

# Simultaneous Extraction of Spin-Averaged and Helicity Light Quark Sea Asymmetries

Christopher Cocuzza (Temple University)

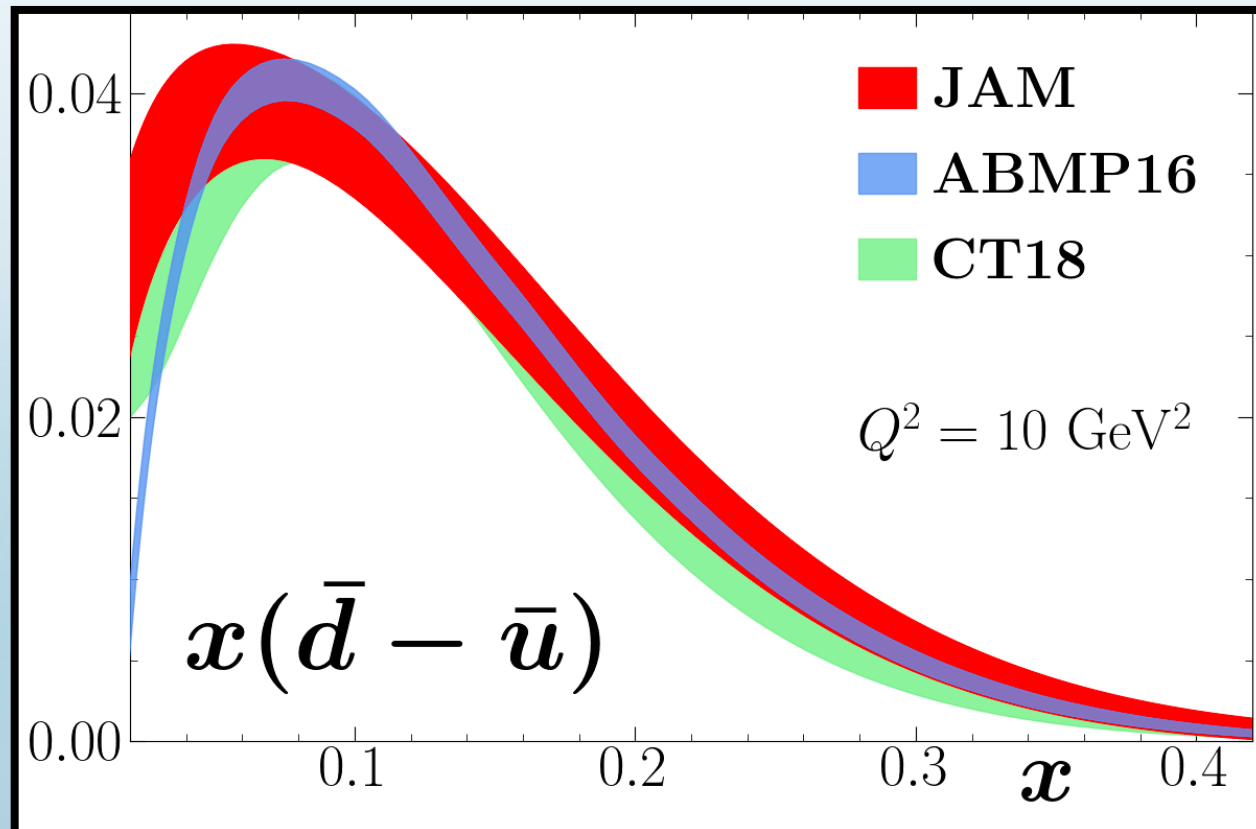
September 5, 2021



Wally Melnitchouk (Jefferson Lab)  
Andreas Metz (Temple University)  
Nobuo Sato (Jefferson Lab)



# Introduction to Sea Asymmetry



Cannot be explained from gluons  
splitting into quark-antiquark pairs

Meson Cloud Models  
Chiral Soliton Models  
Statistical Models

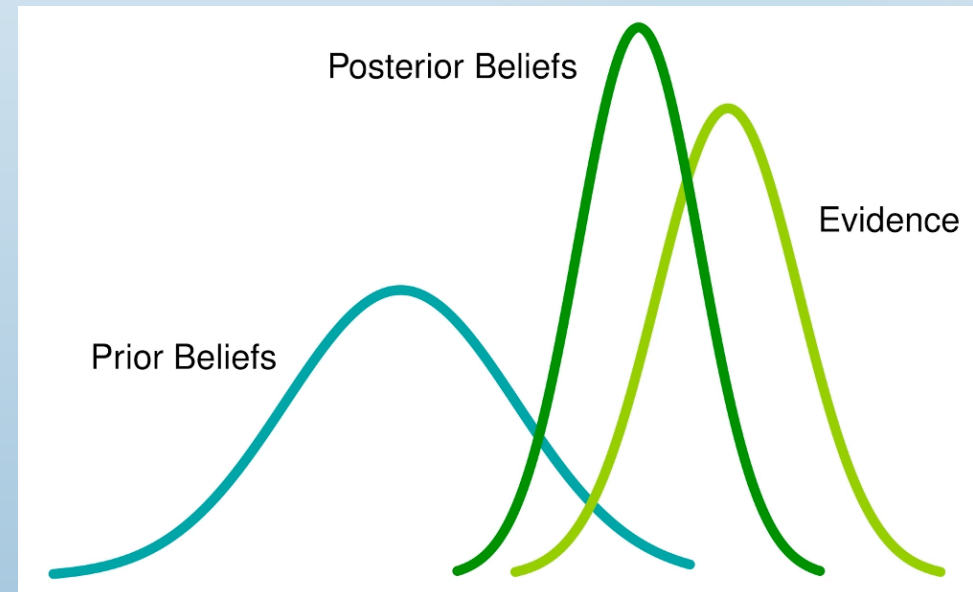
Still questions at high  $x > 0.2$  and  
for helicity asymmetry

# Part 1:

# JAM Methodology



*T. Bayes*



# JAM Collaboration

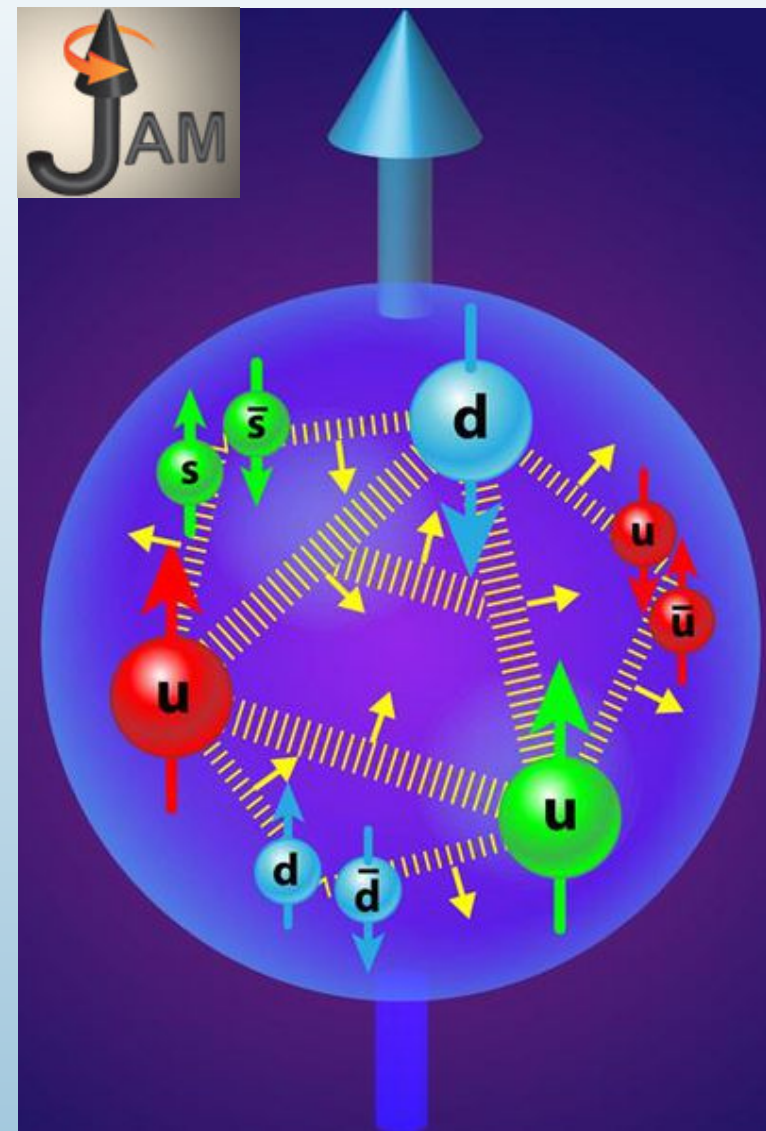
3-dimensional structure of nucleons:

- Parton distribution functions (PDFs)
- Fragmentation functions (FFs)
- Transverse momentum dependent (TMD) distributions
- Generalized parton distributions (GPDs)

Collinear factorization in perturbative QCD

Simultaneous determinations of PDFs, FFs, etc.

Monte Carlo methods for Bayesian inference





# Parameters to Observables

Parameterize PDFs at input scale  $Q_0^2 = m_c^2$

$$f_i(x) = Nx^\alpha(1-x)^\beta(1+\gamma\sqrt{x}+\eta x)$$

Evolve PDFs using DGLAP

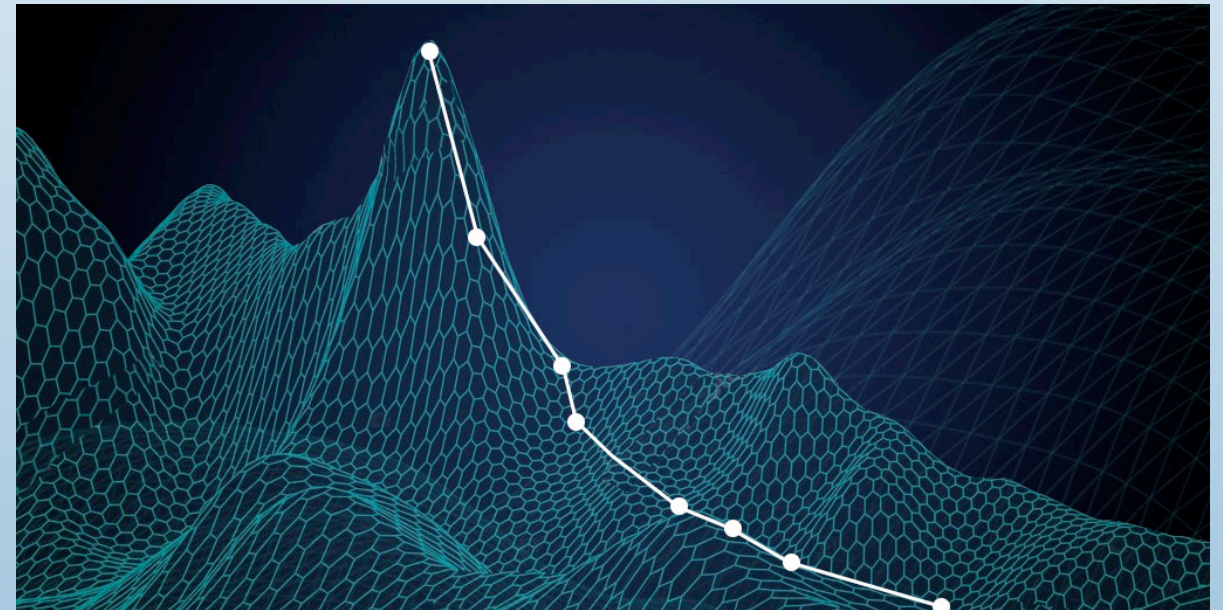
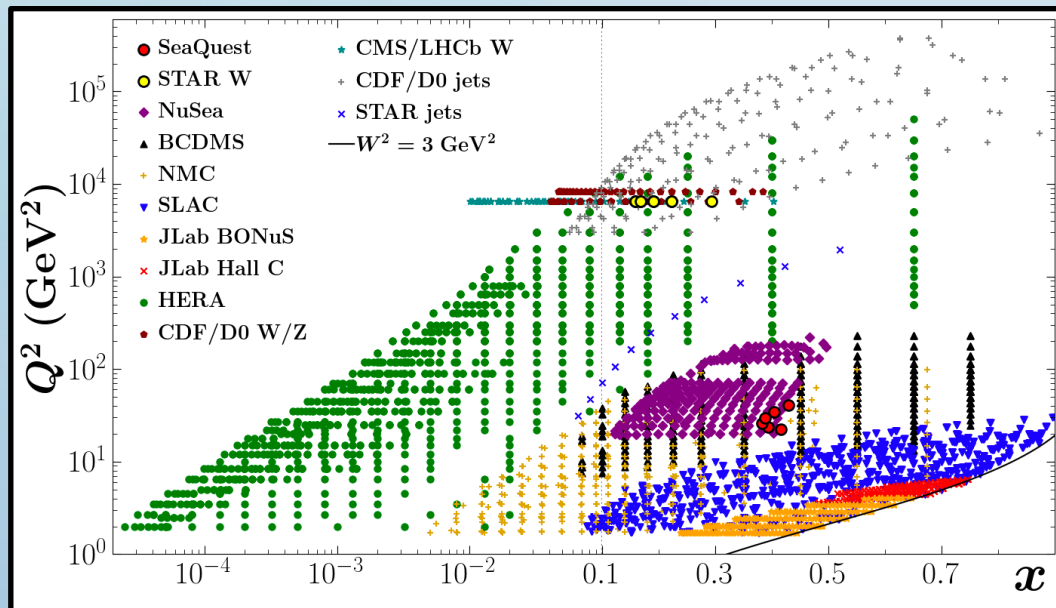
$$\frac{d}{d \ln(\mu^2)} f_i(x, \mu) = \sum_j \int_x^1 \frac{dz}{z} P_{ij}(z, \mu) f_j\left(\frac{x}{z}, \mu\right)$$

