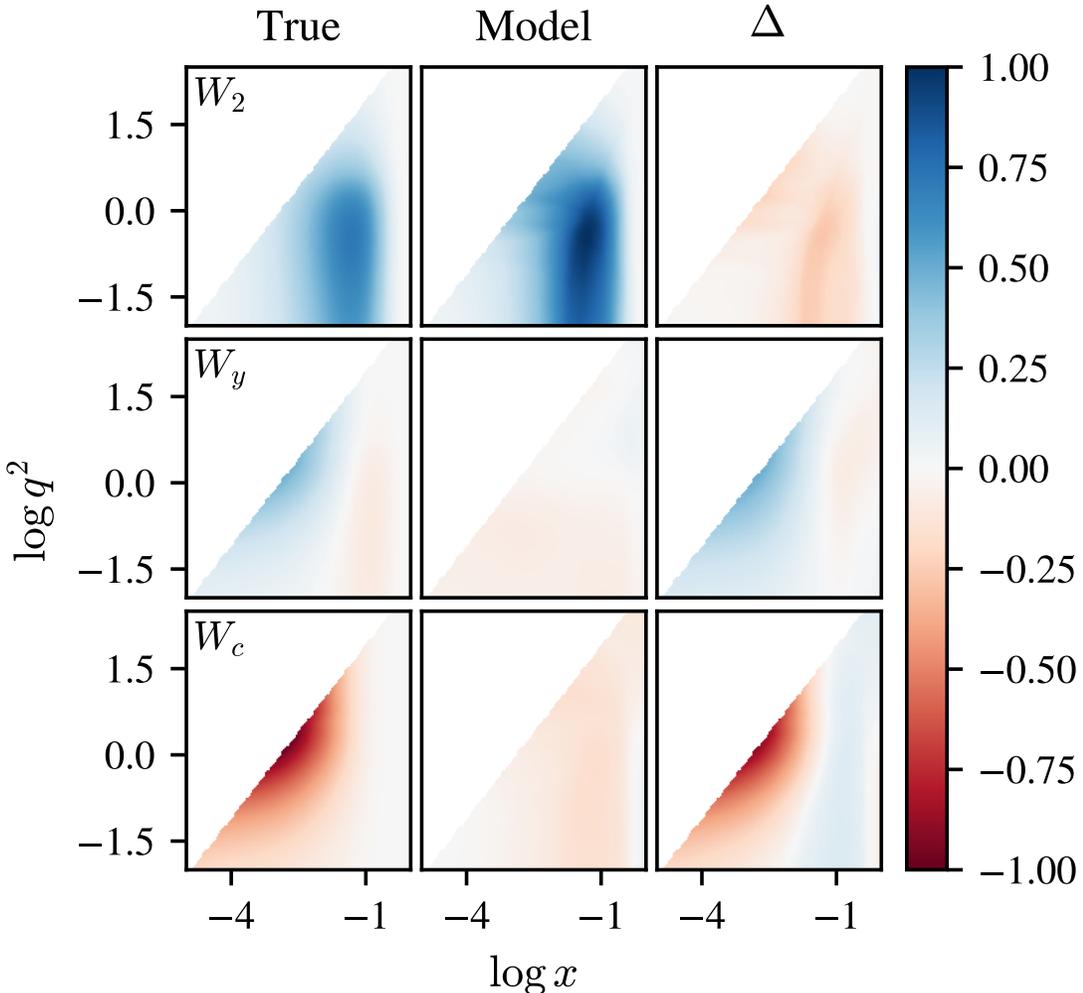$$\omega^* = \max_\omega \mathcal{L}(\omega)$$

