
Simultaneous QCD analysis of quantum probability distributions

Wally Melnitchouk

— Jefferson Lab —

Overview

- JAM (Jefferson Lab Angular Momentum) collaboration studies the parton structure of hadrons through extraction of “quantum probability distributions” (PDFs, FFs, TMDs) via global QCD analysis using Monte Carlo-based methods
- Methodology is based on Bayesian statistics and Monte Carlo sampling of the parameter space
- Inter-dependence of observables on distributions requires simultaneous extraction of unpolarized and polarized PDFs & fragmentation functions

Bibliography

- *Strange quark suppression from a simultaneous Monte Carlo analysis of PDFs and fragmentation functions*
Nobuo Sato, Carlota Andres, Jake Ethier, WM: [arXiv:1905.03788 \[hep-ph\]](#)
- *First Monte Carlo global QCD analysis of pion parton distributions*
Patrick Barry, Nobuo Sato, WM, Chueng Ji: [Phys. Rev. Lett. **121**, 152001 \(2018\)](#)
- *First Monte Carlo global analysis of nucleon transversity with lattice QCD constraints*
Huey-Wen Lin, WM, Alexei Prokudin, Nobuo Sato, Harvey Shows: [Phys. Rev. Lett. **120**, 152502 \(2018\)](#)
- *First simultaneous extraction of spin-dependent PDFs and fragmentation functions from a global QCD analysis*
Jake Ethier, Nobuo Sato, WM: [Phys. Rev. Lett. **119**, 132001 \(2017\)](#)
- *First Monte Carlo analysis of fragmentation functions from e^+e^- annihilation*
Nobuo Sato, Jake Ethier, WM, Hirai, Kumano, Alberto Accardi: [Phys. Rev. D **94**, 114004 \(2016\)](#)
- *Iterative Monte Carlo analysis of spin-dependent parton distributions*
Nobuo Sato, WM, Sebastian Kuhn, Jake Ethier, Alberto Accardi: [Phys. Rev. D **93**, 074005 \(2016\)](#)
- *Constraints on spin-dependent parton distributions at large x from global QCD analysis*
Pedro Jimenez-Delgado, Harut Avakian, WM: [Phys. Lett. B **738**, 263 \(2014\)](#)
- *Impact of hadronic and nuclear corrections on global analysis of spin-dependent PDFs*
Pedro Jimenez-Delgado, Alberto Accardi, WM: [Phys. Rev. D **89**, 034025 \(2014\)](#)

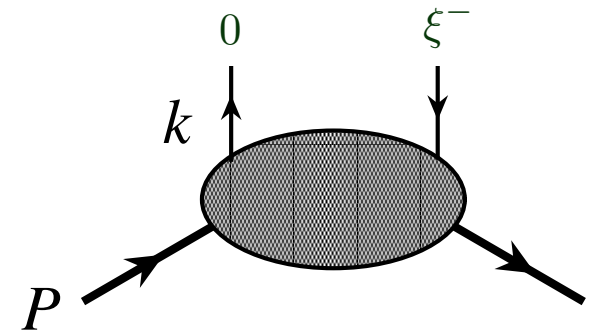
Parton distributions in hadrons

- Parton distribution functions (PDFs) are light-cone correlation functions

$$q(x) = \int_{-\infty}^{\infty} d\xi^- e^{-ixP^+\xi^-} \langle P | \bar{\psi}(\xi^-) \gamma^+ \mathcal{W}(\xi^-, 0) \psi(0) | P \rangle$$

- light cone momentum fraction $x = \frac{k^+}{P^+}$
- Wilson line (gauge invariance)

$$\mathcal{W}(\xi^-, 0) = \exp \left\{ -ig \int_0^{\xi^-} d\eta^- \mathcal{A}^+(\eta^-) \right\}$$



- In $\mathcal{A}^+ = 0$ gauge, in fast-moving frame PDF has a probabilistic interpretation as a particle density

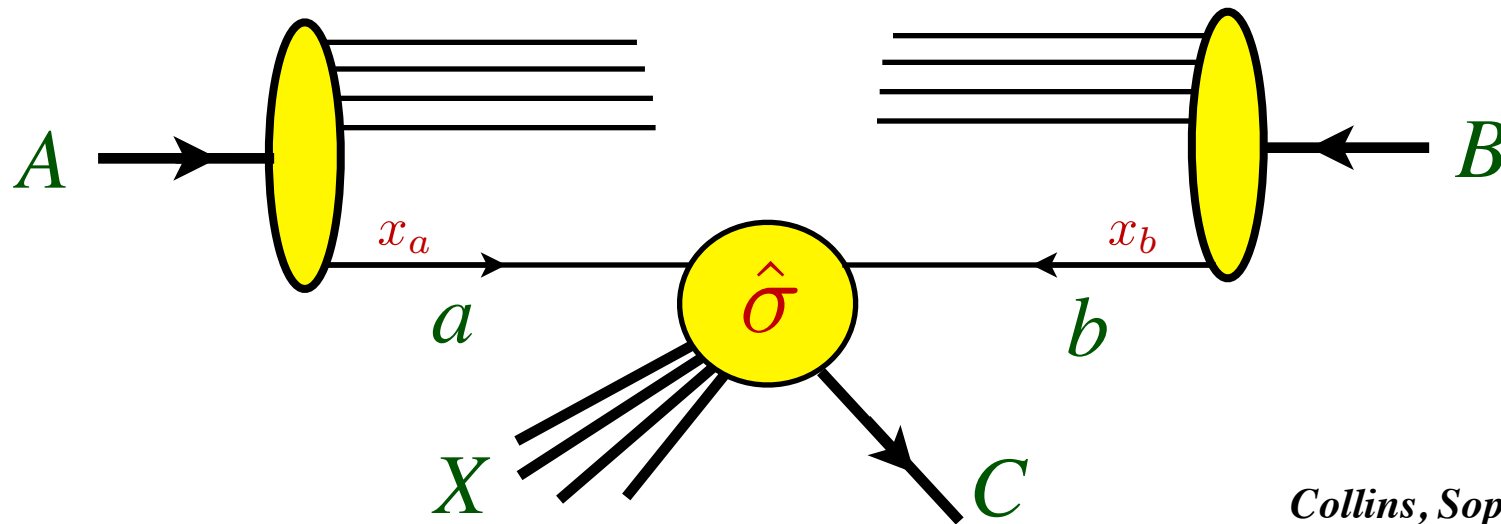
$$\int_{-1}^1 dx q(x) = \langle P | \bar{\psi}(0) \gamma^+ \psi(0) | P \rangle \approx \langle P | \psi^\dagger(0) \psi(0) | P \rangle$$

↑
↑
↑

number density
 $\bar{\psi} \gamma^0 \psi \approx \bar{\psi} \gamma^z \psi$
number operator

Parton distributions in hadrons

■ Inclusive high-energy particle production $AB \rightarrow CX$



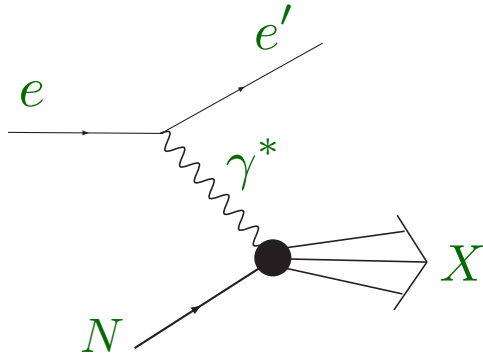
→ QCD factorization: separation of hard (perturbative, calculable) from soft (nonperturbative, parametrized) physics

$$\sigma_{AB \rightarrow CX}(p_A, p_B) = \sum_{a,b} \int dx_a dx_b \underbrace{f_{a/A}(x_a, \mu)}_{\dots\dots} \underbrace{f_{b/B}(x_b, \mu)}_{\dots\dots} \times \sum_n \alpha_s^n(\mu) \hat{\sigma}_{ab \rightarrow CX}^{(n)}(x_a p_A, x_b p_B, Q/\mu)$$

→ process-independent parton distribution functions $f_{a/A}$ characterizing structure of bound state A

Parton distributions in hadrons

- Most information on PDFs obtained from lepton-hadron deep-inelastic scattering (DIS)



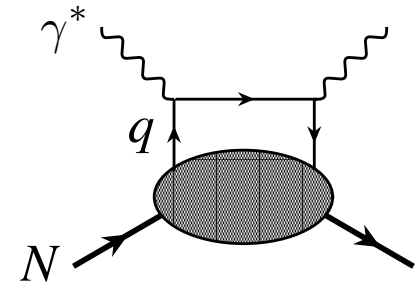
$$\frac{d^2\sigma}{d\Omega dE'} = \frac{4\alpha^2 E'^2 \cos^2 \frac{\theta}{2}}{Q^4} \left(2 \tan^2 \frac{\theta}{2} \frac{F_1}{M} + \frac{F_2}{\nu} \right)$$

$$x_B = \frac{Q^2}{2M\nu} \quad \begin{aligned} Q^2 &= \vec{q}^2 - \nu^2 \\ \nu &= E - E' \end{aligned}$$

→ structure function given as convolution of hard Wilson coefficient with PDF

$$F_2(x_B, Q^2) = x_B \sum_q e_q^2 \int_{x_B}^1 \frac{dx}{x} C_q\left(\frac{x_B}{x}, \alpha_s\right) q(x, Q^2)$$

$$\rightarrow x_B \sum_q e_q^2 q(x_B, Q^2)$$



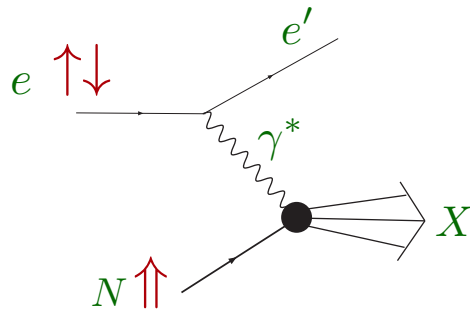
for leading order approximation $C_q \rightarrow \delta\left(1 - \frac{x_B}{x}\right)$

Parton distributions in hadrons

- Spin-dependent PDFs are defined similarly

$$\begin{aligned}\Delta q(x) &= q^\uparrow(x) - q^\downarrow(x) \\ &= \int_{-\infty}^{\infty} d\xi^- e^{-ixP^+\xi^-} \langle P | \bar{\psi}(\xi^-) \gamma^+ \gamma_5 \mathcal{W}(\xi^-, 0) \psi(0) | P \rangle\end{aligned}$$

→ measured in polarized lepton-nucleon DIS



$$A_{\parallel} = \frac{\sigma^{\uparrow\downarrow} - \sigma^{\uparrow\uparrow}}{\sigma^{\uparrow\downarrow} + \sigma^{\uparrow\uparrow}} = D(A_1 + \eta A_2)$$

$$A_1 \approx \frac{2x_B g_1(x_B, Q^2)}{F_2(x_B, Q^2)}$$

→ polarized structure function in terms of spin-dependent PDFs

$$g_1(x_B, Q^2) = \frac{1}{2} \sum_q e_q^2 \int_{x_B}^1 \frac{dx}{x} \Delta C_q\left(\frac{x_B}{x}, \alpha_s\right) \Delta q(x, Q^2)$$

$$\rightarrow \frac{1}{2} \sum_q e_q^2 \Delta q(x_B, Q^2) \quad \text{at leading order}$$

Global PDF analysis

- Universality of PDFs allows data from many different processes (DIS, SIDIS, weak boson/jet production in pp , Drell-Yan ...) to be analyzed simultaneously
 - distributions parametrized using a specific functional form, with parameters fitted to data
- Extraction of PDFs is challenging because usually there exist multiple solutions — “inverse problem”
 - PDFs are not directly measured, but inferred from observables involving convolutions with other functions

Bayesian approach to global analysis

- Analysis of data requires estimating expectation values E and variances V of “observables” \mathcal{O} (functions of PDFs) which are functions of parameters

$$E[\mathcal{O}] = \int d^n a \mathcal{P}(\vec{a}|\text{data}) \mathcal{O}(\vec{a})$$

$$V[\mathcal{O}] = \int d^n a \mathcal{P}(\vec{a}|\text{data}) [\mathcal{O}(\vec{a}) - E[\mathcal{O}]]^2$$

“Bayesian master formulas”