Calculate Observables

$$d\sigma_{\text{DY}} = \sum_{i,j} H_{ij}^{\text{DY}} \otimes f_i \otimes f_j$$

# Part 2:

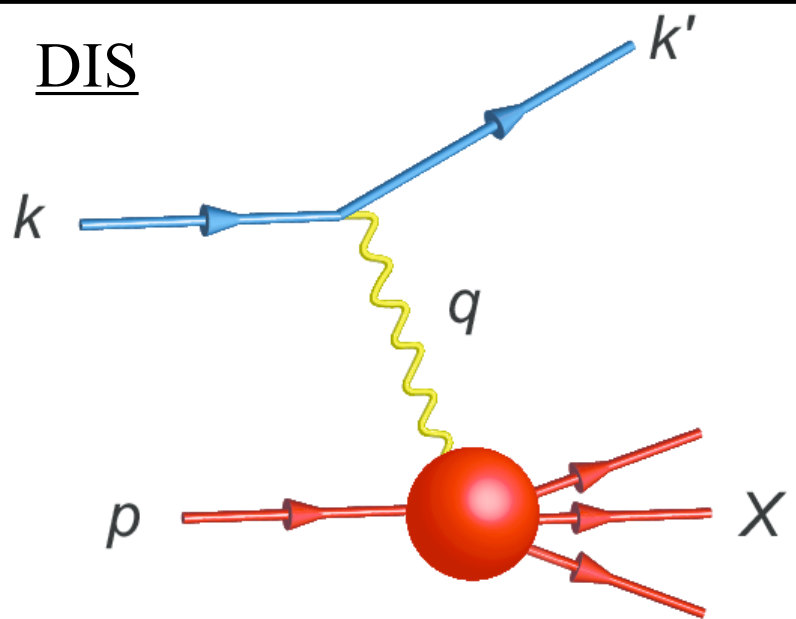
# Data and Fitting



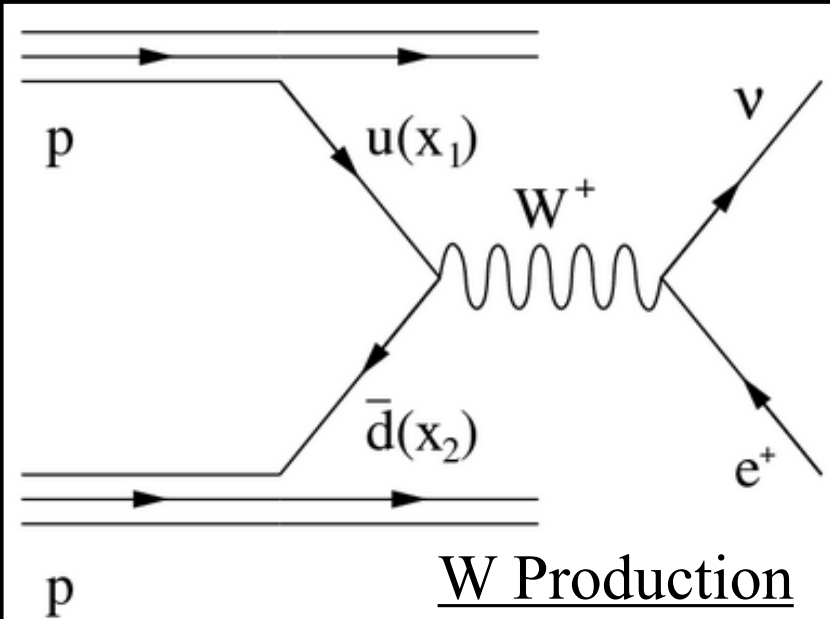
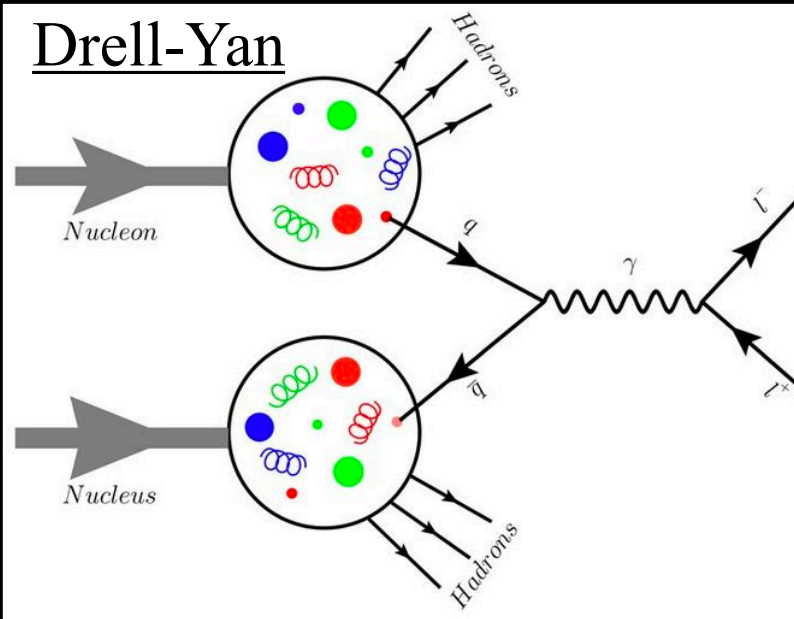
# A Global Analysis

*Simultaneous extraction of  
spin-averaged and helicity PDFs*

DIS



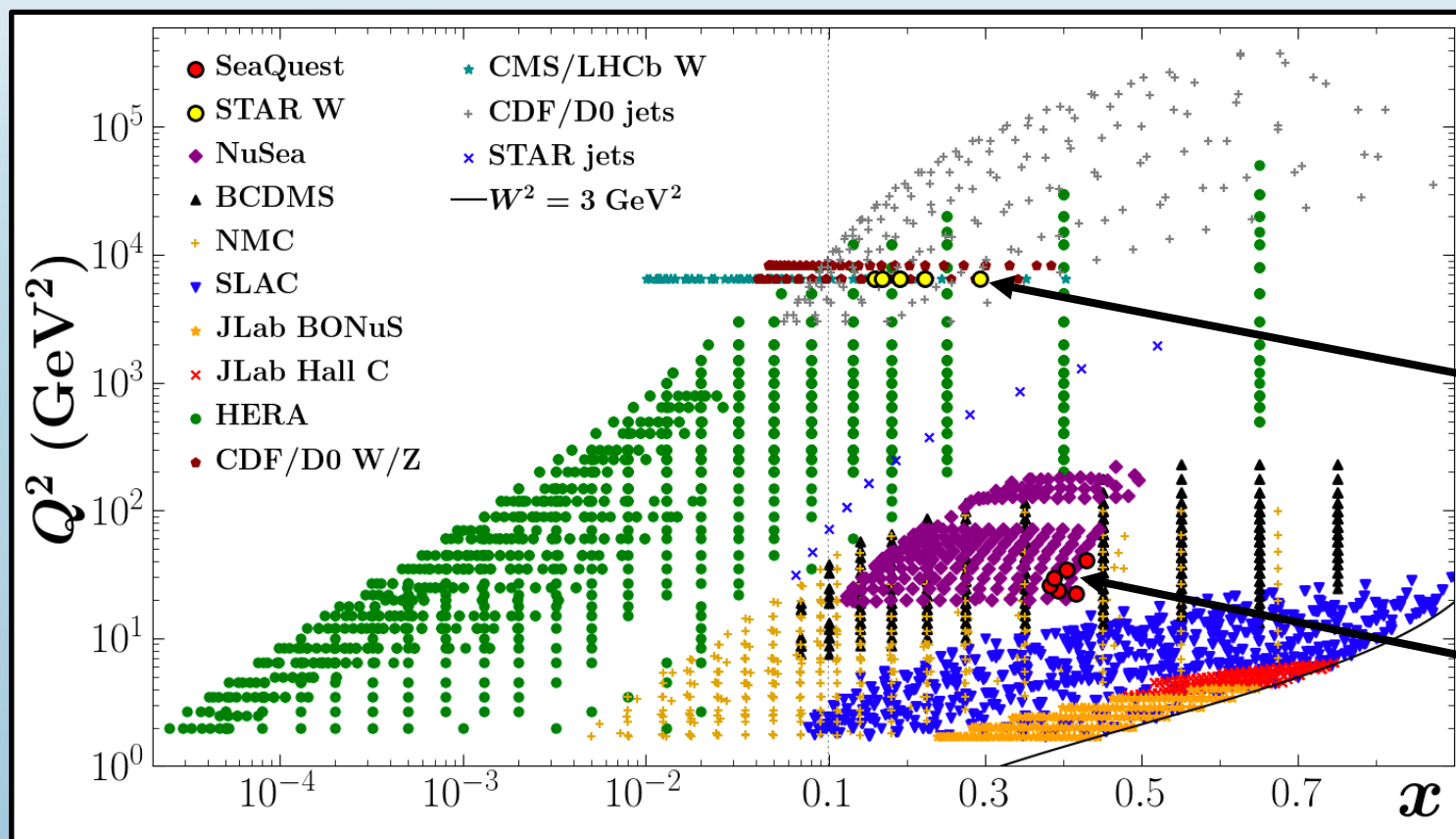
Drell-Yan



W Production

# Kinematic Coverage (Spin-Averaged)

<b>Deep Inelastic Scattering</b>	BCDMS, NMC, SLAC, HERA, Jefferson Lab	3863 points
<b>Drell-Yan</b>	Fermilab E866, E906	205 points
<b>W/Z Boson Production</b>	CDF/D0, STAR, LHCb, CMS	153 points
<b>Jets</b>	CDF/D0, STAR	200 points



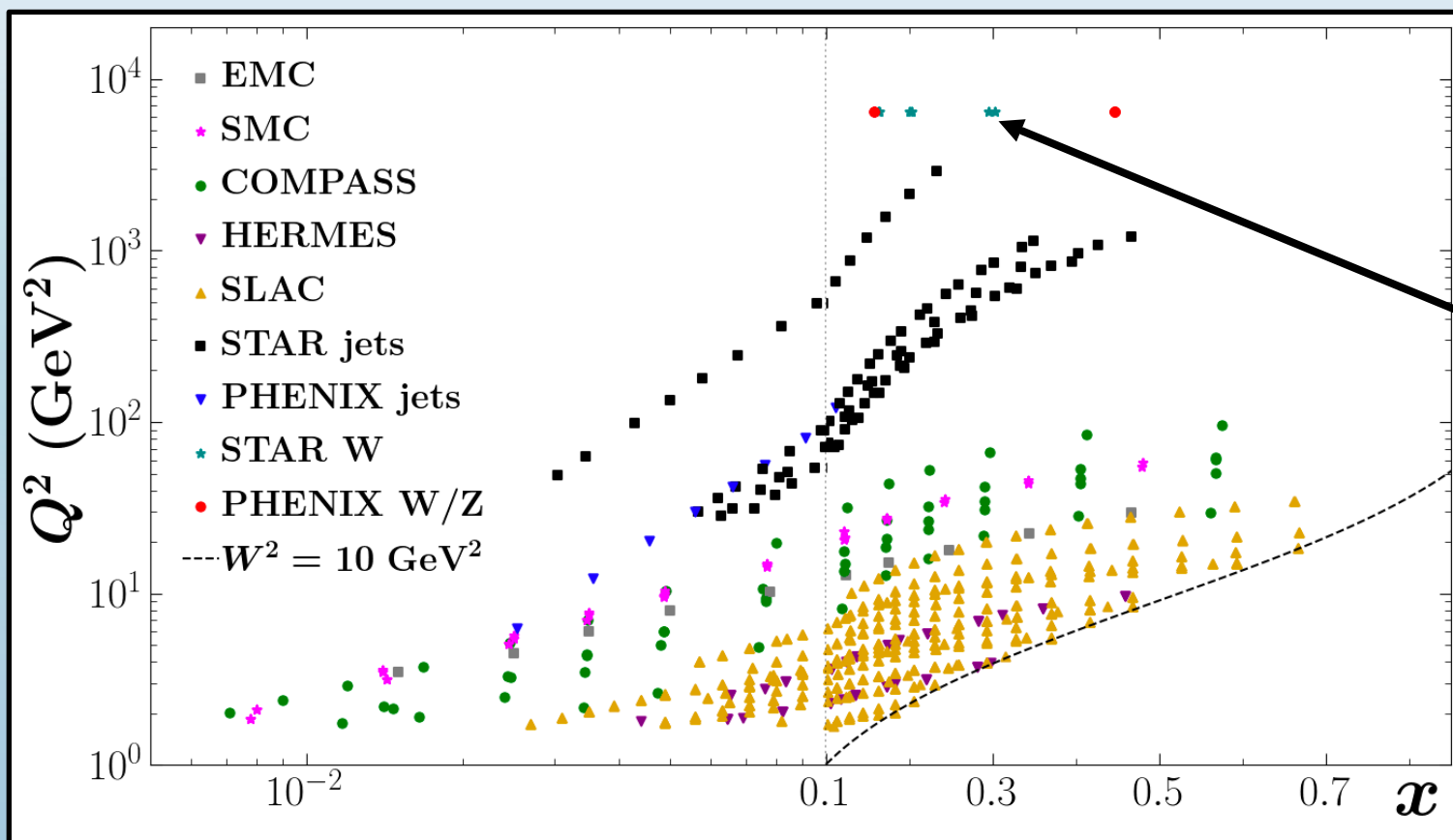
New STAR data

New SeaQuest data



# Kinematic Coverage (Helicity)

Deep Inelastic Scattering	COMPASS, EMC, HERMES, SLAC, SMC	365 points
W/Z Boson Production	STAR, PHENIX	18 points
Jets	STAR, PHENIX	61 points