Bootstrap over FD events to construct confidence intervals

ML model provides closely comparable results as if true cross section were known exactly!
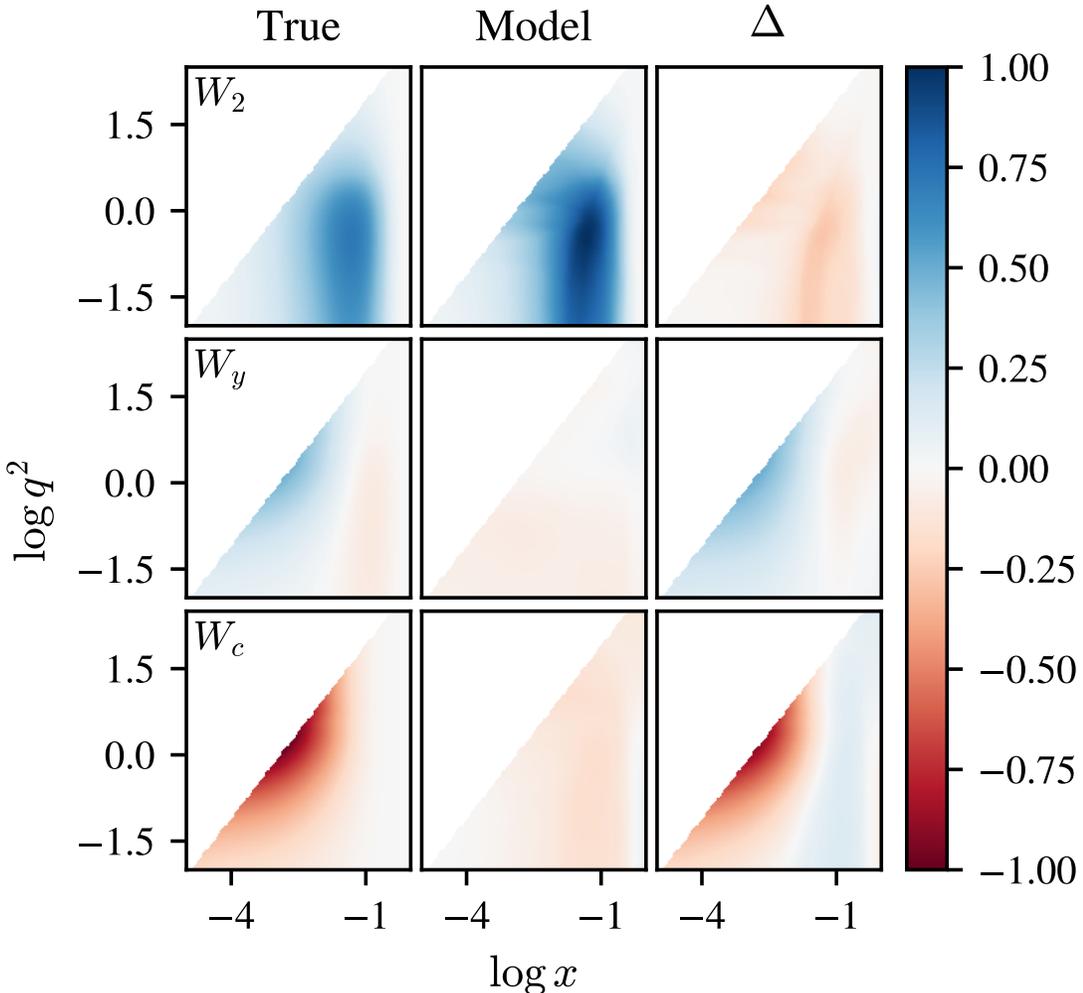
Extracting structure functions?