- Using Bayes’ theorem, probability distribution \mathcal{P} given by

$$\mathcal{P}(\vec{a}|\text{data}) = \frac{1}{Z} \mathcal{L}(\text{data}|\vec{a}) \pi(\vec{a})$$

in terms of the likelihood function \mathcal{L} and priors π

Bayesian approach to global analysis

■ Likelihood function

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

is a Gaussian form in the data, with χ^2 function

$$\chi^2(\vec{a}) = \sum_i \left(\frac{\text{data}_i - \text{theory}_i(\vec{a})}{\delta(\text{data})} \right)^2$$

with priors $\pi(\vec{a})$ and evidence Z

$$Z = \int d^n a \mathcal{L}(\text{data}|\vec{a}) \pi(\vec{a})$$

→ Z tests if *e.g.* an n -parameter fit is statistically different from $(n+1)$ -parameter fit

Bayesian approach to global analysis

- Standard method for evaluating E, V via maximum likelihood

→ maximize probability distribution

$$\mathcal{P}(\vec{a}|\text{data}) \rightarrow \vec{a}_0$$

→ if \mathcal{O} is linear in parameters, and if probability is symmetric in all parameters

$$E[\mathcal{O}(\vec{a})] = \mathcal{O}(\vec{a}_0), \quad V[\mathcal{O}(\vec{a})] \rightarrow \text{Hessian} \quad H_{ij} = \frac{1}{2} \frac{\partial^2 \chi^2(\vec{a})}{\partial a_i \partial a_j} \Big|_{\vec{a}=\vec{a}_0}$$

- In practice, since in general $E[f(\vec{a})] \neq f(E[\vec{a}])$, maximum likelihood method sometimes fails

→ need more robust (Monte Carlo) approach

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

Bayesian approach to global analysis

- First group to use MC for global PDF analysis was NNPDF, using neural network to parametrize $P(x)$ in

Ball, Forte et al. (2002)

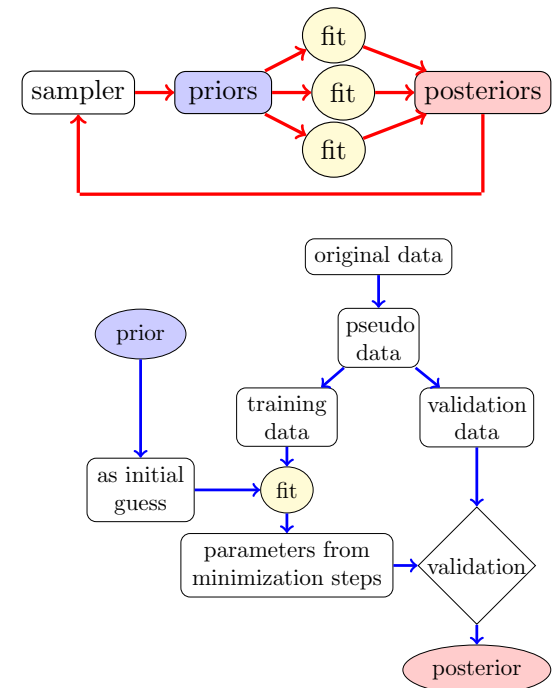
$$f(x) = N x^\alpha (1 - x)^\beta P(x)$$

- “unbiased”? not really... “pre-processing” coefficients...
there is no such thing as an unbiased PDF fit!

Accardi, WM, Nocera, Sato et al. (2019)

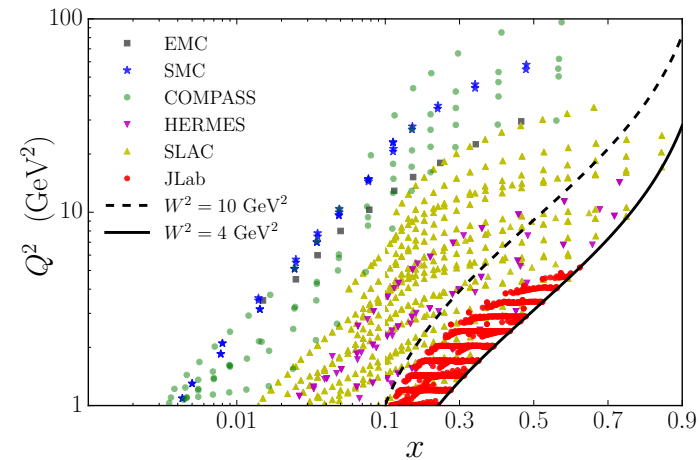
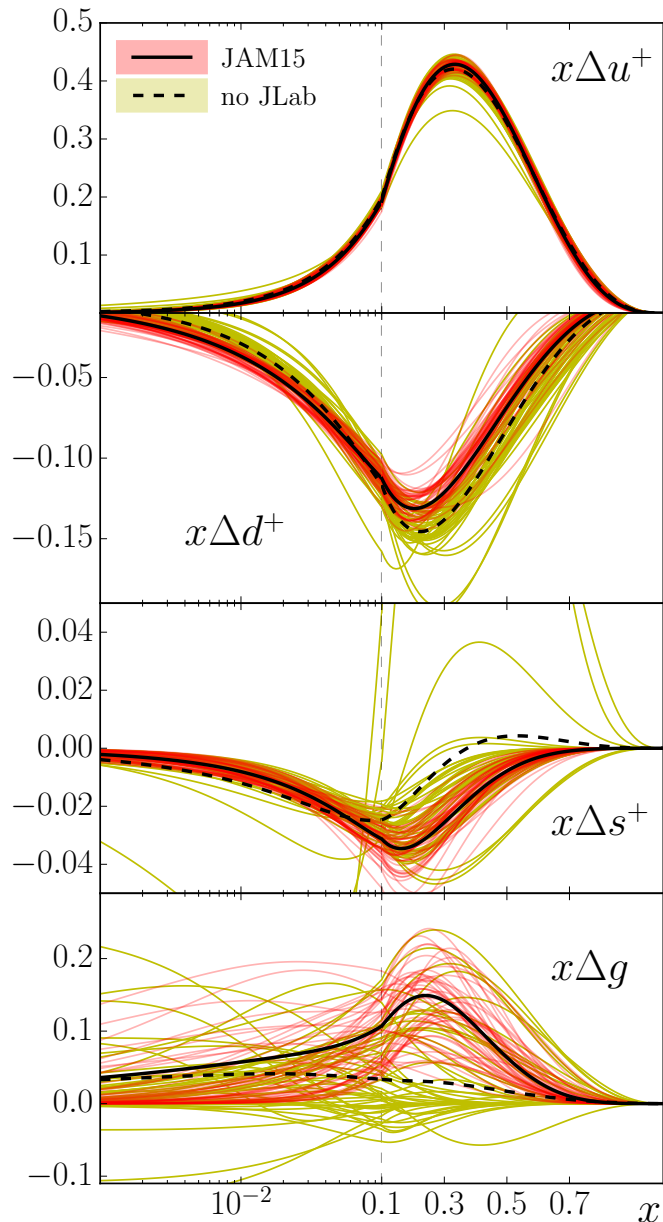
- JAM — iterative, multi-step Monte Carlo

- traditional functional form for distributions, but sample much larger parameter space
- no assumptions for exponents
- iterate until convergence (posteriors = priors)
- robust determination of PDF uncertainties



First application of IMC — spin structure

- First JAM MC analysis studied impact of JLab data on spin structure of the nucleon

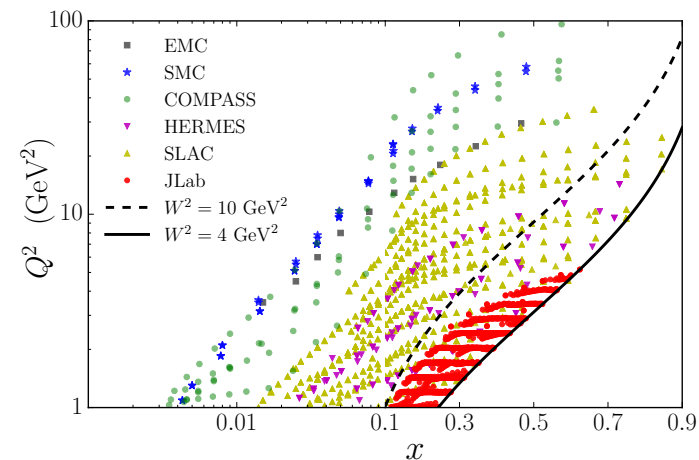
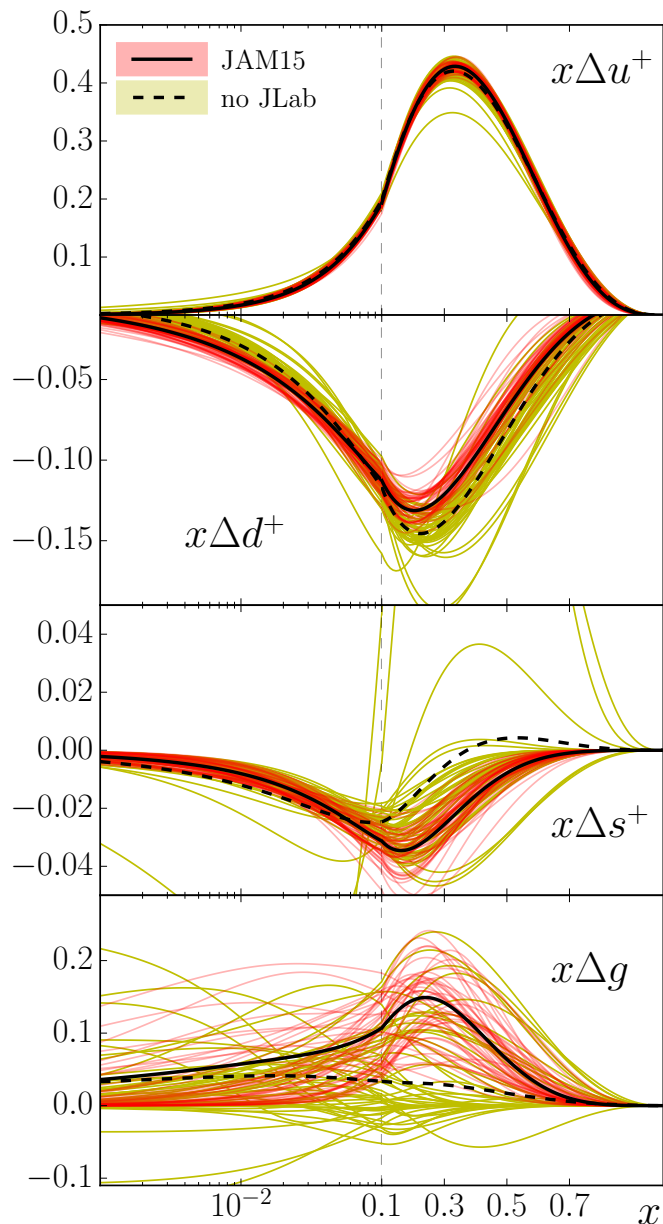


Sato, WM, Kuhn, Ethier, Accardi (2016)

- inclusion of JLab data increases # data points by factor ~ 2
- reduced uncertainty in Δs^+ , Δg through Q^2 evolution
- s -quark polarization *negative* from inclusive DIS data (assuming SU(3) symmetry)

First application of IMC — spin structure

- First JAM MC analysis studied impact of JLab data on spin structure of the nucleon



Sato, WM, Kuhn, Ethier, Accardi (2016)

- inclusion of JLab data increases # data points by factor ~ 2
- reduced uncertainty in Δs^+ , Δg through Q^2 evolution
- s -quark polarization *negative* from inclusive DIS data (assuming SU(3) symmetry)

First application of IMC — spin structure

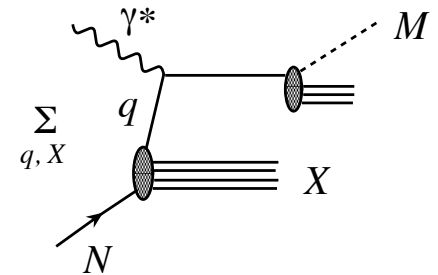
■ Inclusive DIS data cannot distinguish between q and \bar{q}

→ 2 observables (g_1^p, g_1^n) can determine up to 2 unknowns,
e.g. $\Delta u + \Delta \bar{u}$, $\Delta d + \Delta \bar{d}$ — sea quarks from Q^2 dependence