STAR + PHENIX  
W/Z Production

# Part 3:

# Spin-Averaged PDFs

## High Energy Physics – Phenomenology

*[Submitted on 2 Sep 2021]*

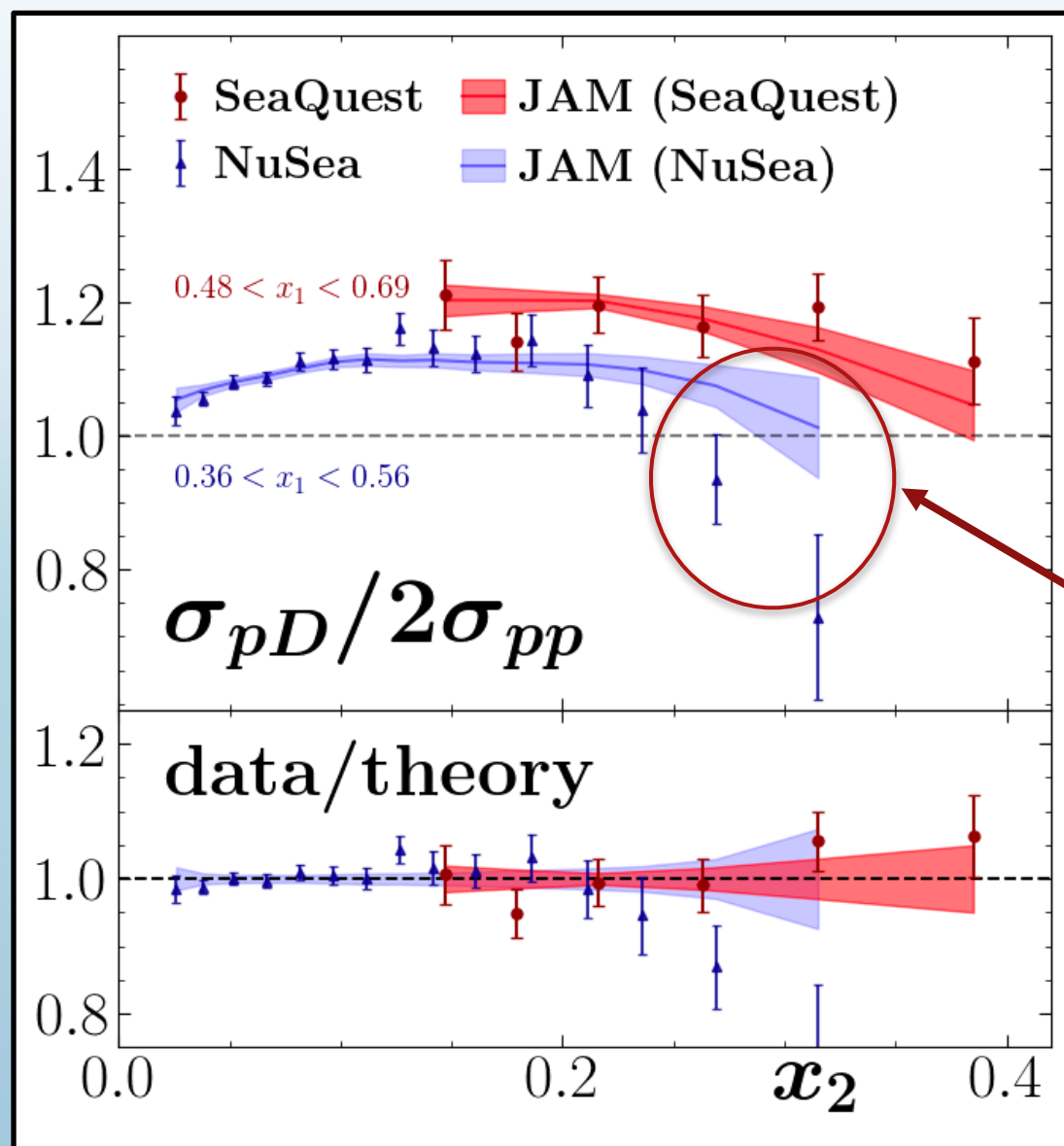
### Bayesian Monte Carlo extraction of sea asymmetry with SeaQuest and STAR data

Christopher Cocuzza, Wally Melnitchouk, Andreas Metz, Nobuo Sato

We perform a global QCD analysis of unpolarized parton distributions within a Bayesian Monte Carlo framework, including the new  $W$ -lepton production data from the STAR Collaboration at RHIC and Drell-Yan di-muon data from the SeaQuest experiment at Fermilab. We assess the impact of these two new measurements on the light antiquark sea in the proton, and the  $\bar{d} - \bar{u}$  asymmetry in particular. The SeaQuest data are found to significantly reduce the uncertainty on the  $\bar{d}/\bar{u}$  ratio at large parton momentum fractions  $x$ , strongly favoring an enhanced  $\bar{d}$  sea up to  $x \approx 0.4$ , in general agreement with nonperturbative calculations based on chiral symmetry breaking in QCD.

<https://arxiv.org/abs/2109.00677>

# SeaQuest and NuSea Quality of Fit

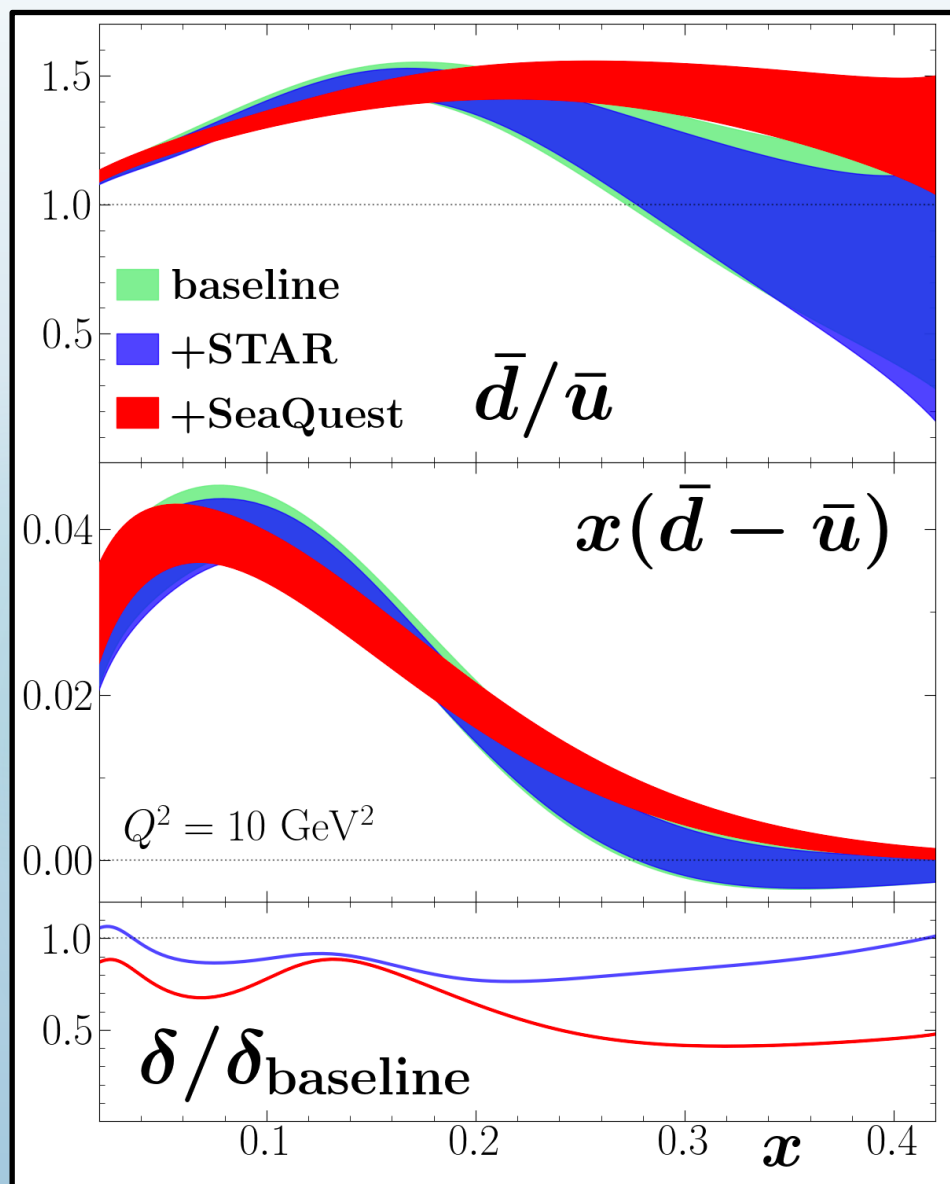


process		$N_{\text{dat}}$	$\chi^2/N_{\text{dat}}$
Drell-Yan			
NuSea	$pp$	184	1.21
NuSea	$pD/2pp$	15	1.30
SeaQuest	$pD/2pp$	6	0.82

$$\left. \frac{\sigma_{pD}}{2\sigma_{pp}} \right|_{x_1 \gg x_2} \approx \frac{1}{2} \left[ 1 + \frac{\bar{d}(x_2)}{\bar{u}(x_2)} \right]$$

Well-known tension  
between NuSea and  
SeaQuest

# Impact from STAR and SeaQuest



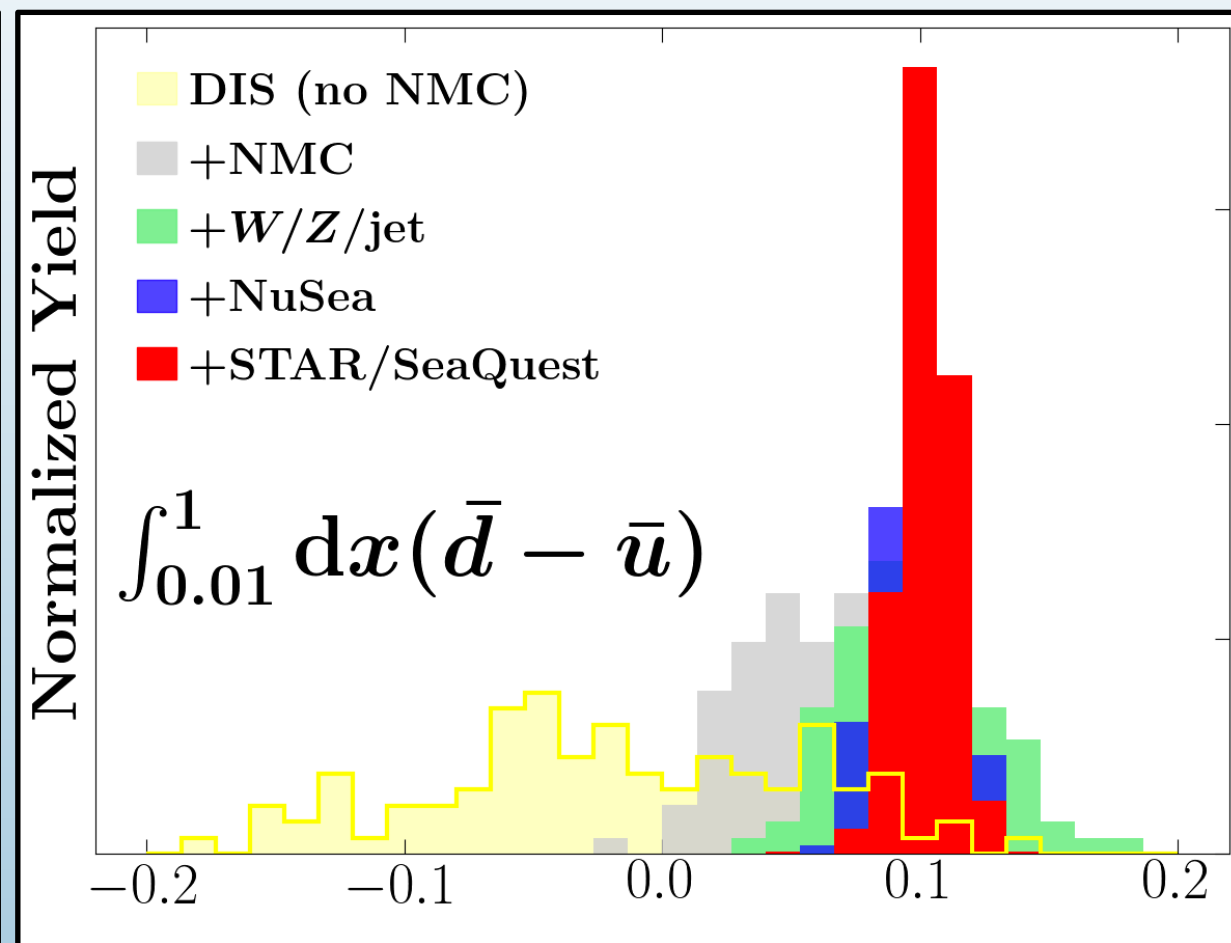
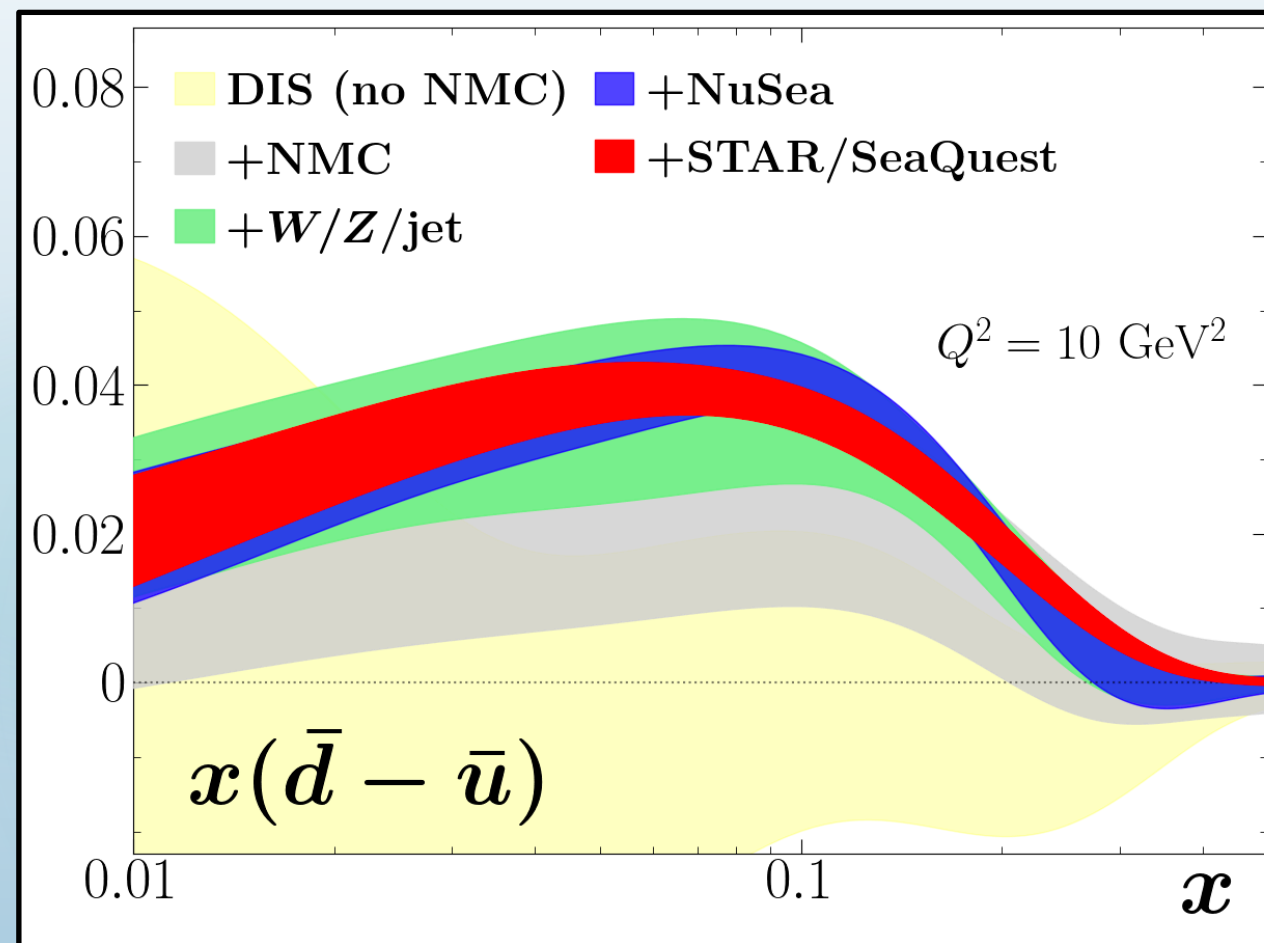
**STAR:** Moderate reduction of uncertainties

