Novel Probes of Nucleon Structure in SIDIS, e+e− & pp **(*FF2019*)**
Duke University, March 14, 2019

Simultaneous analysis of PDFs and fragmentation functions — JAM19 —

Wally Melnitchouk

JLab Angular Momentum collaboration

https://www.jlab.org/jam

Overview

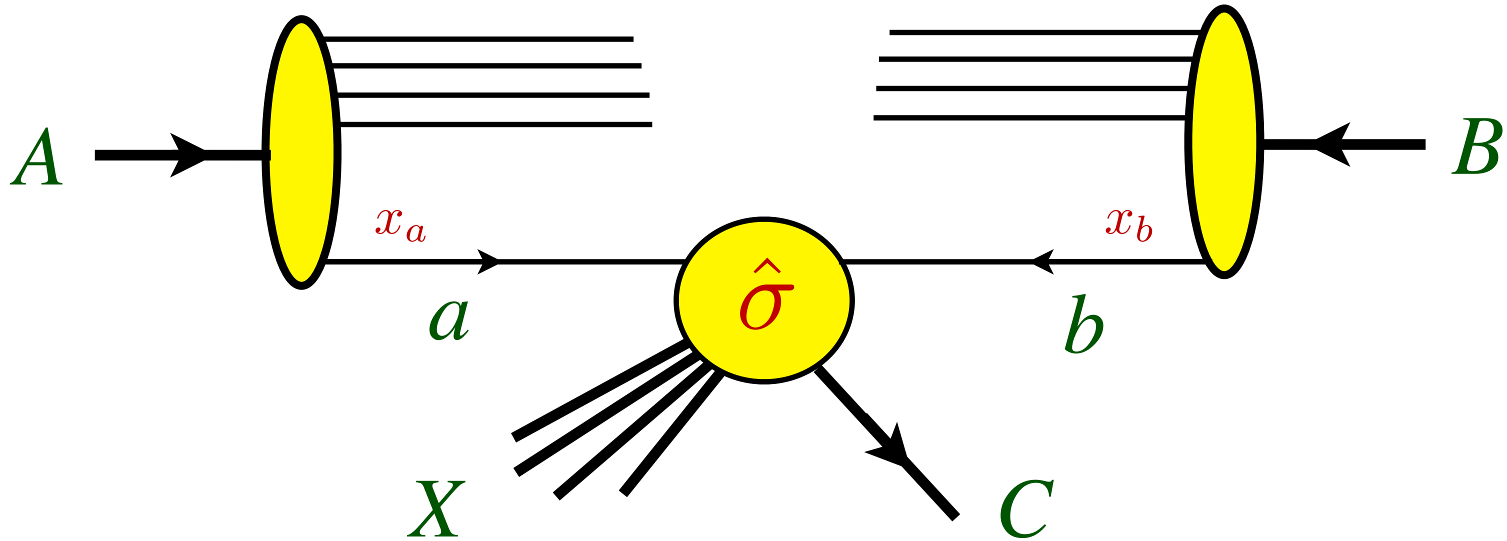
- JAM (Jefferson Lab Angular Momentum) collaboration aims to study the quark-gluon structure of hadrons through extraction of quantum probability distributions (PDFs, FFs, TMDs) via *global QCD analysis* using *Monte Carlo* based methods

- Methodology based on *Bayesian* statistics and Monte Carlo sampling of the parameter space

- Inter-dependence of observables on distributions requires *simultaneous* extraction of PDFs & fragmentation functions

Parton distributions in hadrons

- Generic process: inclusive particle production $A\,B \to C\,X$



$$\sigma_{AB\to CX}(p_A, p_B) = \sum_{a,b} \int dx_a\, dx_b\, f_{a/A}(x_a, \mu)\, f_{b/B}(x_b, \mu)$$

$$\times \sum_n \alpha_s^n(\mu)\, \hat{\sigma}^{(n)}_{ab\to CX}\,(x_a p_A, x_b p_B, Q/\mu)$$

$\longrightarrow$ extraction of functions $f_{a/A}$ characterizing structure of bound state A is a typical *"inverse problem"*

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

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

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

"Bayesian master formulas"

- Using Bayes' theorem, *probability distribution* $\mathcal{P}$ given by

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

in terms of the *likelihood function* $\mathcal{L}$

Bayesian approach to global analysis

- Likelihood function

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

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

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

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

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

$\longrightarrow$ 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

  $\longrightarrow$ maximize probability distribution

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

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

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

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

  $\longrightarrow$ need more robust (Monte Carlo) approach

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

Monte Carlo methods

- First group to use MC for global PDF analysis was NNPDF, using neural network to parametrize $P(x)$ in *Forte et al. (2002)*