→ semi-inclusive DIS sensitive to Δq & $\Delta \bar{q}$

$$\sim \sum_q e_q^2 [\Delta q(x) D_q^h(z) + \Delta \bar{q}(x) D_{\bar{q}}^h(z)]$$

fragmentation functions



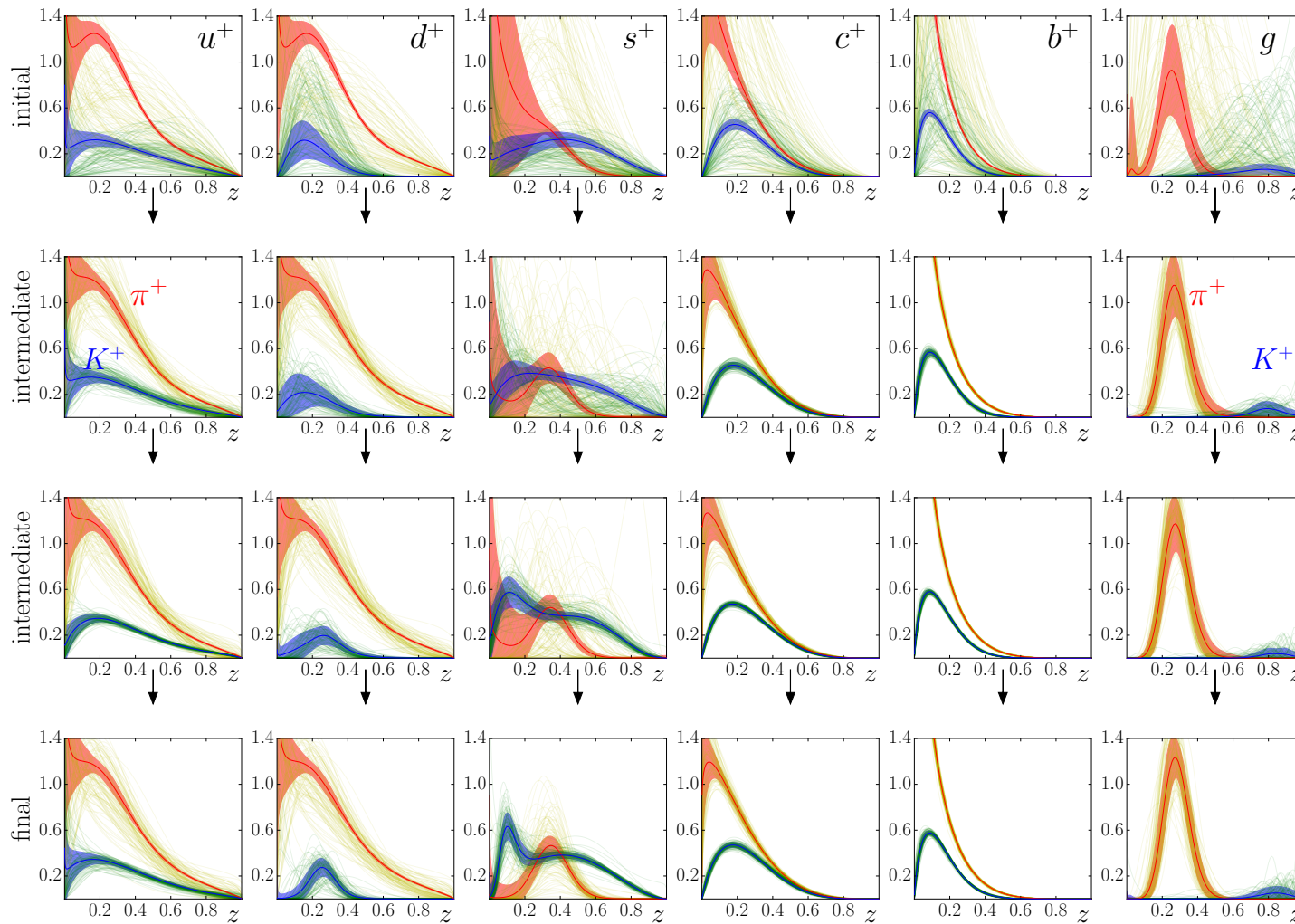
■ Global analysis of DIS + SIDIS data gives different *sign* for strange quark polarization for different fragmentation functions!

→ $\Delta s > 0$ for “DSS” FFs, but $\Delta s < 0$ for “HKNS” FFs

→ need to understand origin of differences in fragmentation functions

IMC analysis of fragmentation functions

- Analysis of single-inclusive e^+e^- annihilation data for π , K production (from DESY, CERN, SLAC & KEK) from $Q \sim 10$ GeV to M_Z



$e^+e^- \rightarrow h X$
single-inclusive
annihilation (SIA)

*Sato, Ethier, WM, Hirai,
Kumano, Accardi (2016)*

→ convergence after ~ 20 iterations

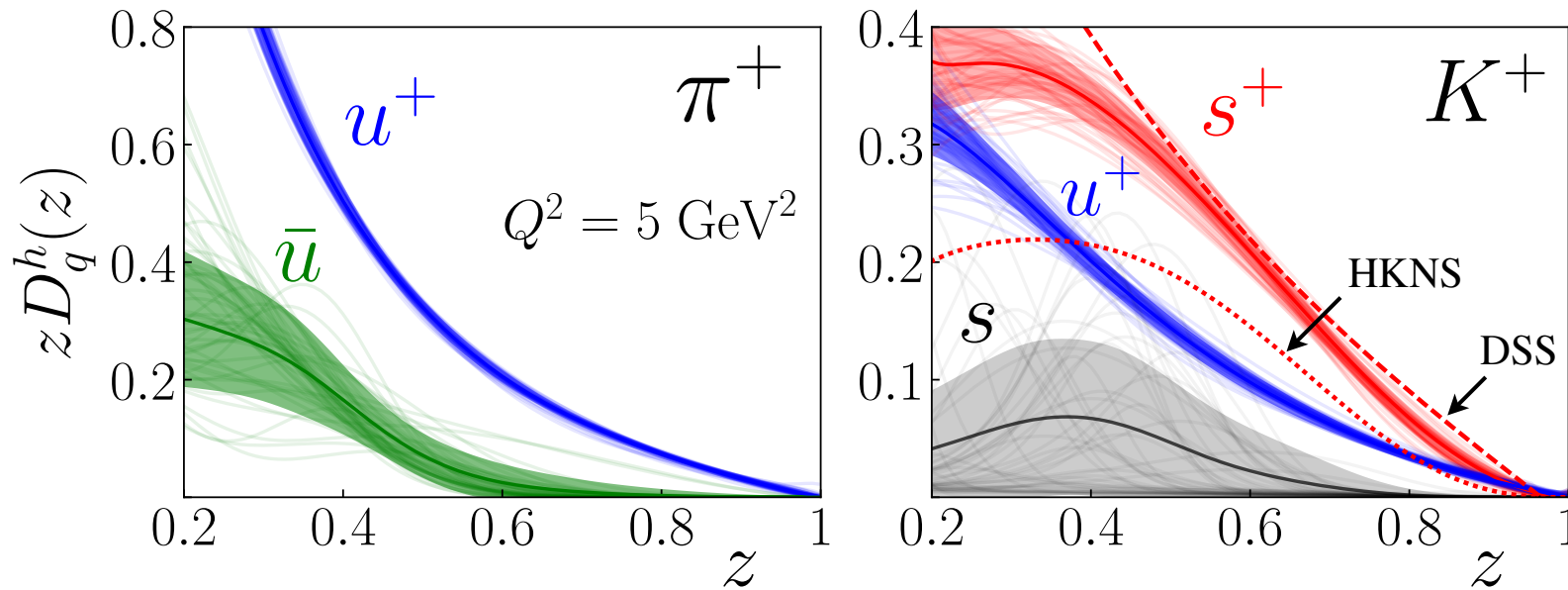
IMC analysis of fragmentation functions

- Analysis of single-inclusive e^+e^- annihilation data for π , K production (from DESY, CERN, SLAC & KEK) from $Q \sim 10$ GeV to M_Z

Experiment	Ref.	Observable	Q (GeV)	Pions			Kaons		
				N_{dat}	norm.	χ^2	N_{dat}	norm.	χ^2
ARGUS	[26]	Inclusive	9.98	35	1.024 (1.058)	51.1 (55.8)	15	1.007	8.5
Belle	[38,39]	Inclusive	10.52	78	0.900 (0.919)	37.6 (21.7)	78	0.988	10.9
BABAR	[40]	Inclusive	10.54	39	0.993 (0.948)	31.6 (70.7)	30	0.992	4.9
TASSO	[23–25]	Inclusive	12-44	29	(*)	37.0 (38.8)	18	(*)	14.3
TPC	[27–29]	Inclusive	29.00	18	1	36.3 (57.8)	16	1	47.8
		uds tag	29.00	6	1	3.7 (4.6)			
		b tag	29.00	6	1	8.7 (8.6)			
		c tag	29.00	6	1	3.3 (3.0)			
HRS	[30]	Inclusive	29.00	2	1	4.2 (6.2)	3	1	0.3
TOPAZ	[37]	Inclusive	58.00	4	1	4.8 (6.3)	3	1	0.9
OPAL	[32,33]	Inclusive	91.20	22	1	33.3 (37.2)	10	1	6.3
		u tag	91.20	5	1.203 (1.203)	6.6 (8.1)	5	1.185	2.1
		d tag	91.20	5	1.204 (1.203)	6.1 (7.6)	5	1.075	0.6
		s tag	91.20	5	1.126 (1.200)	14.4 (11.0)	5	1.173	1.5
		c tag	91.20	5	1.174 (1.323)	10.7 (6.1)	5	1.169	13.2
		b tag	91.20	5	1.218 (1.209)	34.2 (36.6)	4	1.177	10.9
		Inclusive	91.20	22	0.987 (0.989)	15.6 (20.4)	18	1.008	6.1
ALEPH	[34]	Inclusive	91.20	22	0.987 (0.989)	15.6 (20.4)	18	1.008	6.1
DELPHI	[35,36]	Inclusive	91.20	17	1	21.0 (20.2)	27	1	3.9
		uds tag	91.20	17	1	13.3 (13.4)	17	1	22.5
		b tag	91.20	17	1	41.9 (42.9)	17	1	9.1
SLD	[31]	Inclusive	91.28	29	1.002 (1.004)	27.3 (36.3)	29	0.994	14.3
		uds tag	91.28	29	1.003 (1.004)	51.7 (55.6)	29	0.994	42.6
		c tag	91.28	29	0.998 (1.001)	30.2 (40.4)	29	1.000	31.7
		b tag	91.28	29	1.005 (1.005)	74.6 (61.9)	28	0.992	134.1
Total:				459		599.3 (671.2)	391		395.0
				$\chi^2/N_{\text{dat}} = 1.31$ (1.46)			$\chi^2/N_{\text{dat}} = 1.01$		

IMC analysis of fragmentation functions

- Analysis of single-inclusive e^+e^- annihilation data for π , K production (from DESY, CERN, SLAC & KEK) from $Q \sim 10$ GeV to M_Z

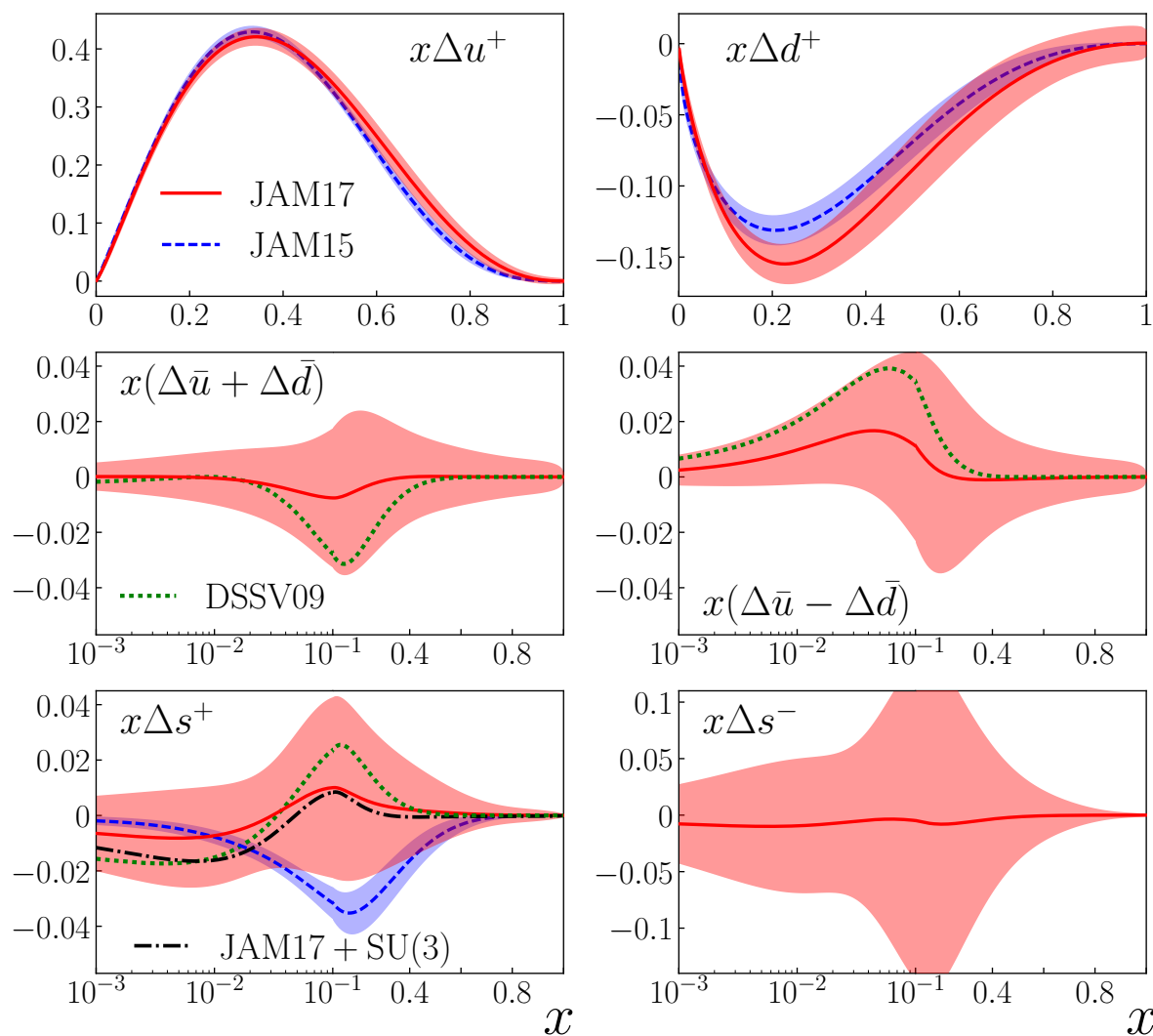


Ethier, Sato, WM (2017)