**SeaQuest:** Large reduction of uncertainties, especially at  $x > 0.2$ .

$\bar{d}/\bar{u} > 1$  up to  $x \approx 0.4$ , in agreement with models

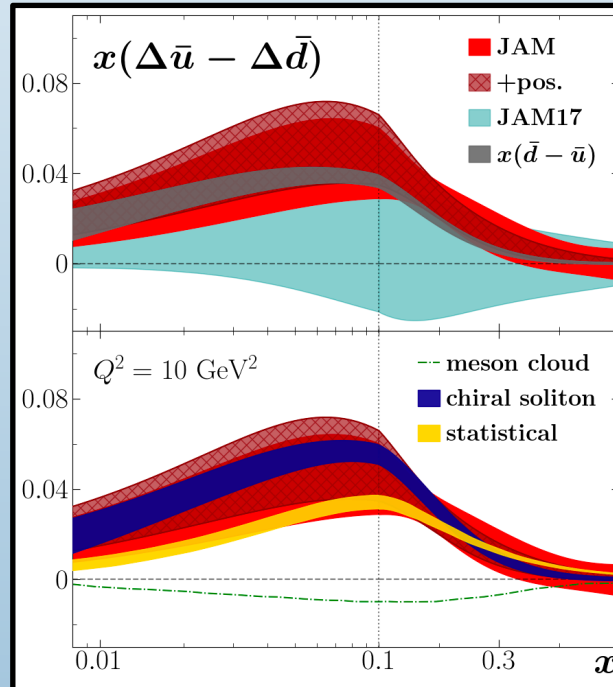


# Sources of Asymmetry

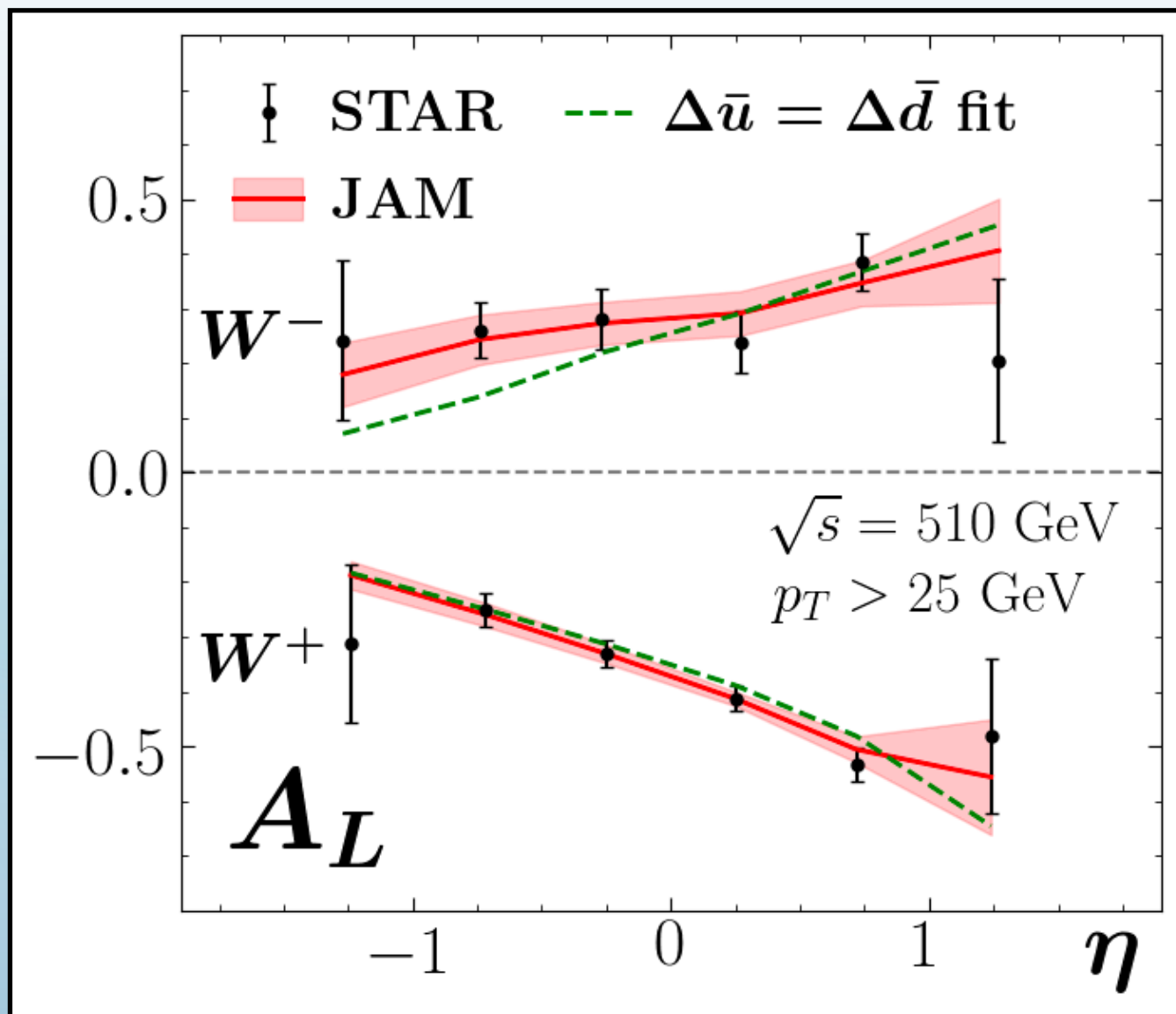


# Part 4:

# Helicity PDFs



# STAR Quality of Fit

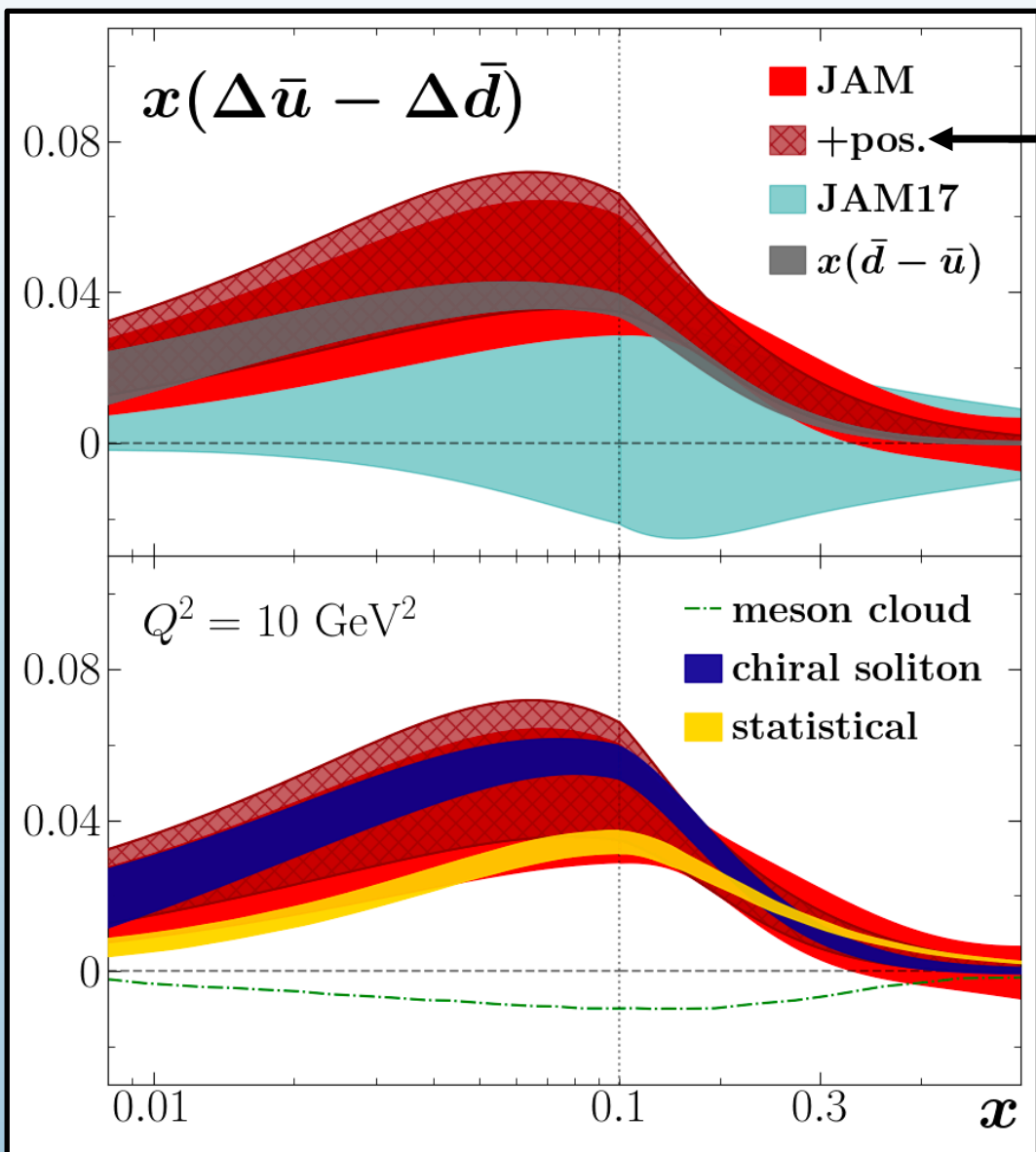


process	$N_{\text{dat}}$	JAM	$\chi^2/N_{\text{dat}}$ +Pos.	$\Delta\bar{u} = \Delta\bar{d}$
STAR $W^\pm$	12	0.45	0.61	1.53
PHENIX $W^\pm/Z$	6	0.47	0.46	0.48
pol. DIS	365	0.93	0.93	0.93
pol. jet	61	1.00	1.03	1.00
<b>total</b>	<b>444</b>	<b>0.92</b>	<b>0.94</b>	<b>0.95</b>

$$A_L^{W^+}(y_W) \propto \frac{\Delta\bar{d}(x_1)u(x_2) - \Delta u(x_1)\bar{d}(x_2)}{\bar{d}(x_1)u(x_2) + u(x_1)\bar{d}(x_2)}$$

$$A_L^{W^-}(y_W) \propto \frac{\Delta\bar{u}(x_1)d(x_2) - \Delta d(x_1)\bar{u}(x_2)}{\bar{u}(x_1)d(x_2) + d(x_1)\bar{u}(x_2)}$$

# Resulting Asymmetry



Positivity Constraints:  
 $|\Delta f(x, Q^2)| < f(x, Q^2)$

**JAM17**: inclusive +  
 semi-inclusive DIS data

Agreement with **Statistical** and  
**Chiral Soliton** models

**Meson Cloud** model is not  
 compatible with extraction

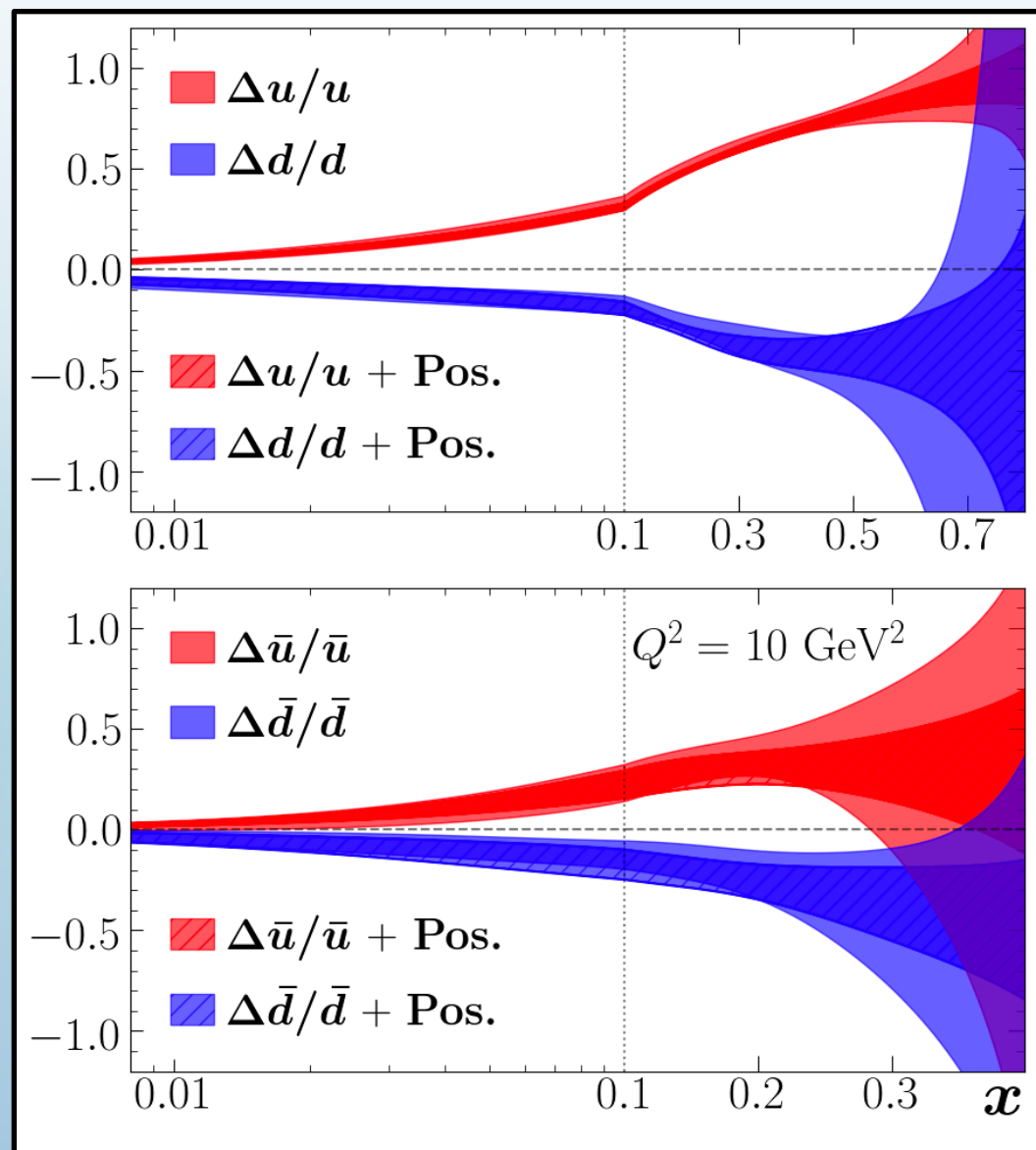
Statistical Model: C. Bourrely and J. Soffer, Nucl. Phys. **A941**, 307-334 (2015)