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

$\longrightarrow$ ***E. Nocera talk***

- Iterative Monte Carlo (IMC), developed by JAM Collaboration, variant of NNPDF, tailored to non-neutral net parametrizations

N. Sato et al. (2016)

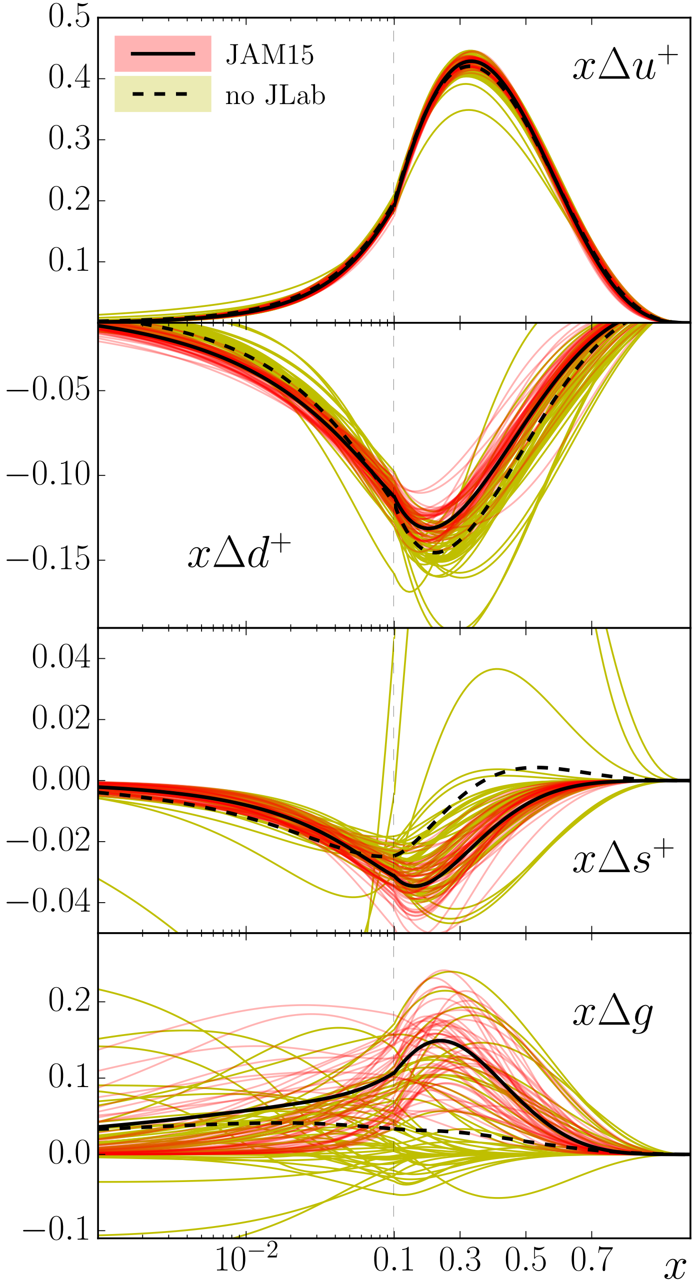
- Markov Chain MC (MCMC) / Hybid MC (HMC) — recent "proof of principle" analysis, ideas from lattice QCD

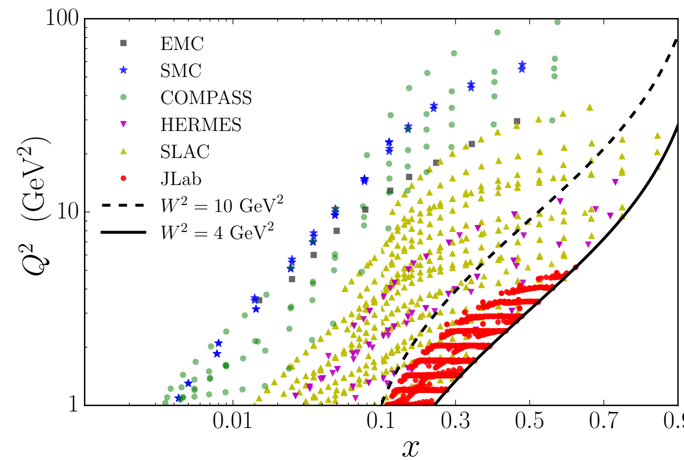
Gbedo, Mangin-Brinet (2017)