- favored $u^+ = u + \bar{u}$ & $s^+ = s + \bar{s}$ FFs well constrained
- larger $s \rightarrow K$ fragmentation *cf.* HKNS suggests less negative Δ_s

Simultaneous spin PDF + FF analysis

- First simultaneous extraction of spin PDFs and FFs, fitting polarized DIS + SIDIS (HERMES, COMPASS) and SIA data



→ no assumption of SU(3) symmetry

$$\Delta u^+ - \Delta d^+ \stackrel{?}{=} g_A \equiv F + D$$

$$\Delta u^+ + \Delta d^+ - 2\Delta s^+ \stackrel{?}{=} a_8 \equiv 3F - D$$

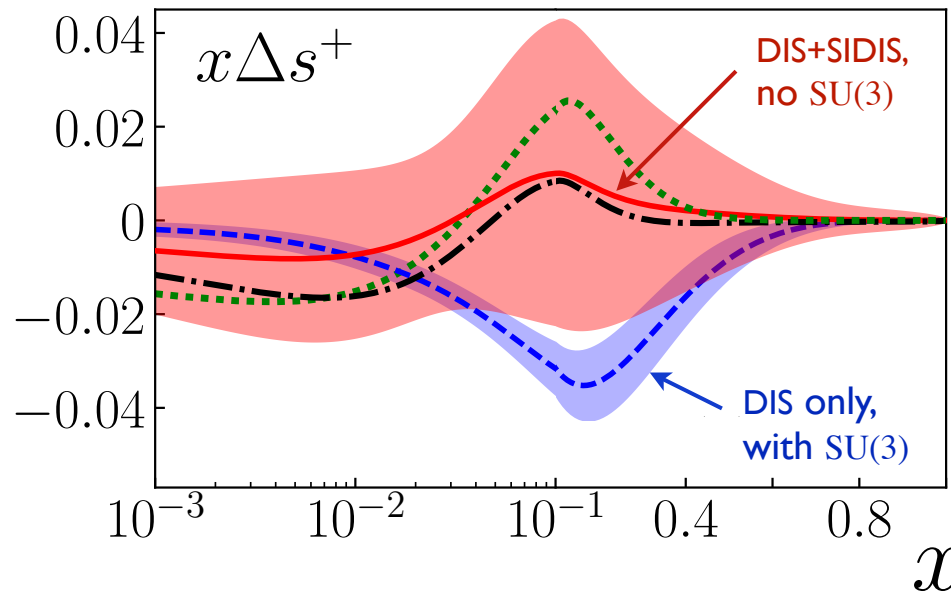
→ Δs slightly > 0 at high x , consistent with zero

→ $\Delta s - \Delta \bar{s}$ & $\Delta \bar{u} - \Delta \bar{d}$ consistent with zero

Ethier, Sato, WM (2017)

Simultaneous spin PDF + FF analysis

- Polarized strangeness in previous, DIS-only analyses was negative at $x \sim 0.1$, induced by SU(3) and parametrization bias

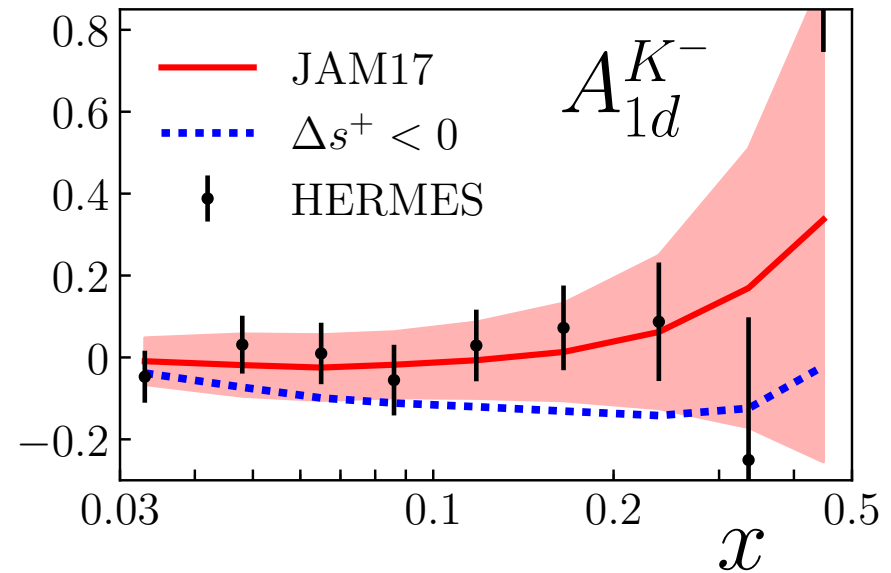
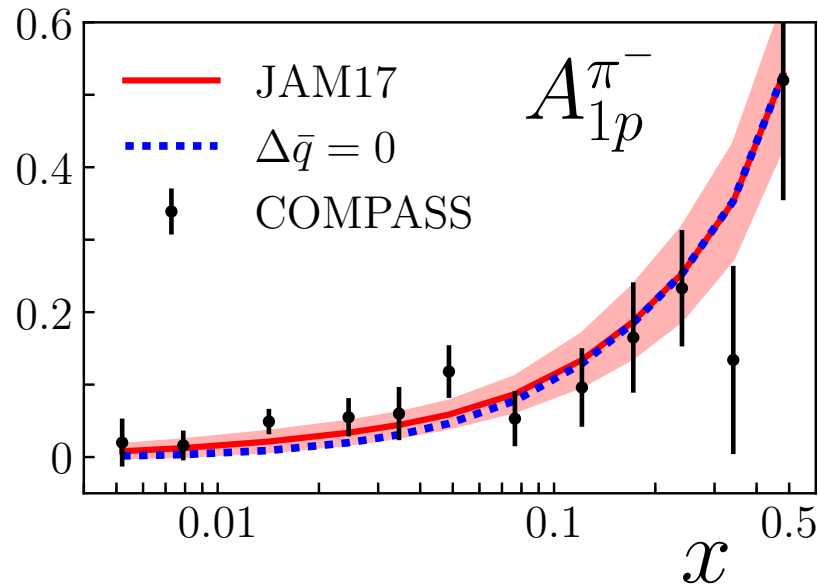


Ethier, Sato, WM (2017)

- weak sensitivity to Δs^+ from DIS data & evolution
- SU(3) pulls Δs^+ to generate moment ~ -0.1
- negative peak at $x \sim 0.1$ induced by fixing $b \sim 6 - 8$

Simultaneous spin PDF + FF analysis

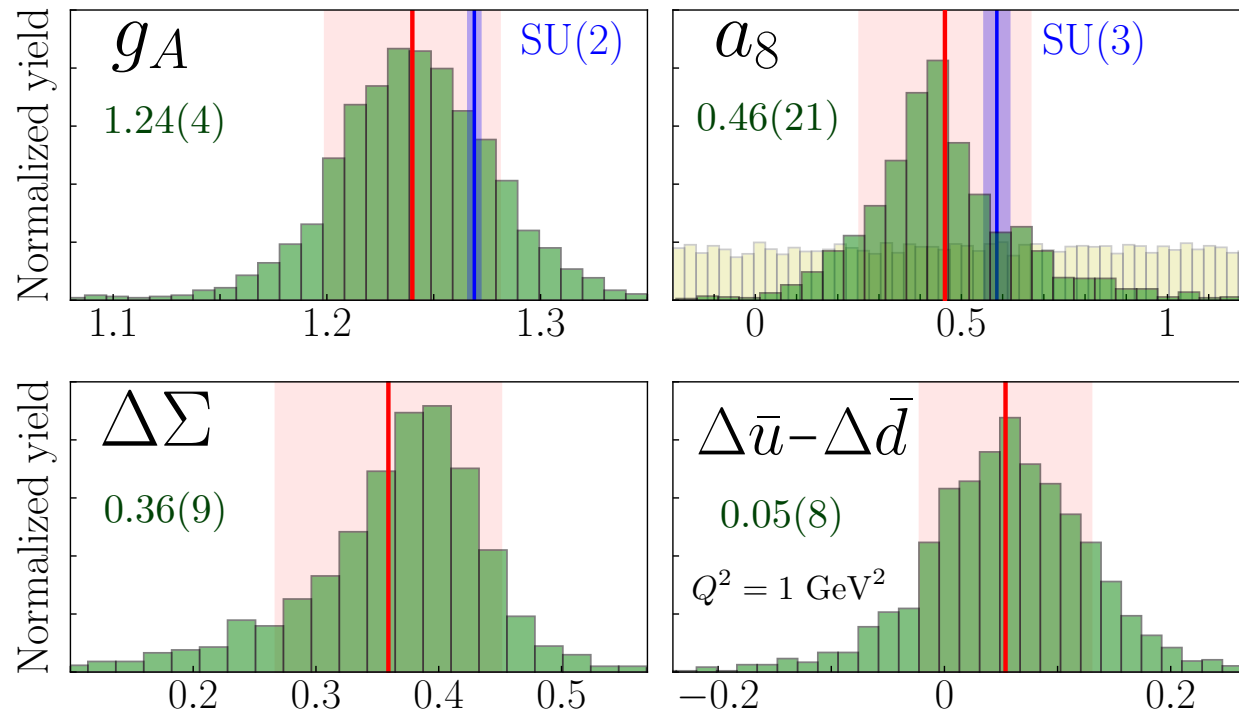
- SIDIS data, especially for K production, clearly prefer a less negative Δ_s



Ethier, Sato, WM (2017)

Simultaneous spin PDF + FF analysis

■ Statistical distribution of lowest moments (axial charges)



Ethier, Sato, WM (2017)

- triplet charge g_A consistent with SU(2) value
- hint of SU(3) breaking in octet charge a_8 *Bass, Thomas (2010)*
- less negative $\Delta s = -0.03(10)$ gives larger total helicity $\Delta\Sigma = 0.36(9)$

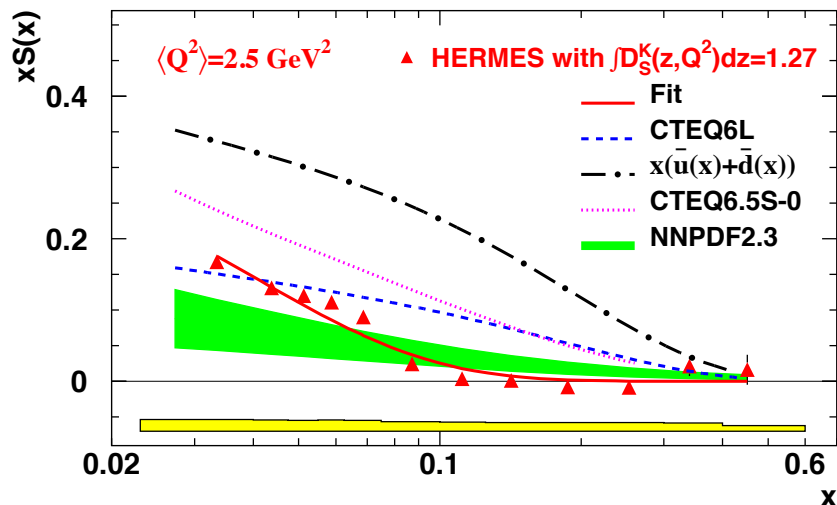
Simultaneous spin PDF + FF analysis

- What impact does unpolarized strange PDF have on the extraction of polarized strange?