Meson Cloud Model: F. G. Cao and A. I. Signal, Phys. Rev. D. **68**, 074002 (2003)

Chiral Soliton Model: M. Wakamatsu and T. Watabe, Phys. Rev. D. **874**, 38-84 (2013)



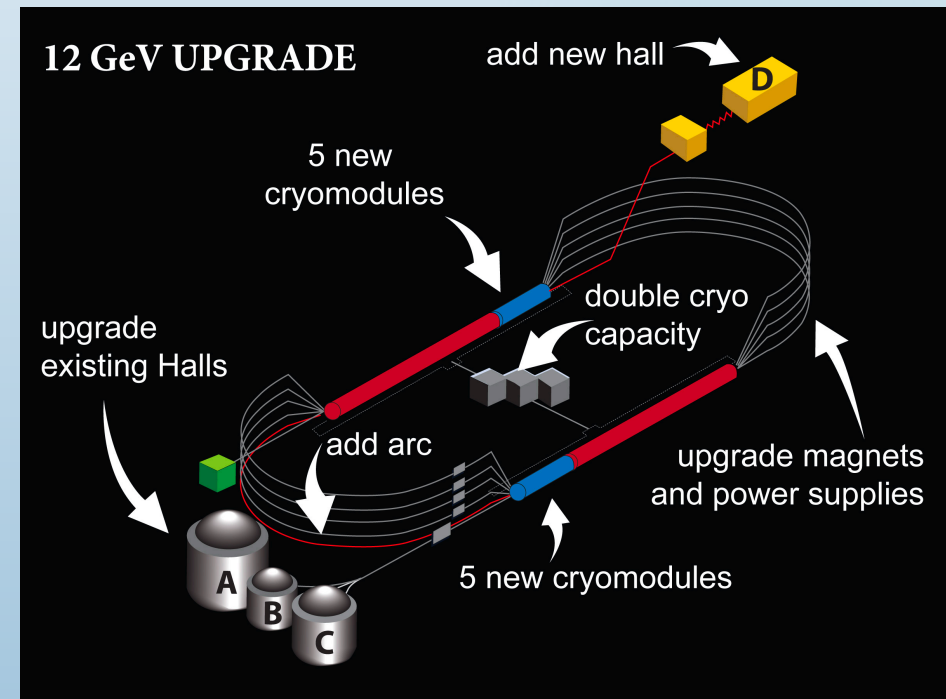
# Quark and Antiquark Polarizations



First self-consistent extraction  
using *simultaneous* fit

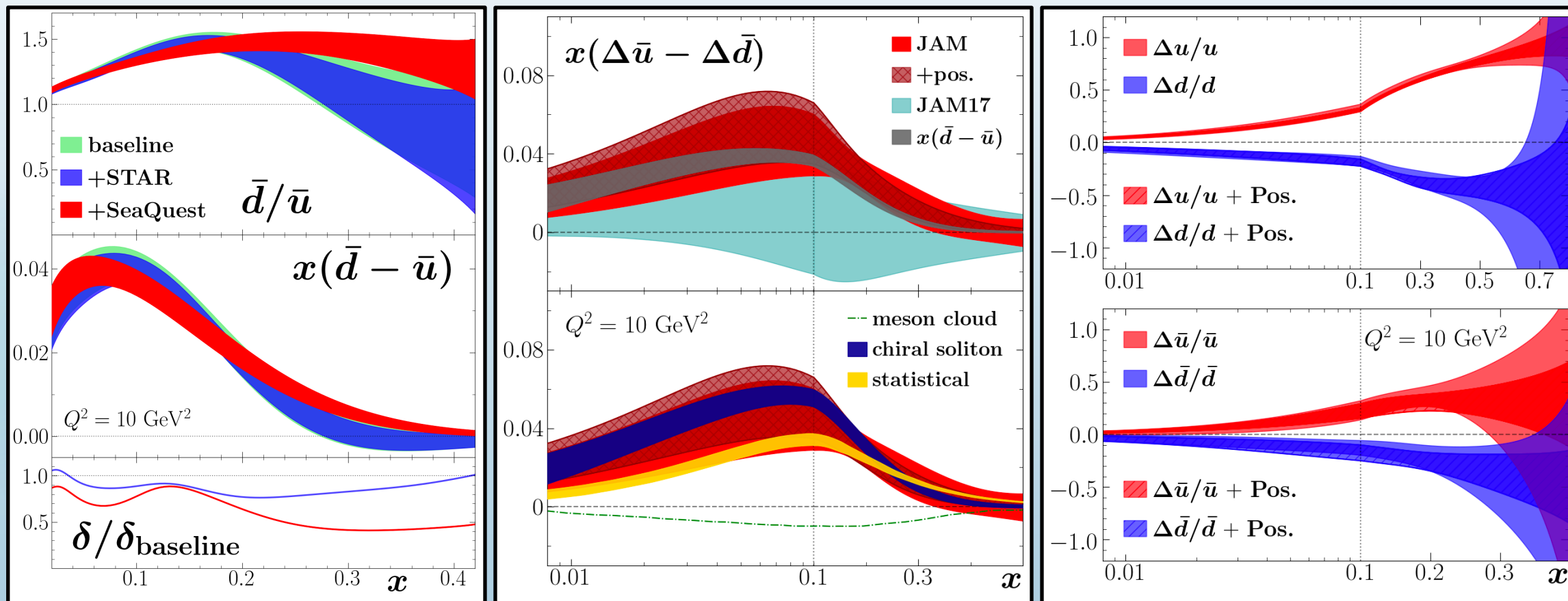
Antiquark ratios have same  
signs as quark ratios

# Conclusions and Outlook



# Results Summary

## First global QCD analysis of SeaQuest and STAR data



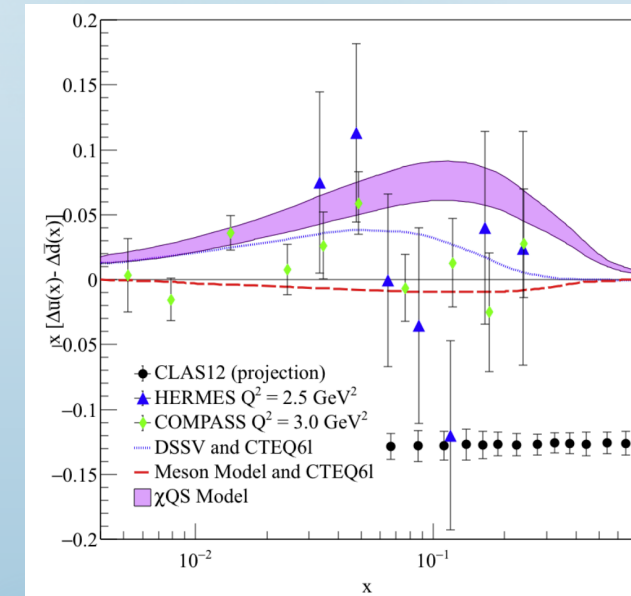
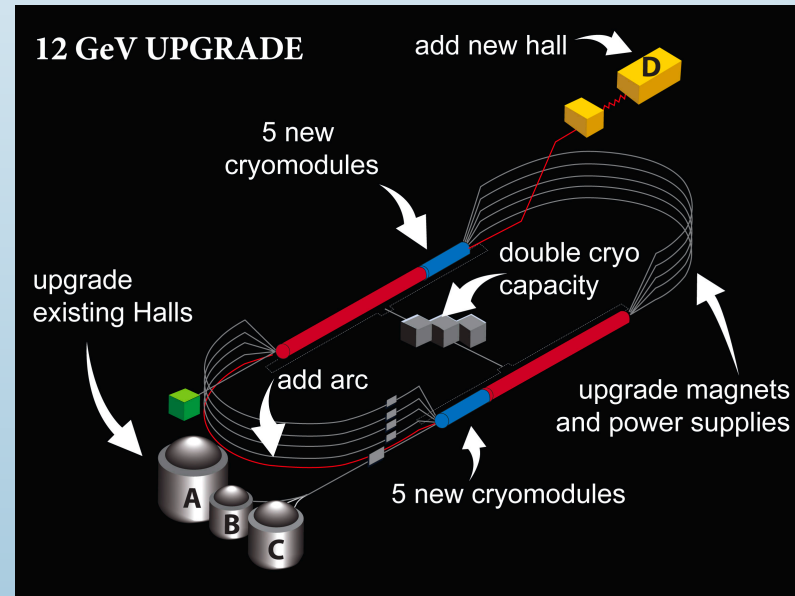
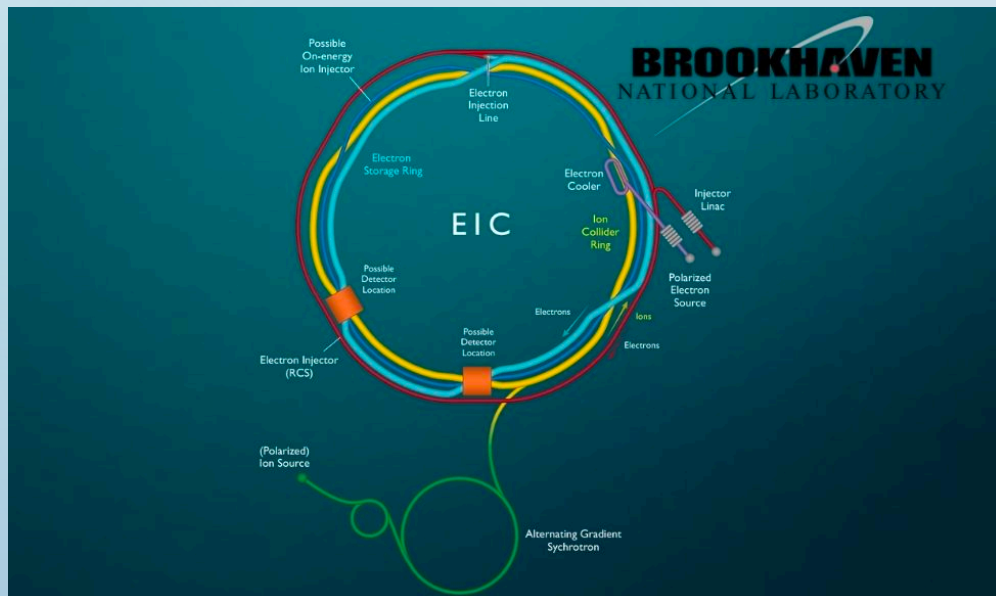
Simultaneous global QCD analysis of spin-averaged and helicity PDFs

# Outlook

Combine analysis with semi-inclusive DIS data from HERMES, COMPASS.

**Jefferson Lab CLAS12: Semi-inclusive DIS**

**EIC: First polarized electron-ion collider**





# Collaboration

Andreas Metz



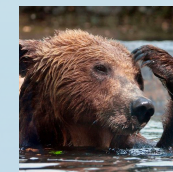
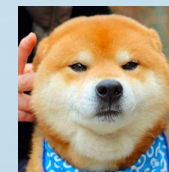
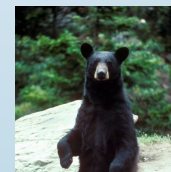
Wally Melnitchouk



Nobuo Sato



Thank you to Jacob Ethier, Yiyu Zhou, and Patrick Barry for helpful discussions





# Extra Slides

---

# Bayes' Theorem

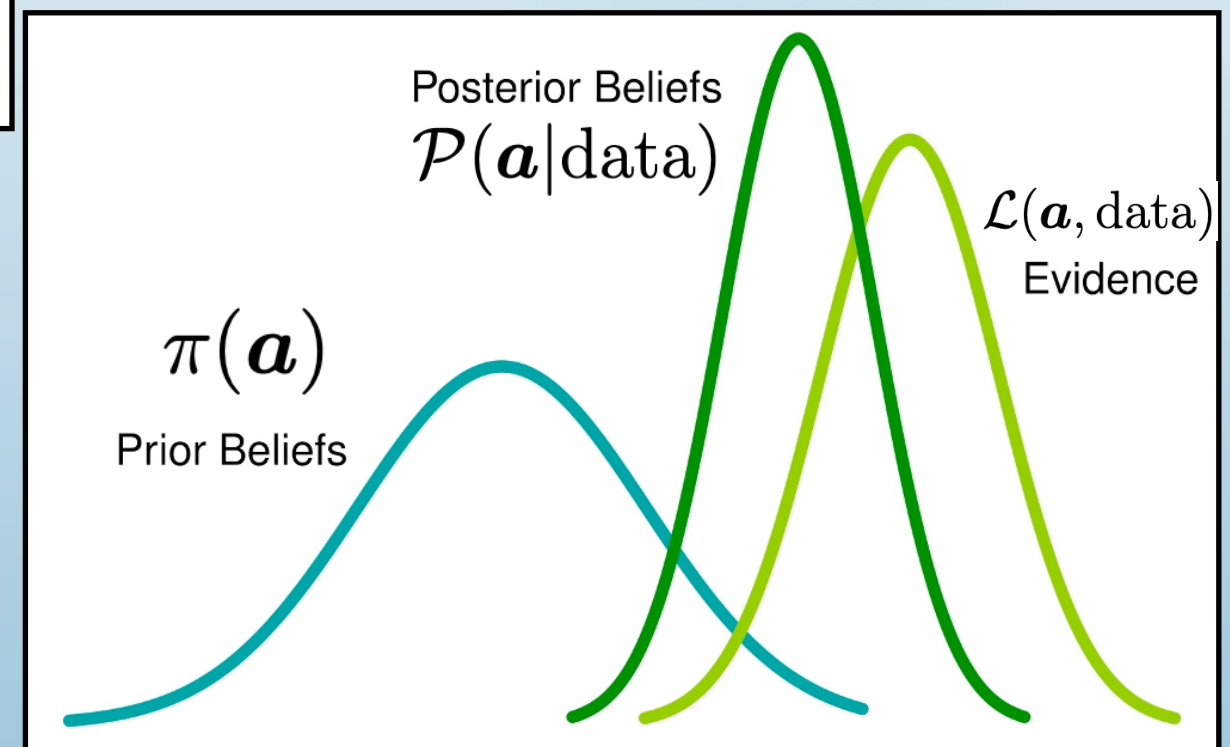
Now that we have calculated  $\chi^2(\mathbf{a}, \text{data}) \dots$