- Nested sampling (NS) — computes integrals in Bayesian master formulas (for $E,\ V,\ Z$) explicitly *Skilling (2004)*

Previous JAM analyses

- First application of IMC — proton spin structure





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

→ inclusion of JLab data increases # data points by factor ~ 2

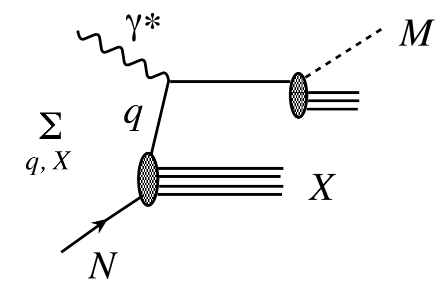
→ reduced uncertainty in Δs^+, Δg through Q^2 evolution

→ s-quark polarization *negative* from inclusive DIS data (assuming SU(3) symmetry)

Previous JAM analyses

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

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



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

  → but need fragmentation functions…

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

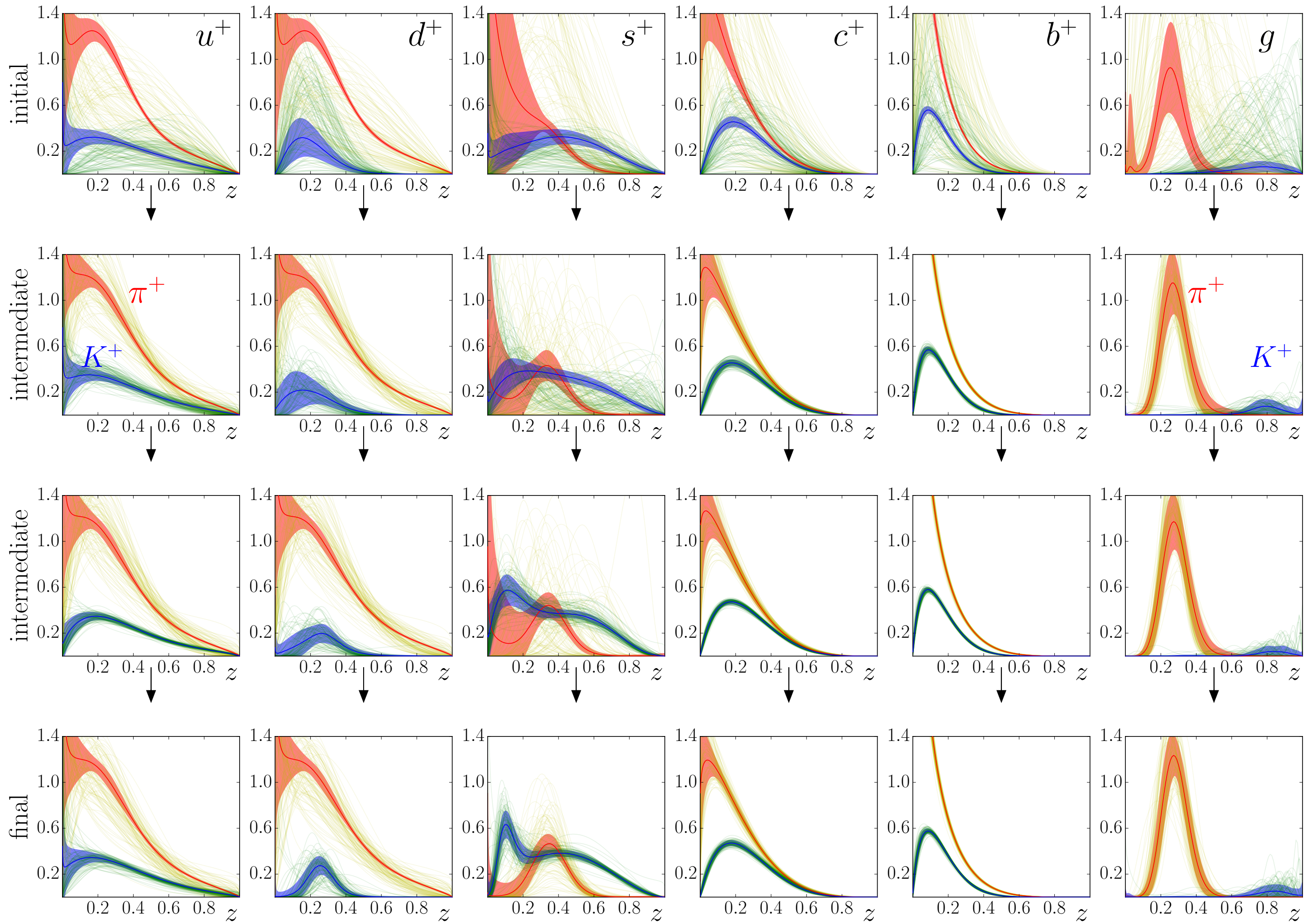
  → $\Delta s > 0$ for "DSS" FFs *de Florian et al. (2007)*