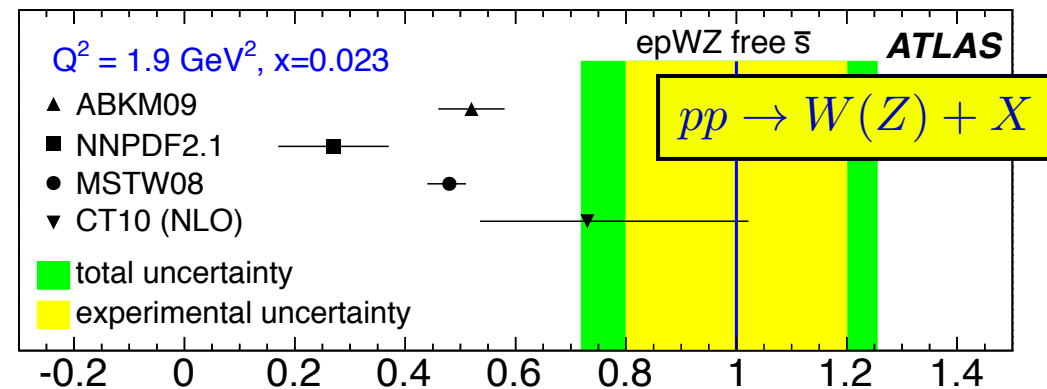
→ only systematic way is to fit unpolarized PDFs, polarized PDFs and fragmentation functions simultaneously...

- Shape of unpolarized strange PDF is interesting (and controversial) in its own right!

→ historically, strange to nonstrange ratio $R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}} \sim 0.4$



PRD 89 (2014) 097101



PRL 109 (2012) 012001

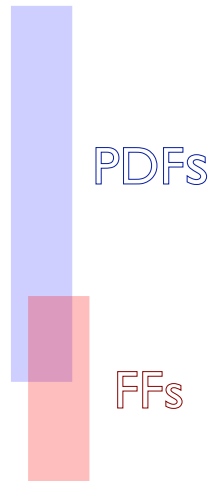
$$r_s = (s + \bar{s}) / 2\bar{d}$$

$$= 1.00^{+0.25}_{-0.28}$$

JAM 2019 analysis

■ Study the impact of SIDIS data on unpolarized PDFs

- unpolarized fixed-target DIS on p , d (SLAC, BCDMS, NMC), HERA collider data (runs I & II)
- Drell-Yan (Fermilab E866), jet production (CDF, D0)
- SIDIS pion & kaon multiplicities for deuteron (COMPASS)
- e^+e^- annihilation (DESY, LEP/CERN, SLAC, KEK)

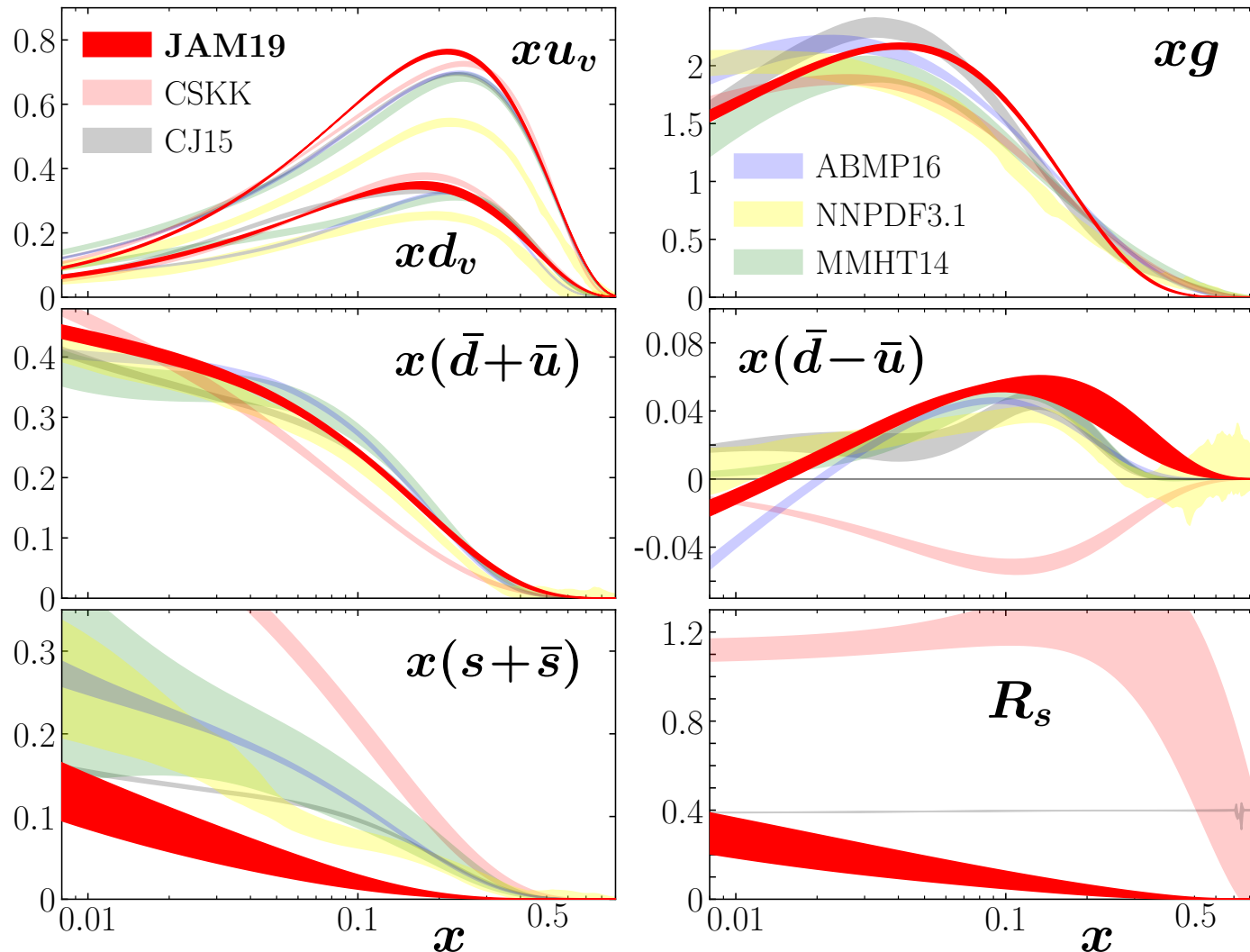


■ 52 shape parameters + 41 “nuisance” parameters for systematic uncertainties (data normalizations)

■ 953 fits to 4366 data points (2680 DIS, 992 SIDIS, 250 DY, 444 SIA)

- such an analysis has never been attempted before...

JAM 2019 analysis

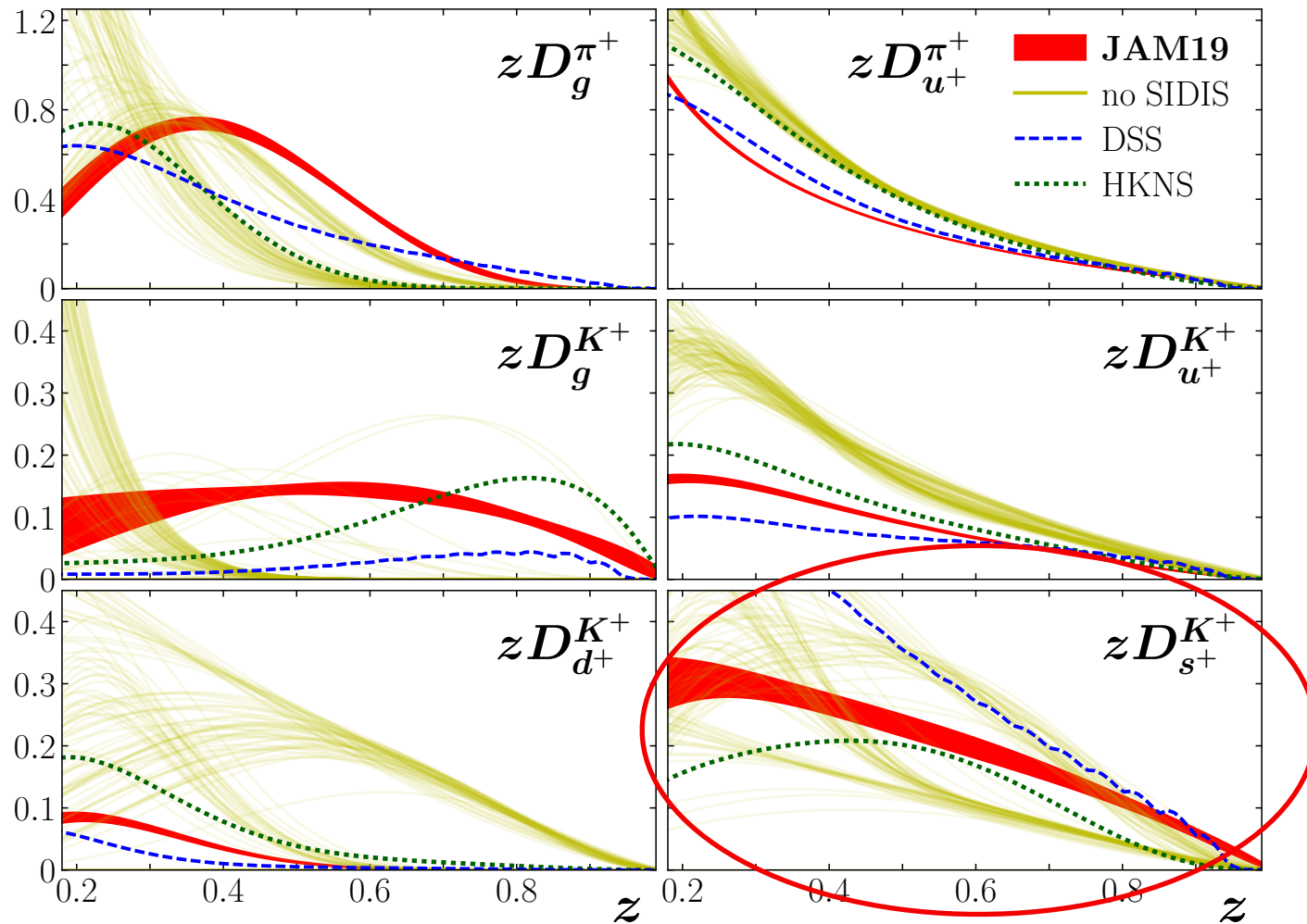


mean reduced
 $\chi^2 = 1.3$
for all data

Sato, Andres, Ethier, WM (2019)