Likelihood Function

$$\mathcal{L}(\mathbf{a}, \text{data}) = \exp \left( -\frac{1}{2} \chi^2(\mathbf{a}, \text{data}) \right)$$

Bayes' Theorem

$$\mathcal{P}(\mathbf{a}|\text{data}) \sim \mathcal{L}(\mathbf{a}, \text{data}) \pi(\mathbf{a})$$



# Data Resampling

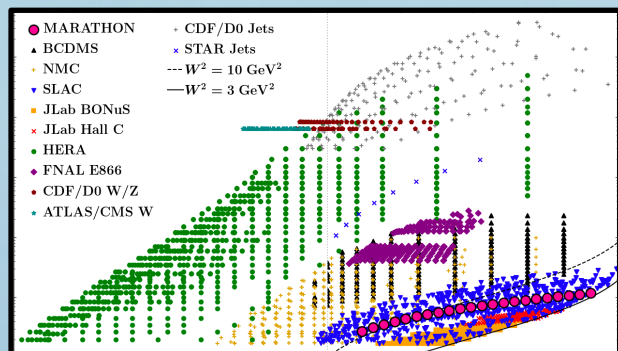
Pseudo-Data

$$\tilde{\sigma} = \sigma + N(0,1) \alpha$$

Uncorrelated  
Uncertainties

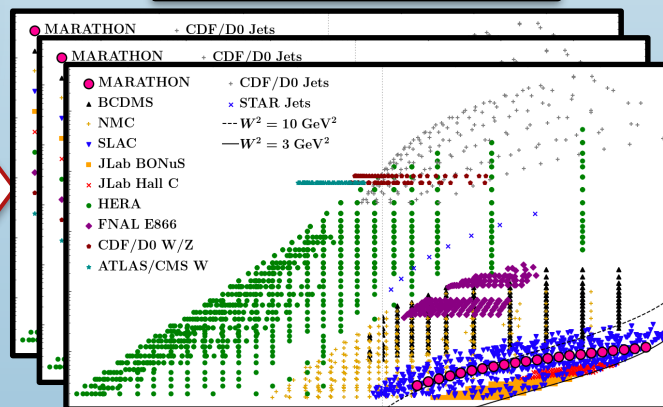
Data

Original Data



DR

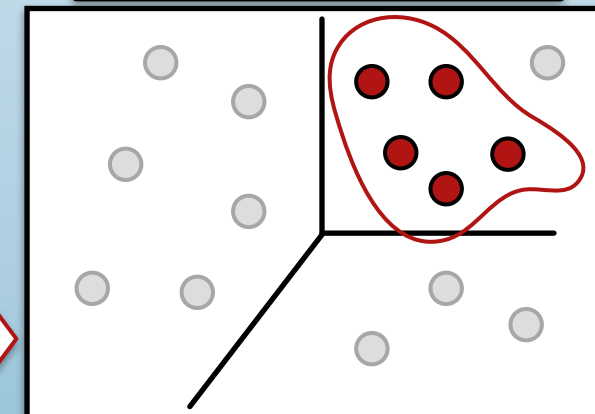
Replica Data


Maximum  
Likelihood

Maximum  
Likelihood

Maximum  
Likelihood

Parameter Space



# Error Quantification

For a quantity  $O(\mathbf{a})$ : (for example, a PDF at a given value of  $(x, Q^2)$ )

$$E[O] = \int d^n a \, \rho(\mathbf{a} \mid \text{data}) \, O(\mathbf{a})$$

$$V[O] = \int d^n a \, \rho(\mathbf{a} \mid \text{data}) \, [O(\mathbf{a}) - E[O]]^2$$

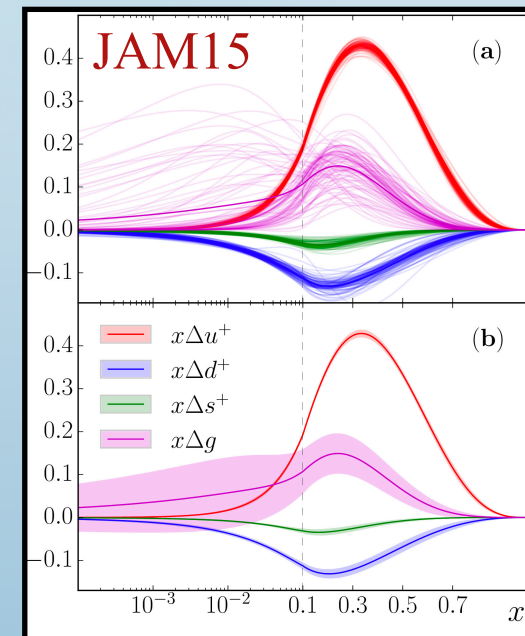
Build an MC ensemble

$$E[O] \approx \frac{1}{N} \sum_k O(\mathbf{a}_k)$$

$$V[O] \approx \frac{1}{N} \sum_k [O(\mathbf{a}_k) - E[O]]^2$$

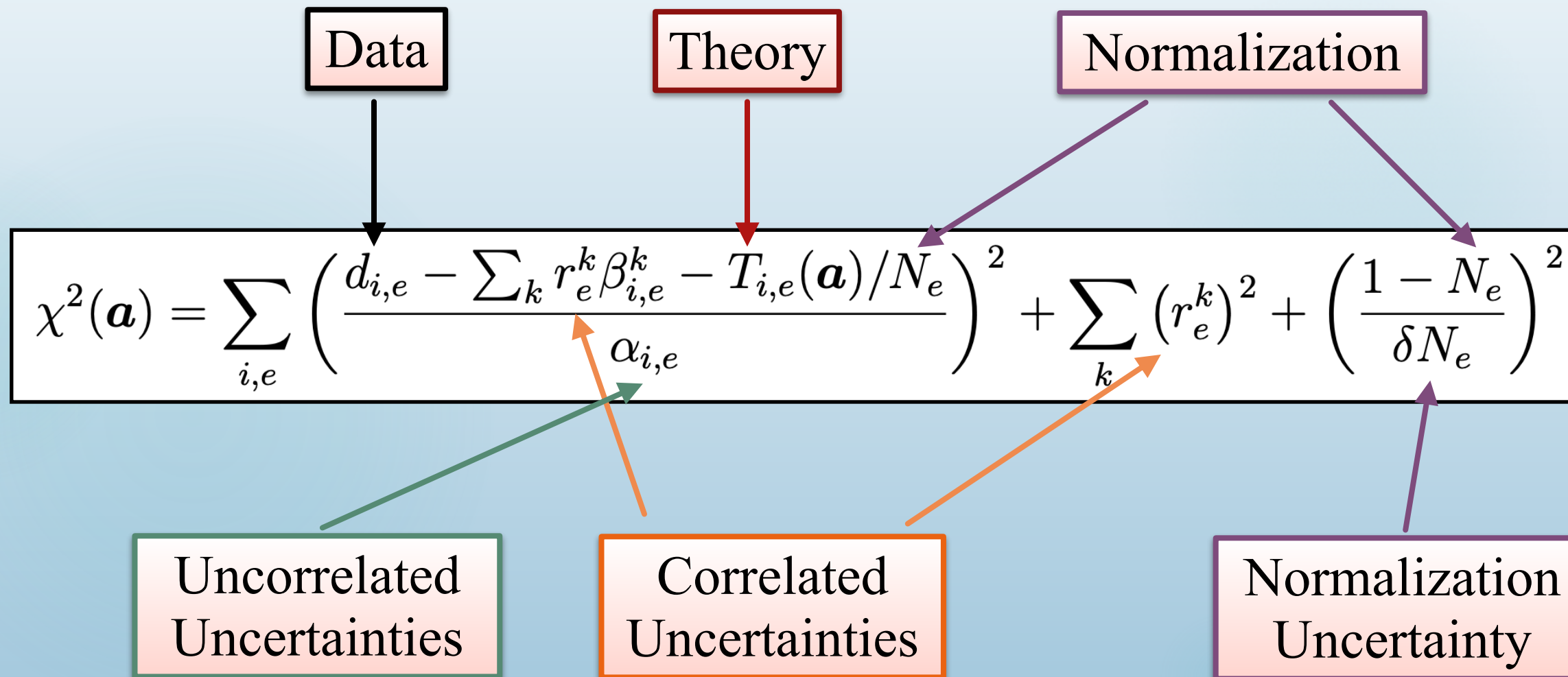
Exact, but  
 $n = \mathcal{O}(100)$ !

Average over  $k$  sets  
of the parameters  
(replicas)



# The $\chi^2$ function

Now that the observables have been calculated...



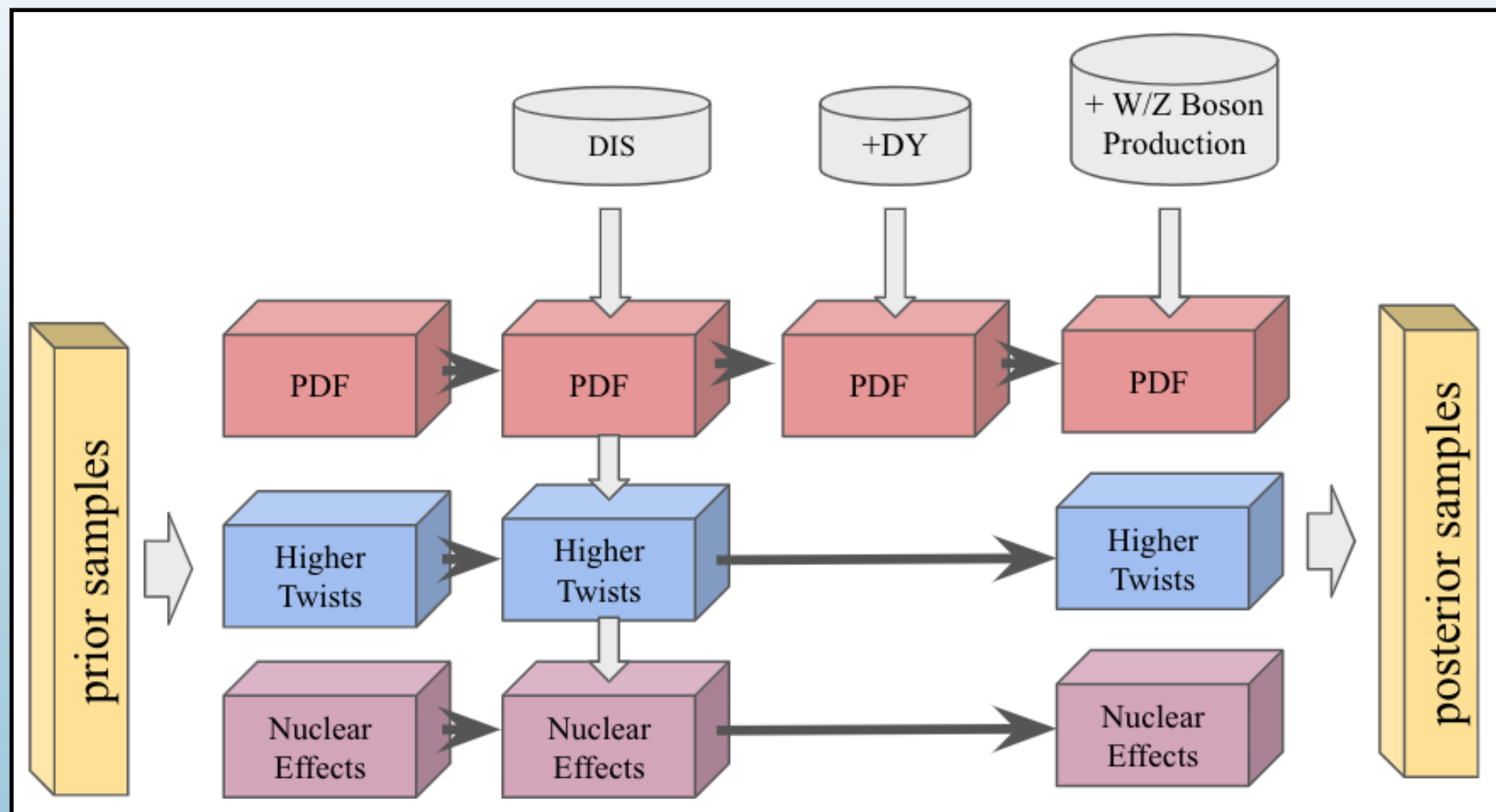
$$\chi^2(\mathbf{a}) = \sum_{i,e} \left( \frac{d_{i,e} - \sum_k r_e^k \beta_{i,e}^k - T_{i,e}(\mathbf{a})/N_e}{\alpha_{i,e}} \right)^2 + \sum_k (r_e^k)^2 + \left( \frac{1 - N_e}{\delta N_e} \right)^2$$

Diagram illustrating the components of the  $\chi^2$  function:

- Data** (points to  $d_{i,e}$ )
- Theory** (points to  $T_{i,e}(\mathbf{a})/N_e$ )
- Normalization** (points to  $N_e$ )
- Uncorrelated Uncertainties** (points to  $\alpha_{i,e}$ )
- Correlated Uncertainties** (points to  $\beta_{i,e}^k$ )
- Normalization Uncertainty** (points to  $\delta N_e$ )