Complication: only constrain
some kinematic regions of SFs
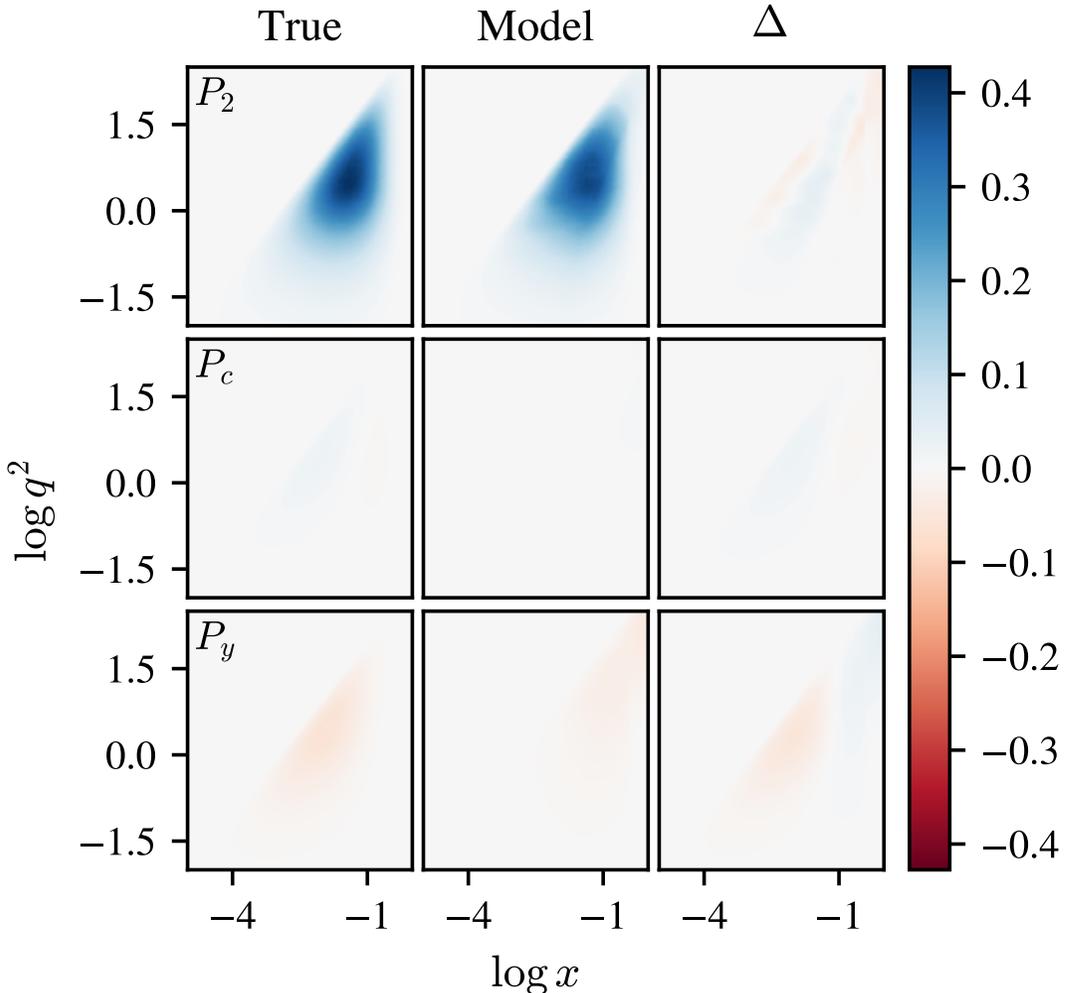
Extracting structure functions?

Complication: only constrain some kinematic regions of SFs

Weight by (local) contribution to ND event distribution in x, Q^2

$$\mathcal{K}_i(x, Q^2) \propto \int dy \, \Phi_{\mathrm{ND}} \, K_i$$

$$P_i = \mathcal{K}_i(x, Q^2) \, W_i(x, Q^2)$$

Outlook

Proof-of-concept demonstrated:

    Can learn cross section from ND data

    Fully data-driven approach to oscillation analyses is possible

    Independent complement to generator-based approaches!

Uncertainty quantification?

- ND data / flux uncertainties, detector effects, etc

- Function inference vs QFT

EIC connections?:

- Incorporate other theory/expt constraints to improve SF extraction?

- Need to extend to treat hadronic info (inclusive processes) → TMDs