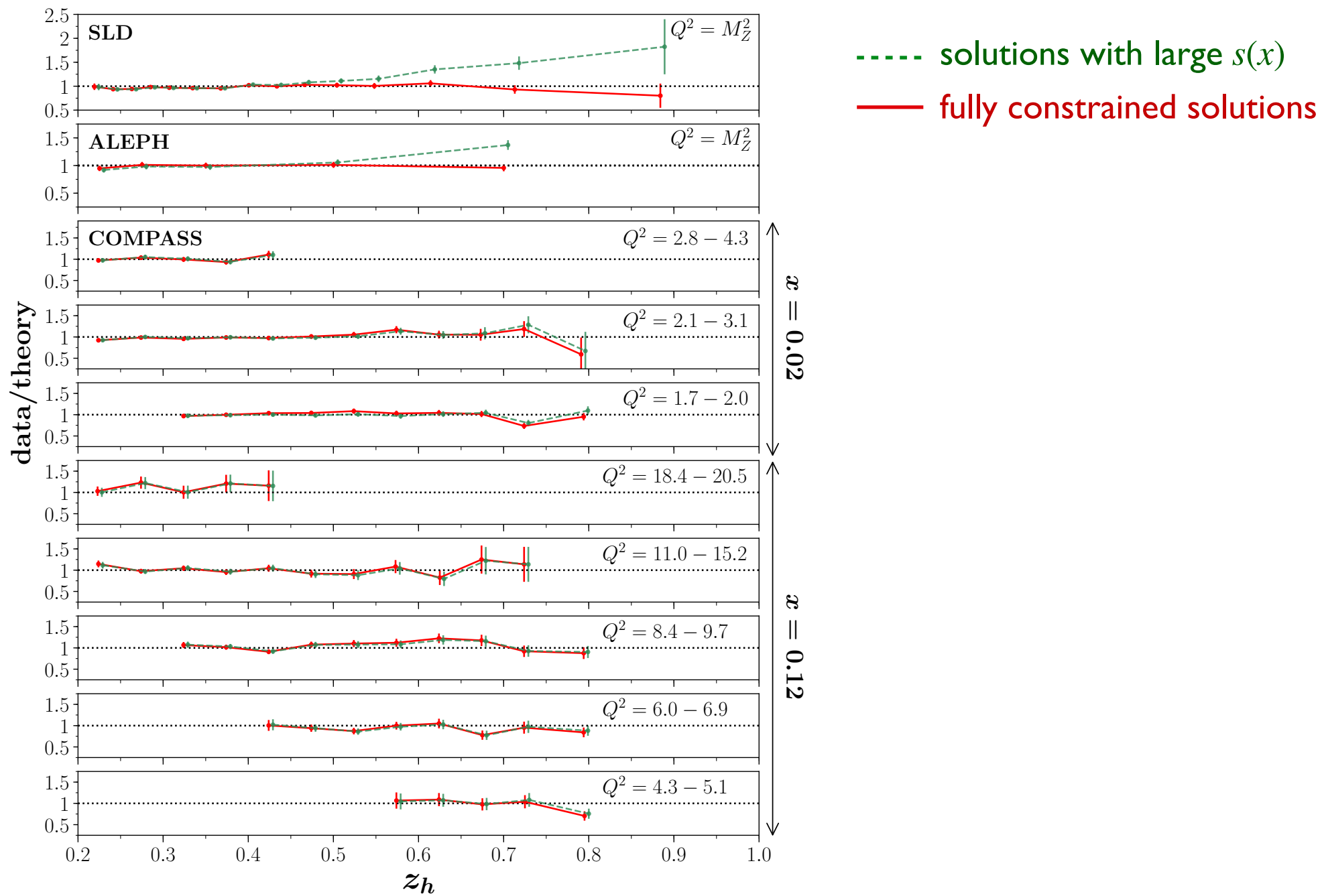
- valence & light sea quark broadly in agreement with other groups
- striking suppression of strange PDF compared to ATLAS extraction

JAM 2019 analysis

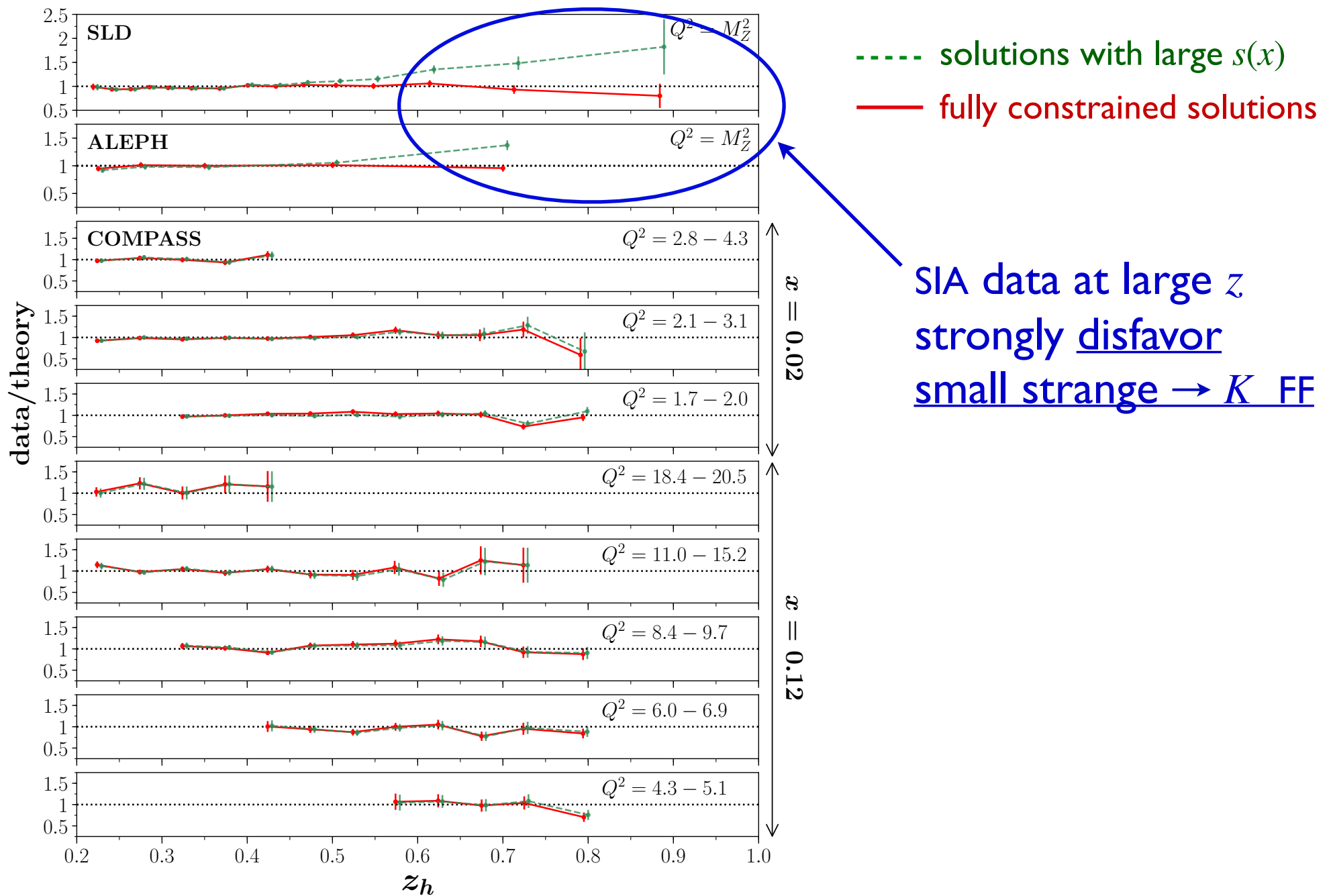


→ SIDIS + SIA data force strange to kaon FF to be larger

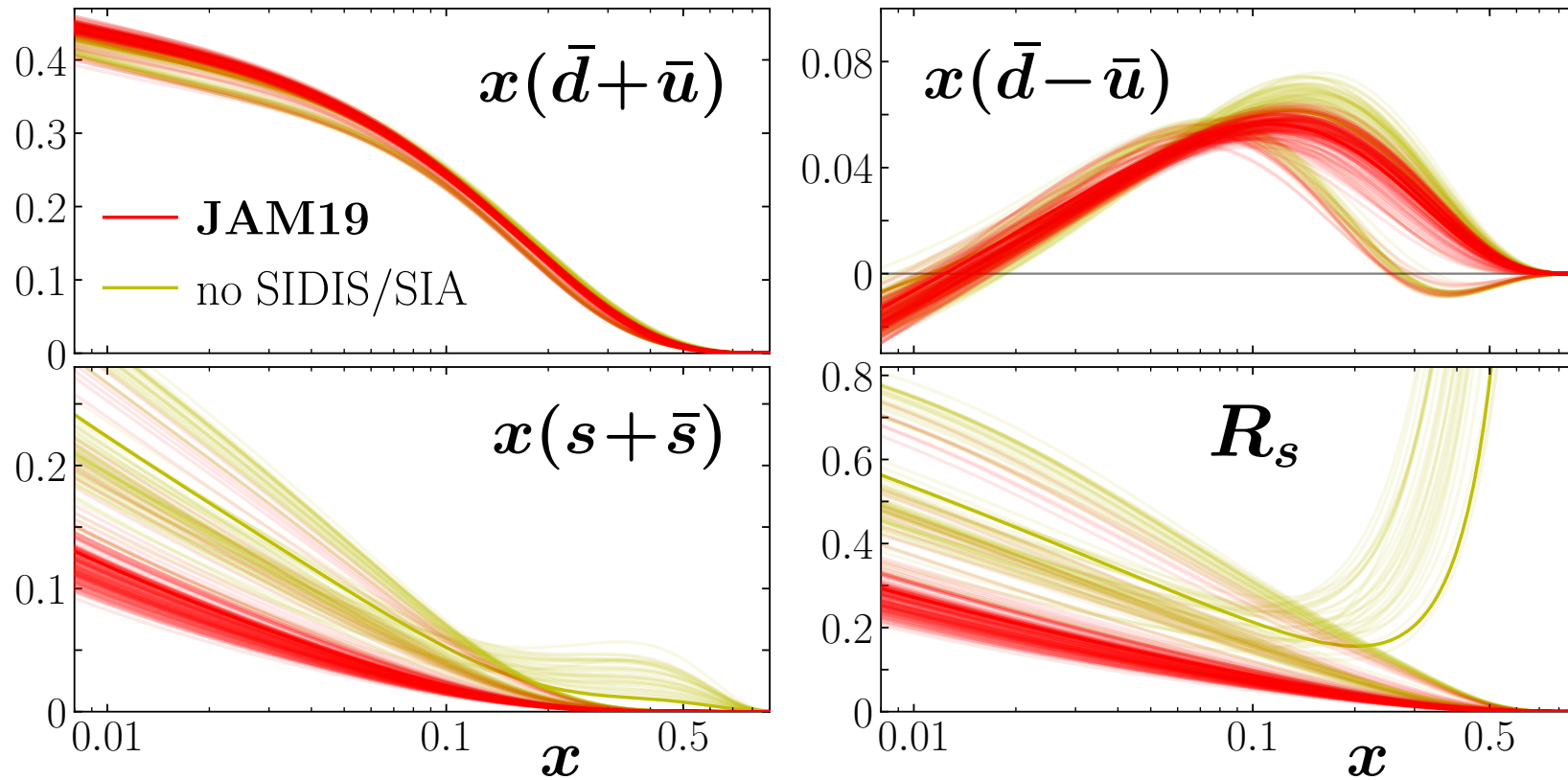
JAM 2019 analysis



JAM 2019 analysis



JAM 2019 analysis



→ vital role played by SIDIS + SIA data in constraining strange PDF

PDFs in lattice QCD

- Recent progress in extracting x dependence of PDFs in lattice QCD from matrix element of nonlocal operator

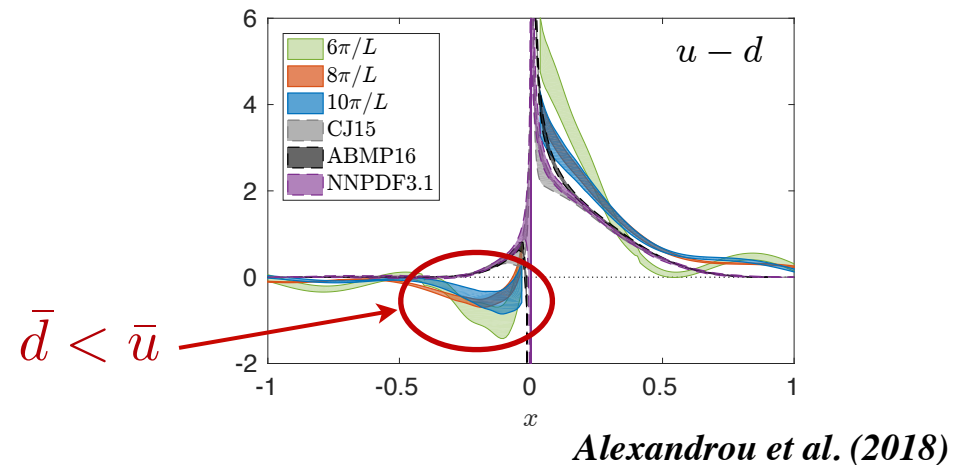
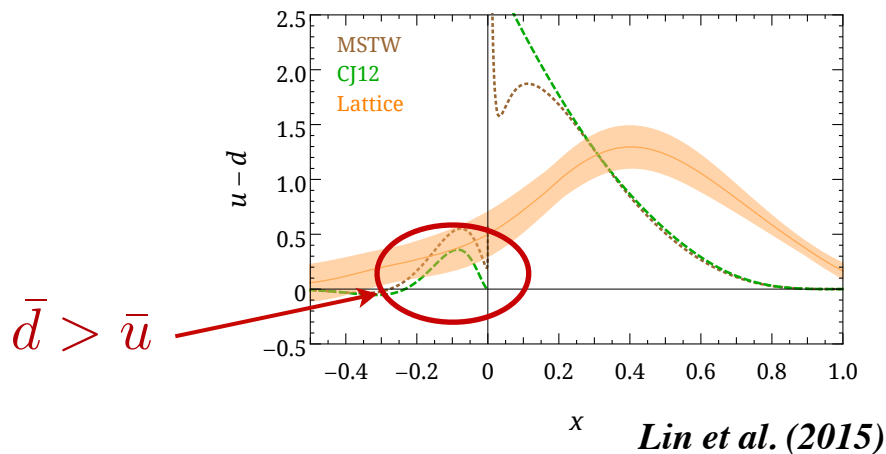
$$h(z, P_z) = \langle P | \bar{\psi}(0, z) \gamma_z \mathcal{W}(z, 0) \psi(0, 0) | P \rangle$$