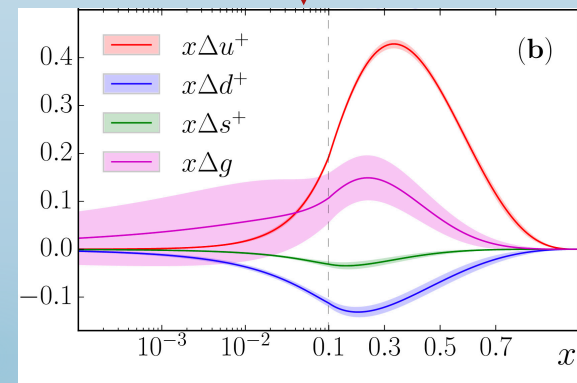
# Multi-Step Strategy



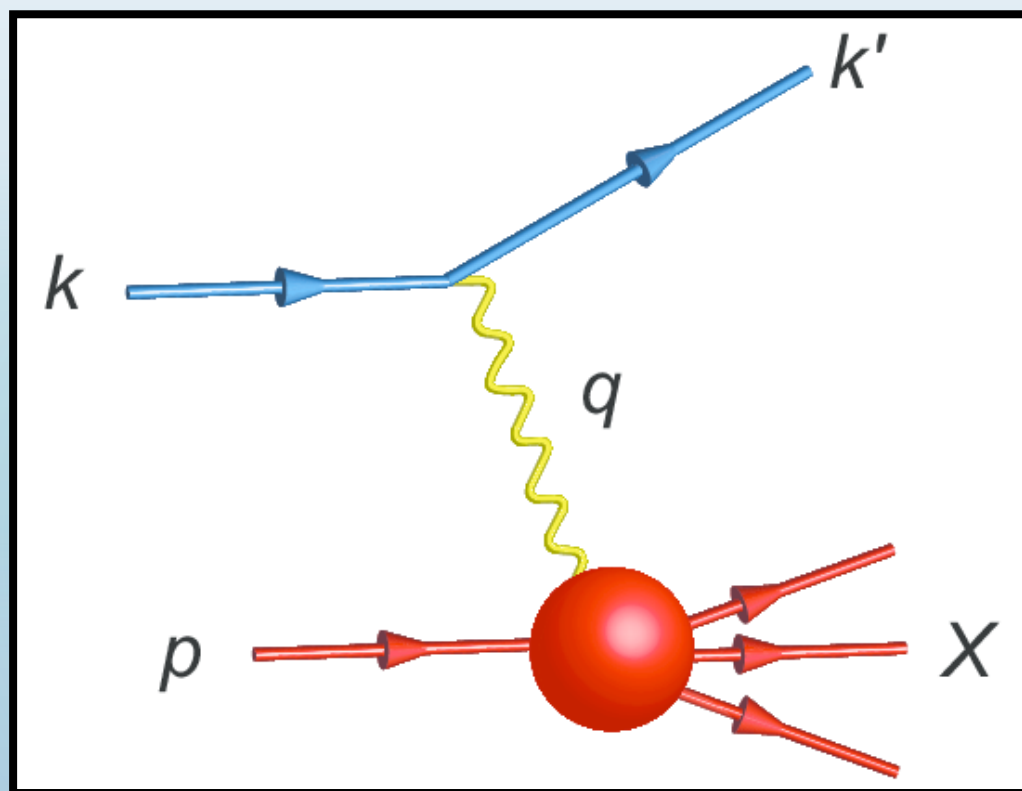
The flowchart illustrates the process of generating new priors from an ensemble of models. It starts with a 'data' node (green rounded rectangle) and a 'priors' node (pink rounded rectangle). The 'data' node branches into three 'pseudo data' nodes (green rounded rectangles): 'pseudo data<sub>1</sub>', 'pseudo data<sub>2</sub>', and 'pseudo data<sub>K</sub>'. Each 'pseudo data' node branches into two nodes: 'T data<sub>i</sub>' (blue rounded rectangle) and 'V data<sub>i</sub>' (pink rounded rectangle). Each 'T data<sub>i</sub>' node leads to a 'fit<sub>i</sub>' node (yellow oval), which then leads to a node containing the set of posterior probabilities  $\{\bar{p}^{(j)}\}$  (blue rounded rectangle). Each 'V data<sub>i</sub>' node leads to a 'validation' node (pink oval), which then leads to a node containing the validation posterior  $\bar{a}^{(i)}$  (blue rounded rectangle). The 'priors' node also branches into three 'pseudo data' nodes. The 'validation' nodes for 'pseudo data<sub>1</sub>' and 'pseudo data<sub>2</sub>' lead to an 'ensemble' node (white rounded rectangle). The 'ensemble' node leads to a 'new priors' node (white rounded rectangle). The 'new priors' node then feeds back into the 'priors' node, completing the loop. A red curved arrow points from the 'ensemble' node to the 'new priors' node.







# Deep Inelastic Scattering



Virtuality:

$$Q^2 = -q^2$$

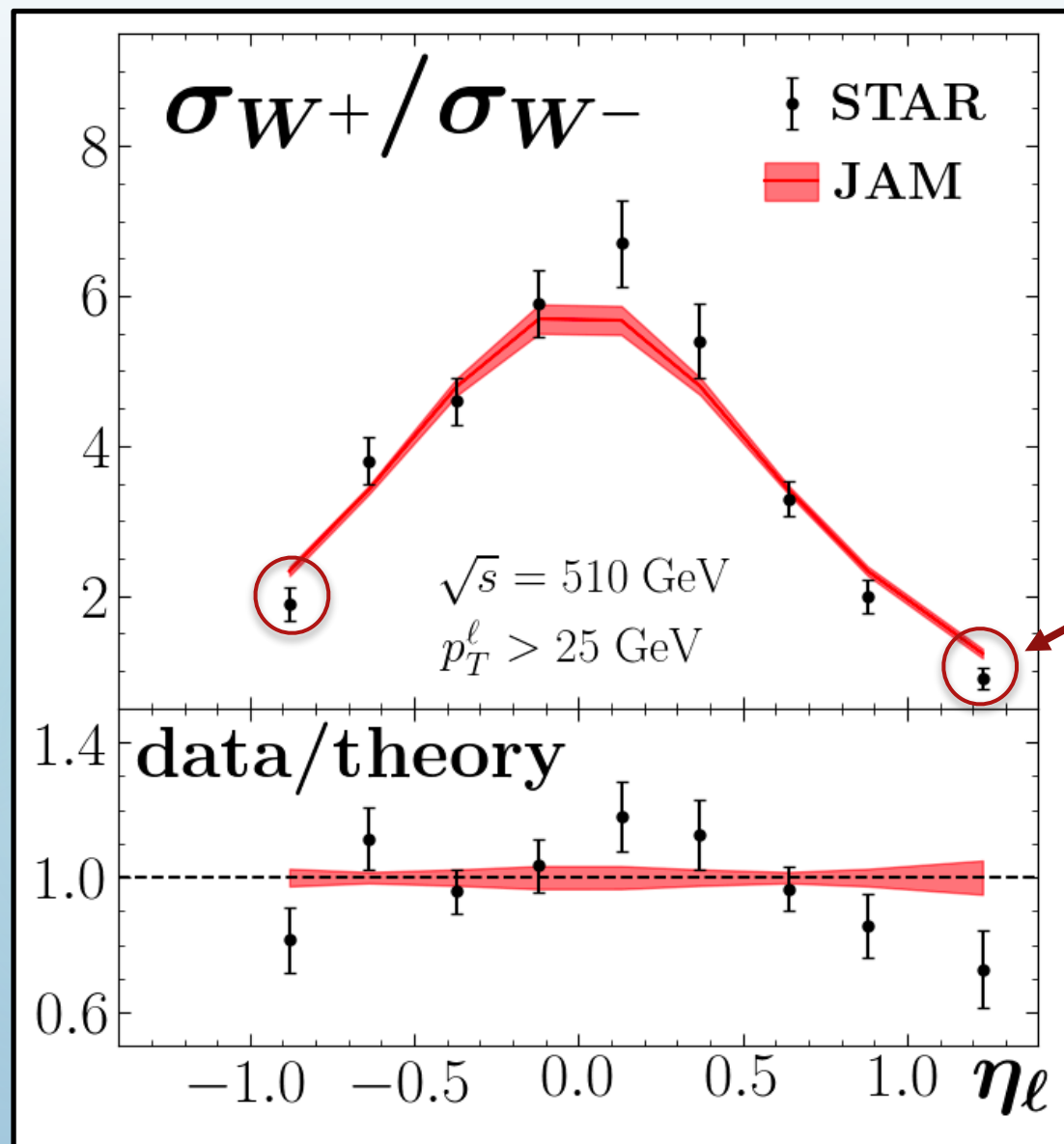
Bjorken  $x$ :

$$x = \frac{Q^2}{2p \cdot q}$$

Invariant mass of  
outgoing particles:

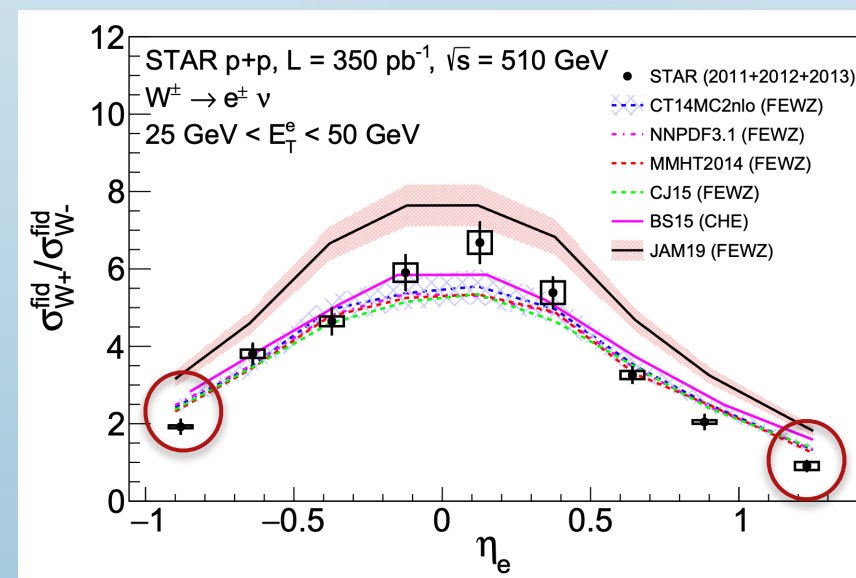
$$W^2 = (p + q)^2$$

# STAR Quality of Fit



process	$N_{\text{dat}}$	$\chi^2/N_{\text{dat}}$
W-lepton		
STAR $W^+/W^-$	9	2.02