$$= \int_{-\infty}^{\infty} dy e^{iyP_z z} \tilde{q}(y, P_z)$$

→ quasi-PDF \tilde{q} related to light-cone PDF via matching kernel \tilde{C}

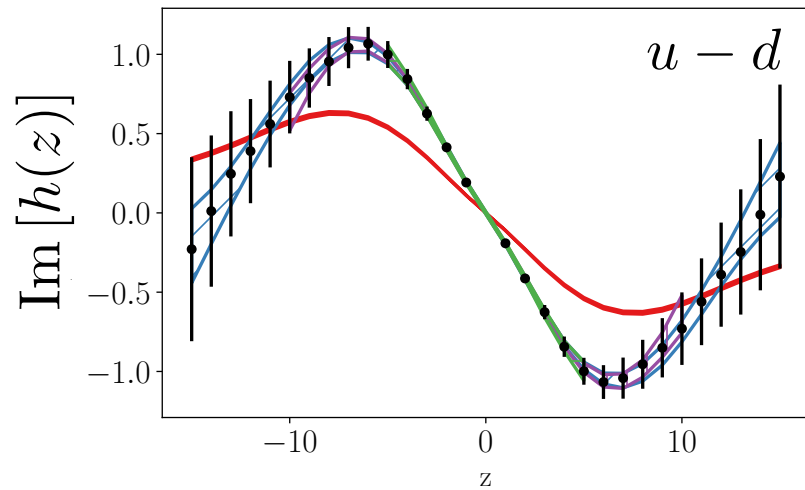
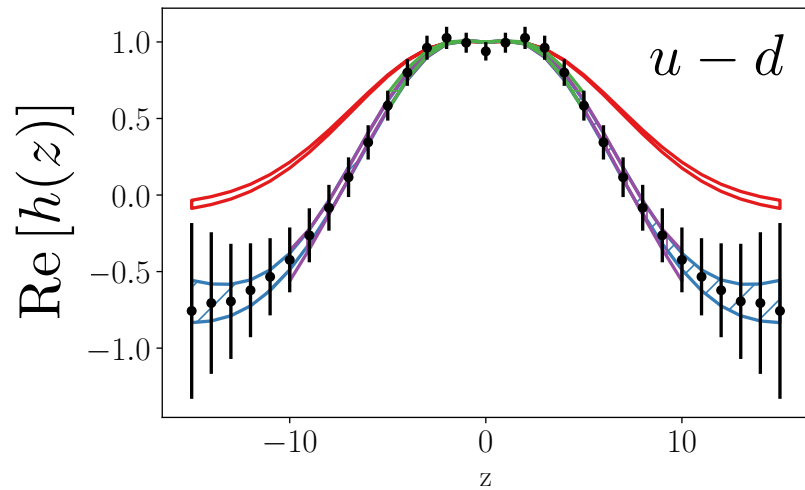
$$q(x, \mu) = \int_{-\infty}^{\infty} \frac{dy}{|y|} \tilde{C}\left(\frac{x}{y}, \mu, P_z\right) \tilde{q}(y, P_z, \mu)$$

- Conflicting results on sign of $\bar{d} - \bar{u}$ asymmetry

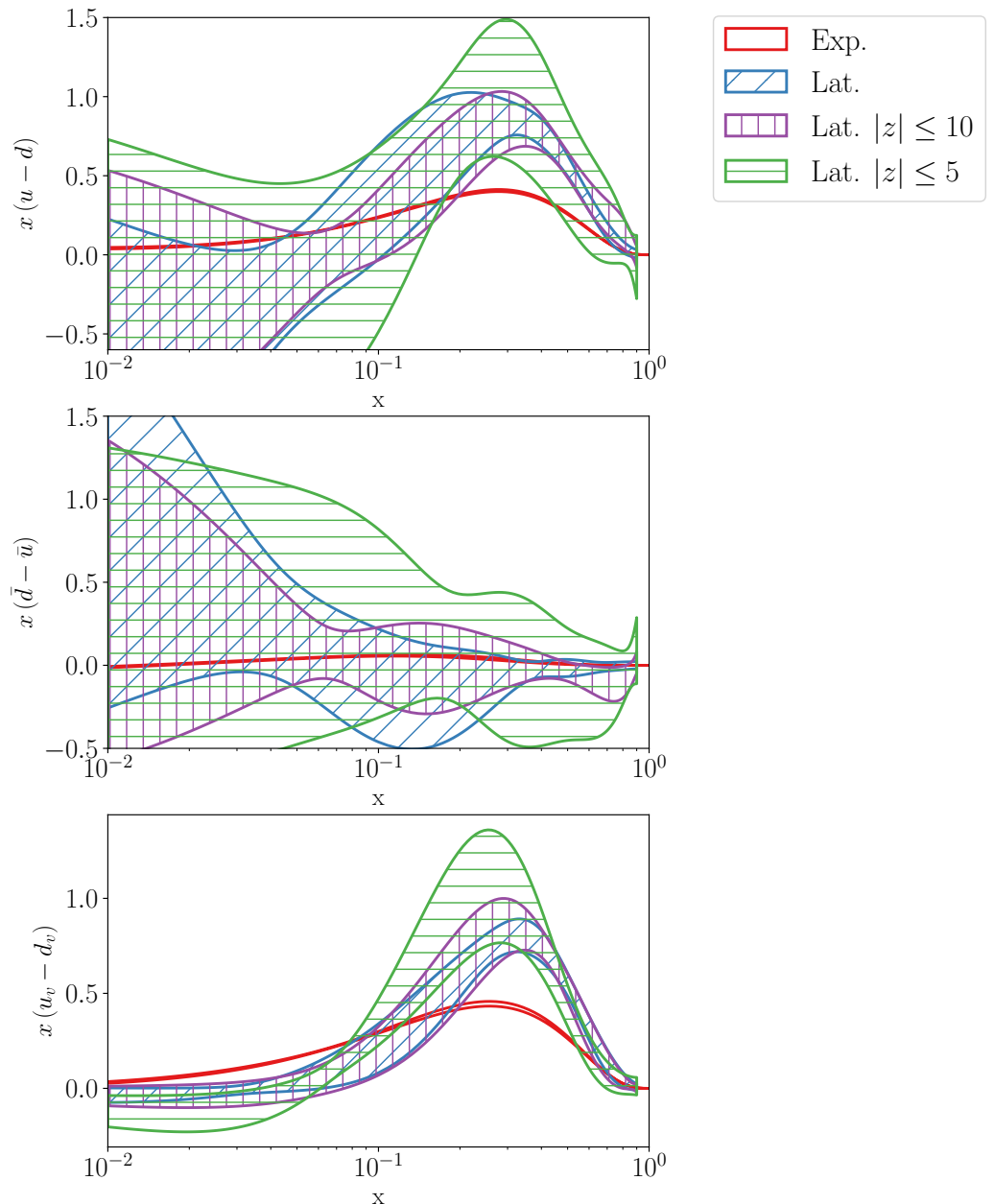


PDFs in lattice QCD

- Fit lattice observable directly within JAM framework

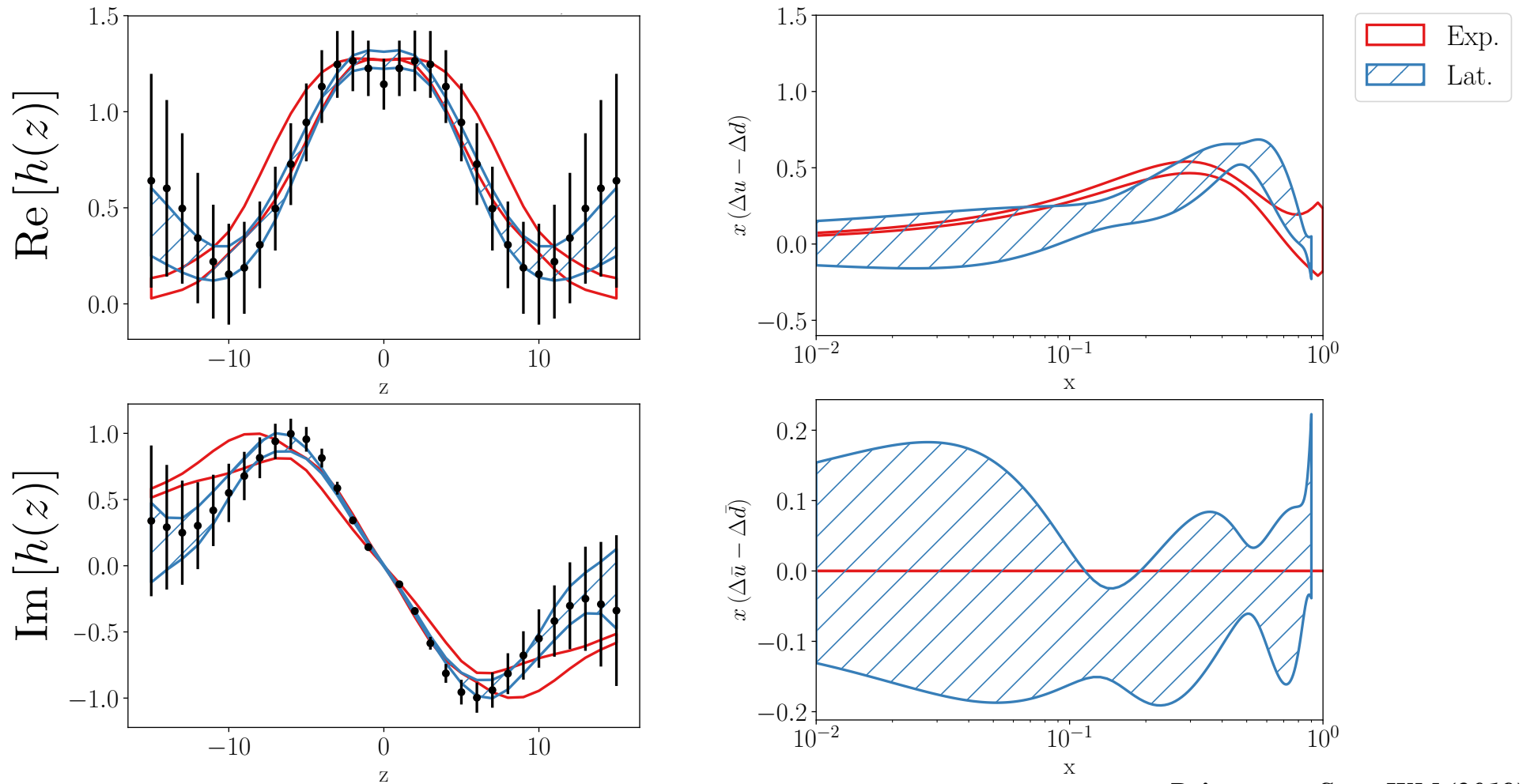


→ cannot determine $\bar{d} - \bar{u}$
from present lattice data



PDFs in lattice QCD

- Fit lattice observable directly within JAM framework

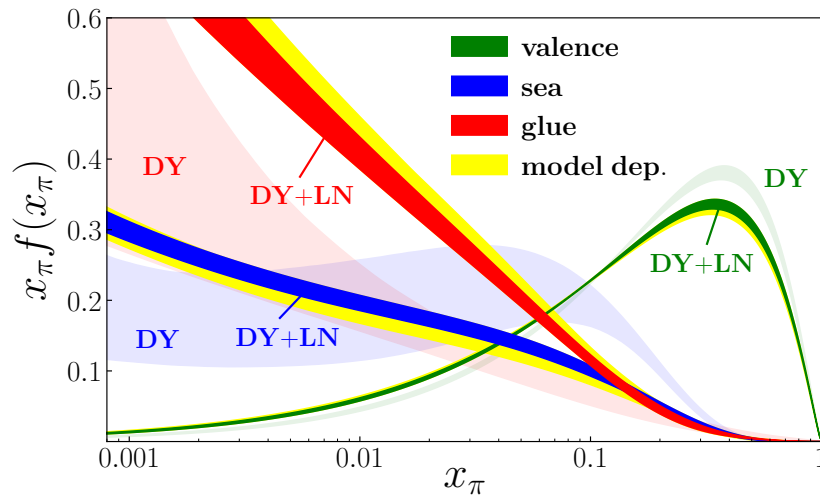


Bringewatt, Sato, WM (2019)

- better agreement between lattice and experiment for polarized PDFs (within larger uncertainties)