Difficult to describe at extreme rapidity

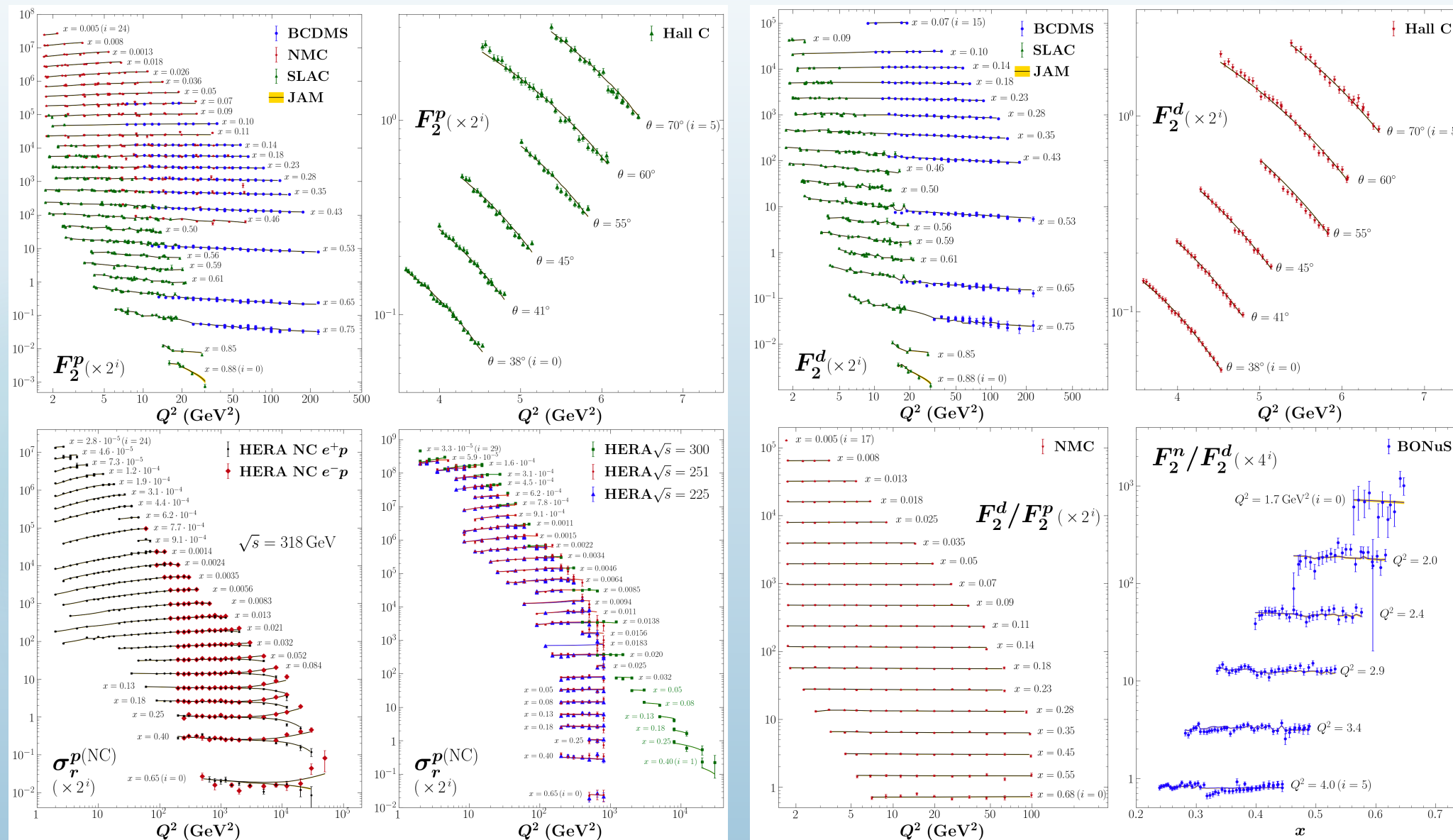


# All $\chi^2/N_{\text{dat}}$ (Spin-Averaged)

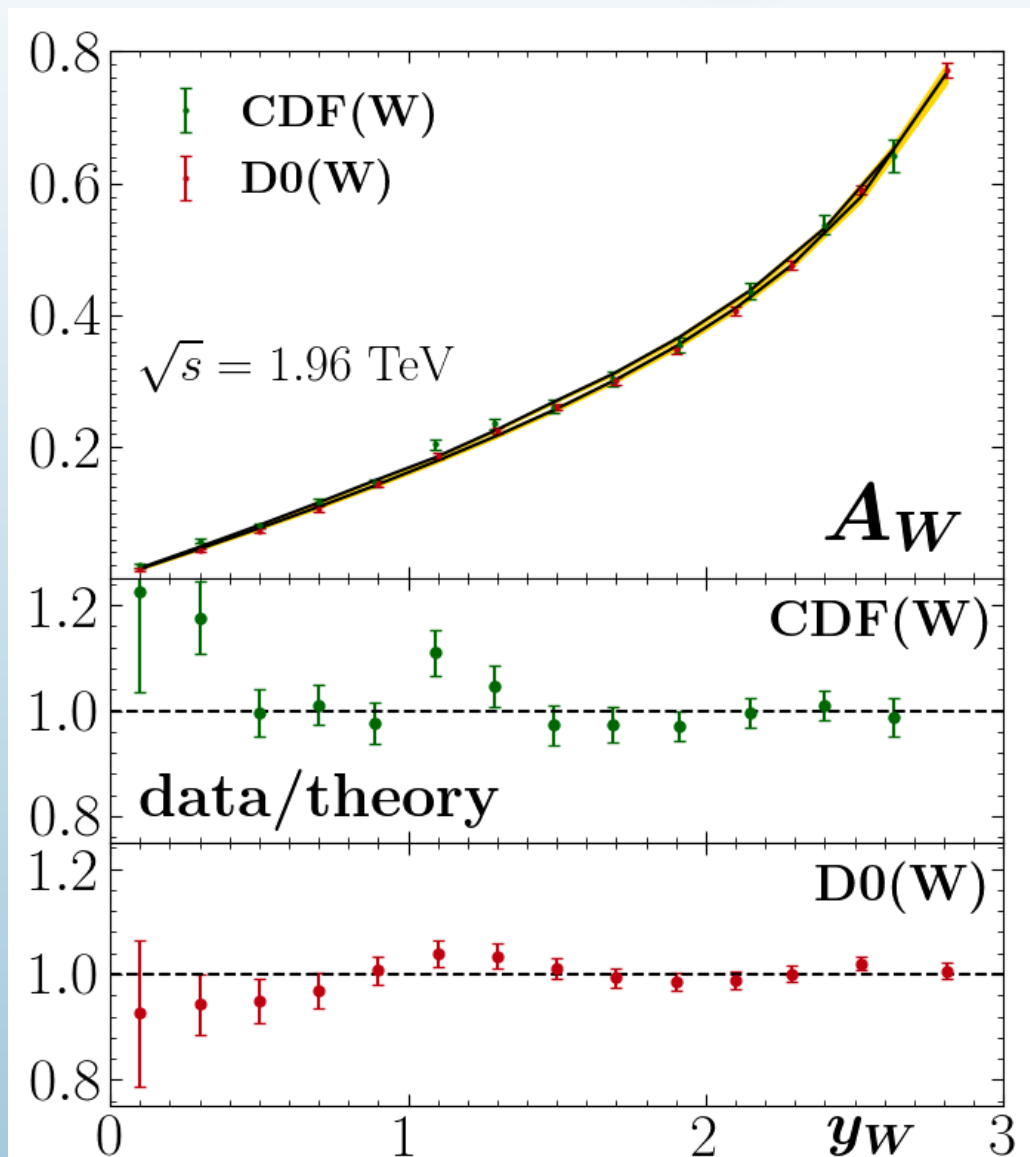
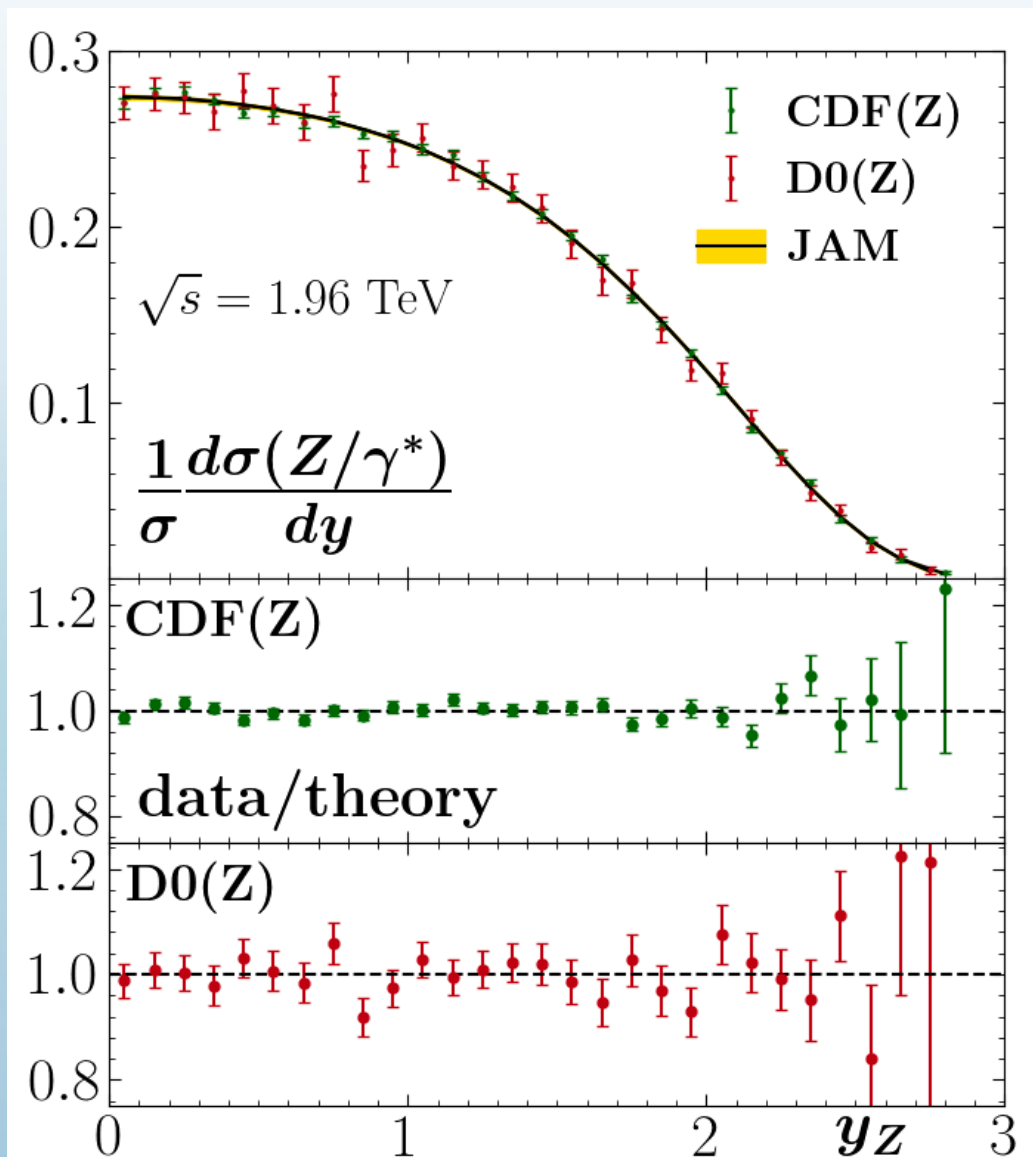
process	$N_{\text{dat}}$	$\chi^2/N_{\text{dat}}$
DIS		
fixed target	2678	1.05
HERA	1185	1.27
Drell-Yan		
NuSea $pp$	184	1.21
NuSea $pD/2pp$	15	1.30
SeaQuest $pD/2pp$	6	0.82
$W$ -lepton		
STAR $W^+/W^-$	9	2.02
CMS charm asym.	45	0.74
LHCb charm asym.	16	0.44
Tevatron $W$ charge asym.	27	1.18
Tevatron $Z$ rapidity	56	0.97
jet	200	1.11
<b>total</b>	<b>4421</b>	<b>1.12</b>



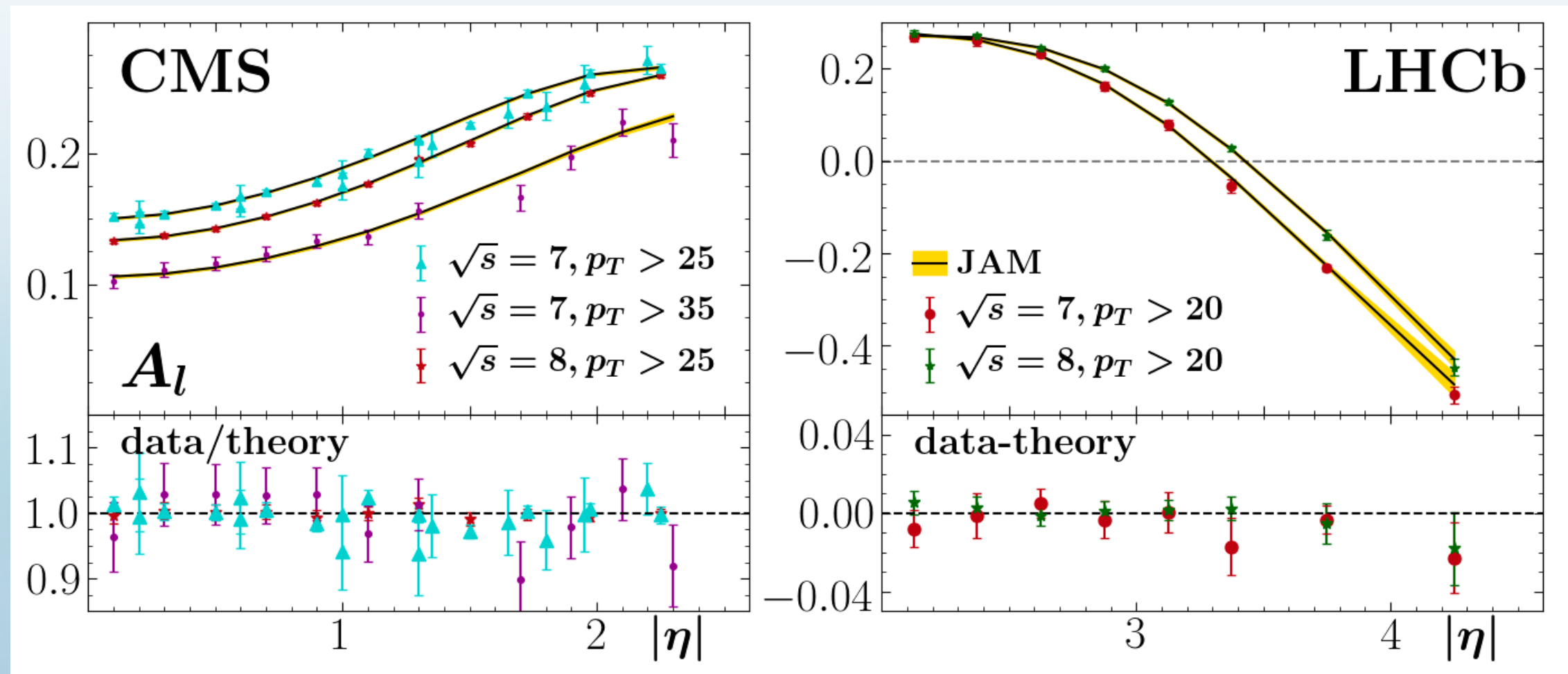
# DIS (Neutral Current)



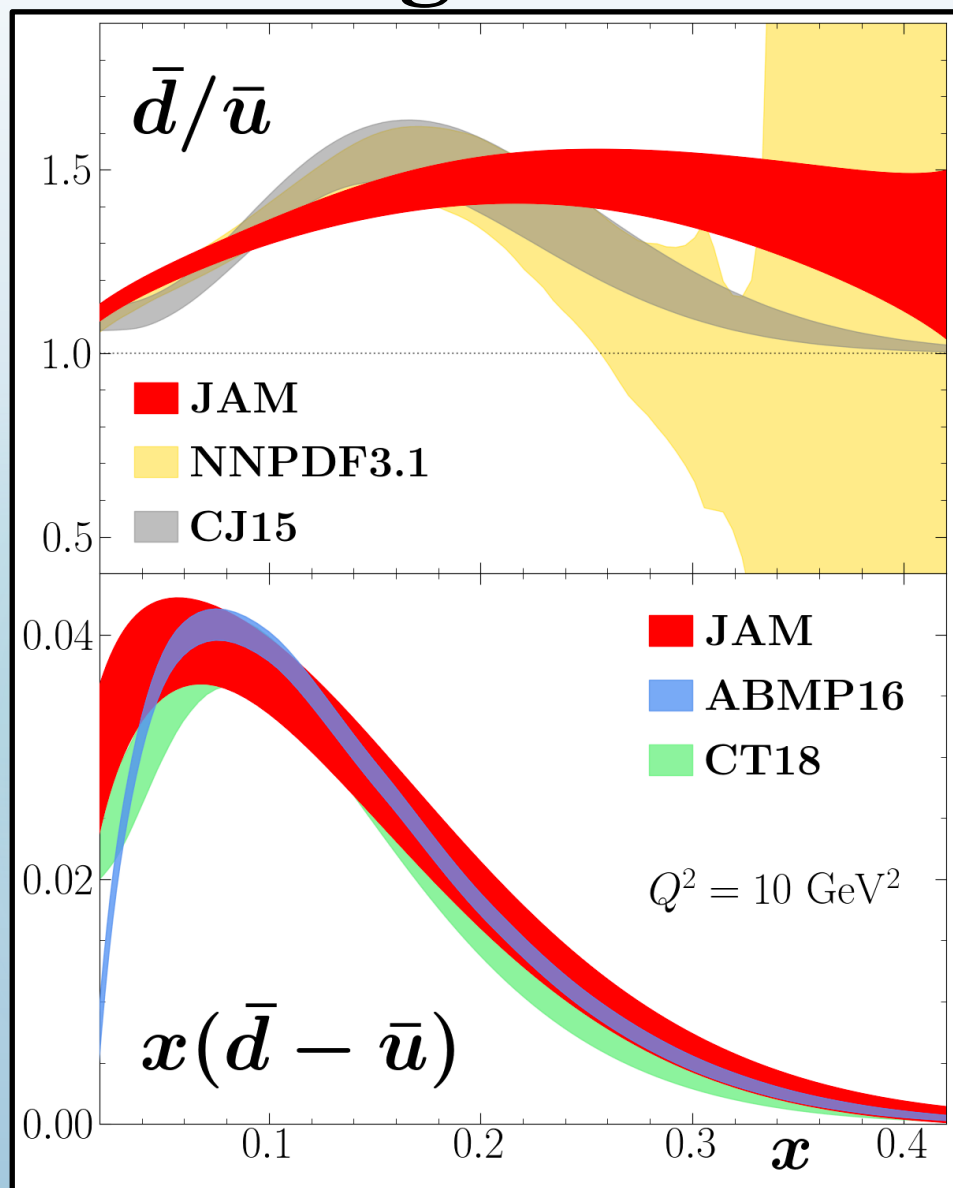
# W/Z Boson Production



# Lepton Production



# Resulting PDFs



Results for asymmetry largely agree with  
 ABMP16, CT18;  
 disagree with NNPDF3.1, CJ15 at high  $x$ .