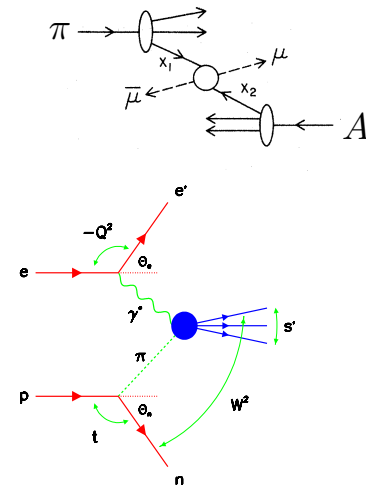
PDFs in the pion

- MC analysis combining pQCD with chiral EFT to fit πN Drell-Yan + leading neutron electroproduction data from HERA

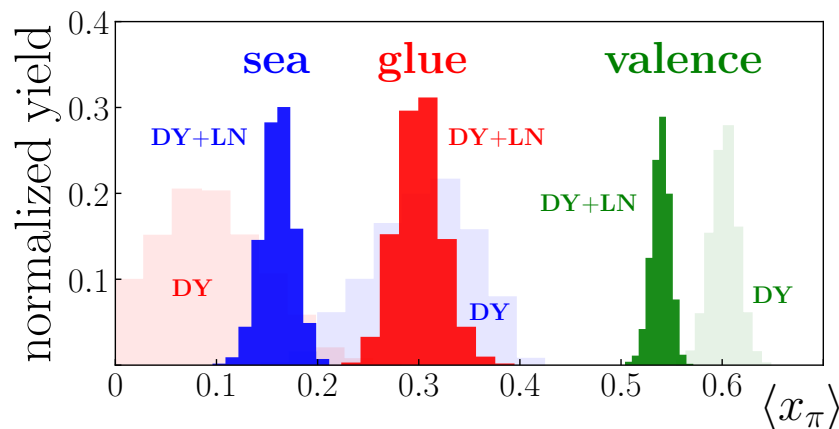


DY = πN Drell-Yan data
(medium/high x)

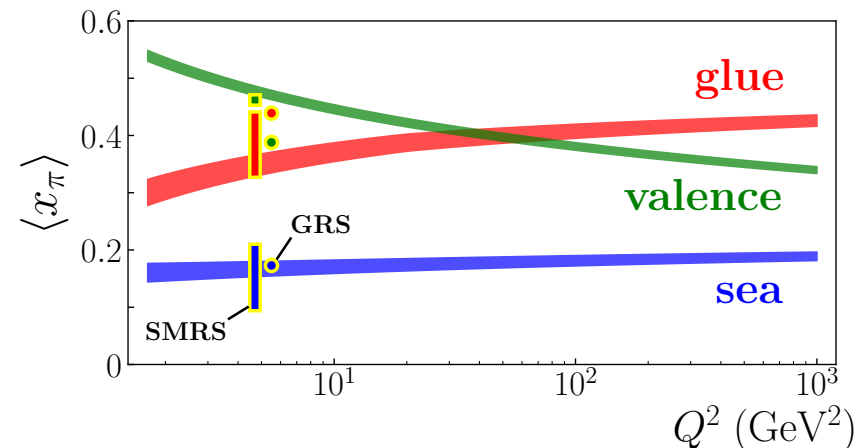
LN = leading neutron
electroproduction
(low x)



- Larger gluon fraction in the pion than without LN constraint

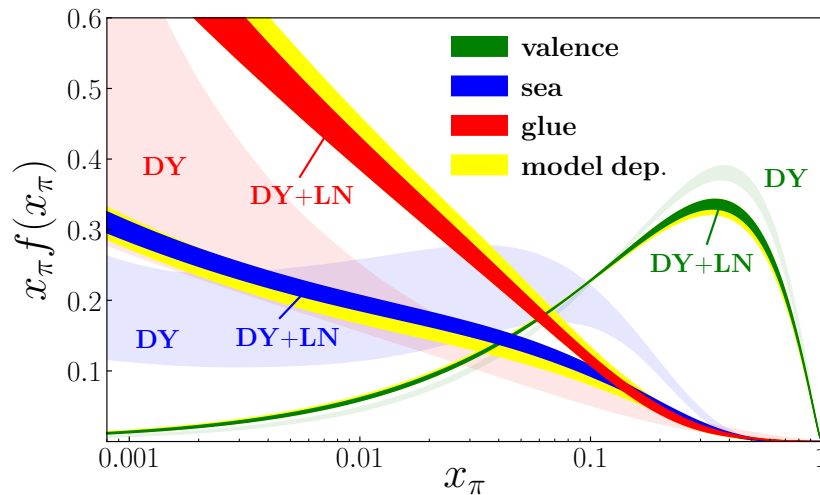


Barry, Sato, WM, C.-R. Ji (2018)



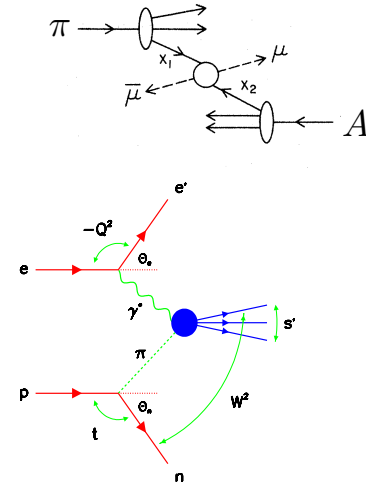
PDFs in the pion

- MC analysis combining pQCD with chiral EFT to fit πN Drell-Yan + leading neutron electroproduction data from HERA



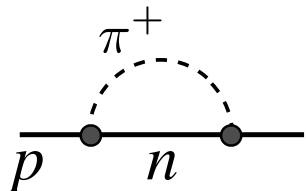
DY = πN Drell-Yan data
(medium/high x)

LN = leading neutron
electroproduction
(low x)

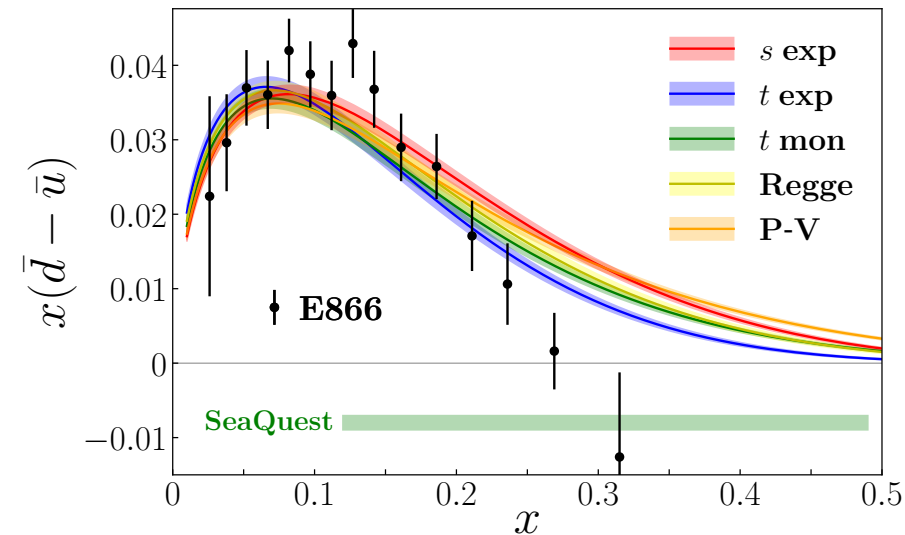


- Provides new insights into the origin of the $\bar{d} - \bar{u}$ asymmetry in the proton

→ chiral effective theory relates asymmetry to structure of pion



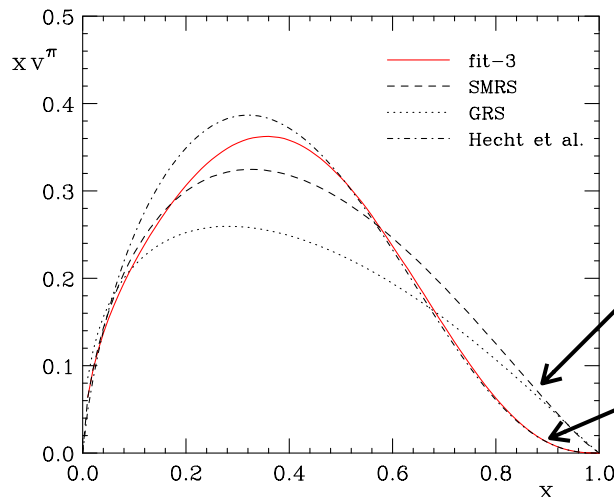
$$(\bar{d} - \bar{u})(x) = \int \frac{dy}{y} f_{\pi+n}(y) \bar{q}^\pi(x/y)$$



Barry, Sato, WM, C.-R. Ji (2018)

PDFs in the pion

- $x \rightarrow 1$ behavior of pion PDF is controversial: $\sim (1-x)$ or $(1-x)^2$?

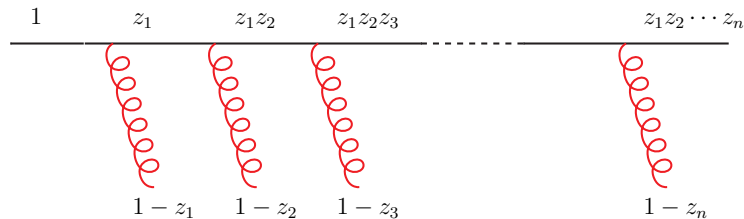


Aicher, Schafer, Vogelsang (2010)

no resummation: more consistent with $\sim (1-x)$

with resummation: more consistent with $\sim (1-x)^2$

- Hard scattering coefficient function kinematically enhanced when $z \rightarrow 1$ because of gluon emissions



Patrick Barry et al. (2019)

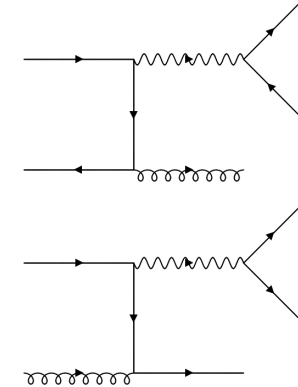
→ effect of resummation on phenomenology?

PDFs in the pion

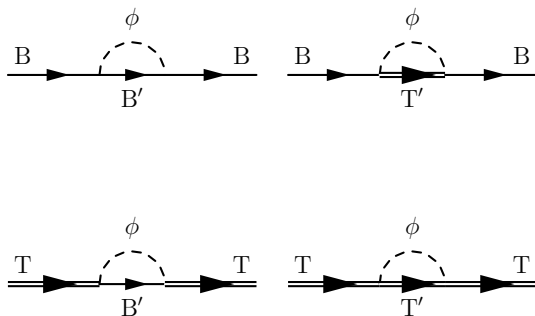
- New analysis examines whether large- q_T DY data can be simultaneously described with q_T -integrated DY + HERA LN data

→ large- q_T photon requires hard gluon to recoil against — sensitivity to gluon PDF in pion at large x !

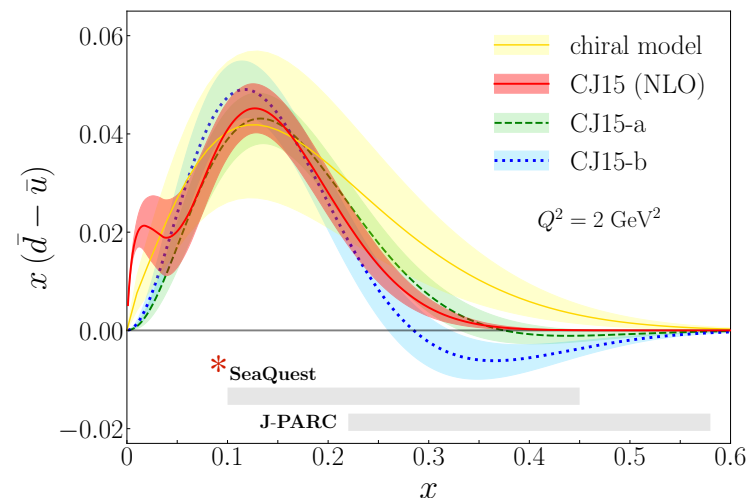
Nina Cao et al. (2019)



- Generalization of chiral model to SU(3) octet & decuplet baryons



Marston Copeland et al. (2019)



* expected Oct. 2019

Outlook

- New paradigm in global analysis — simultaneous determination of collinear distributions using MC sampling of parameter space
- **Next steps:** simultaneous analysis of all collinear distributions — unpolarized & polarized PDFs and FFs (including jet, W production, ... data)
- **Longer-term:** technology developed here will be applied to global QCD analysis of transverse momentum dependent (TMD) distributions — map out full 3-d image of hadrons