  $\Delta s < 0$ for "HKNS" FFs *Hirai et al. (2007)*

  → need to understand origin of differences in fragmentation functions

Previous JAM analyses

- First MC analysis of fragmentation functions



$e^+e^- \to h\,X$

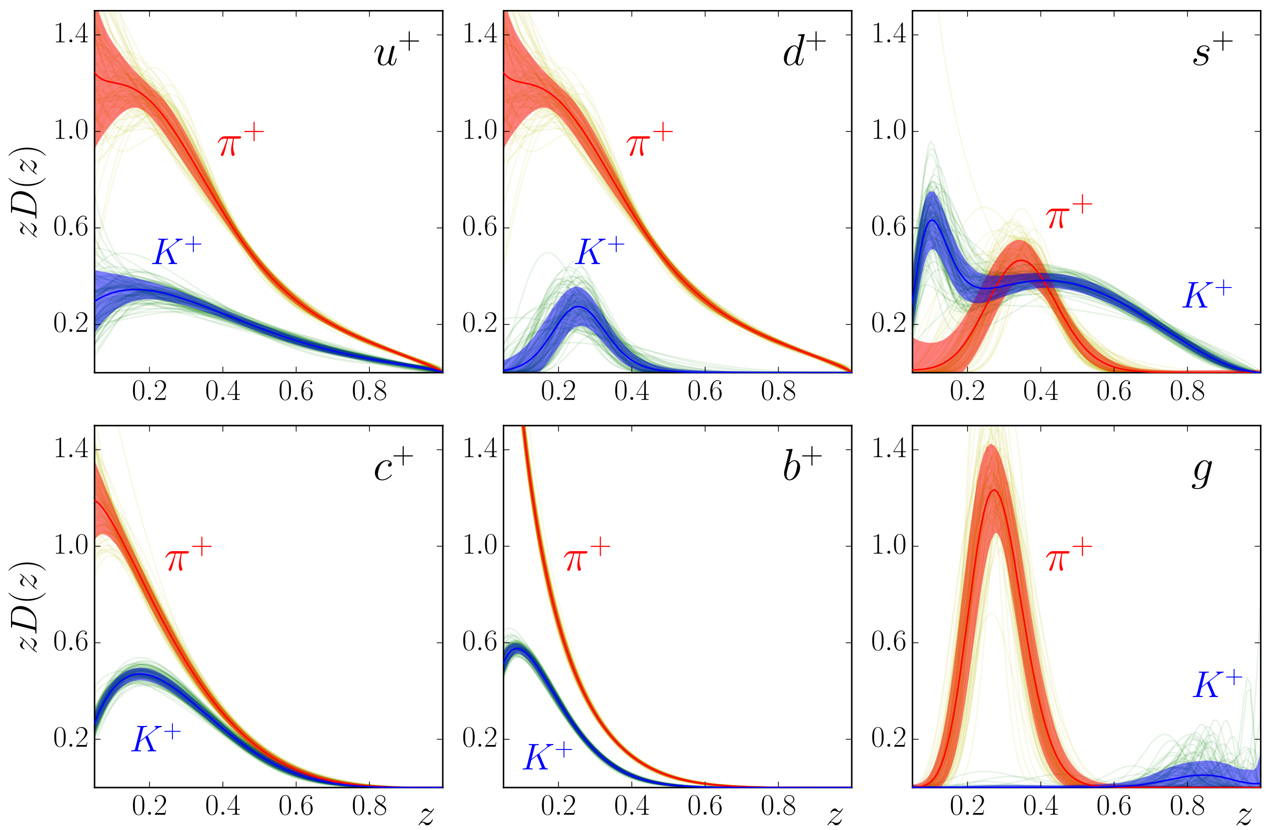
single-inclusive annihilation (SIA)

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

$\longrightarrow$ convergence after ~ 20 iterations

Previous JAM analyses

- First MC analysis of fragmentation functions



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

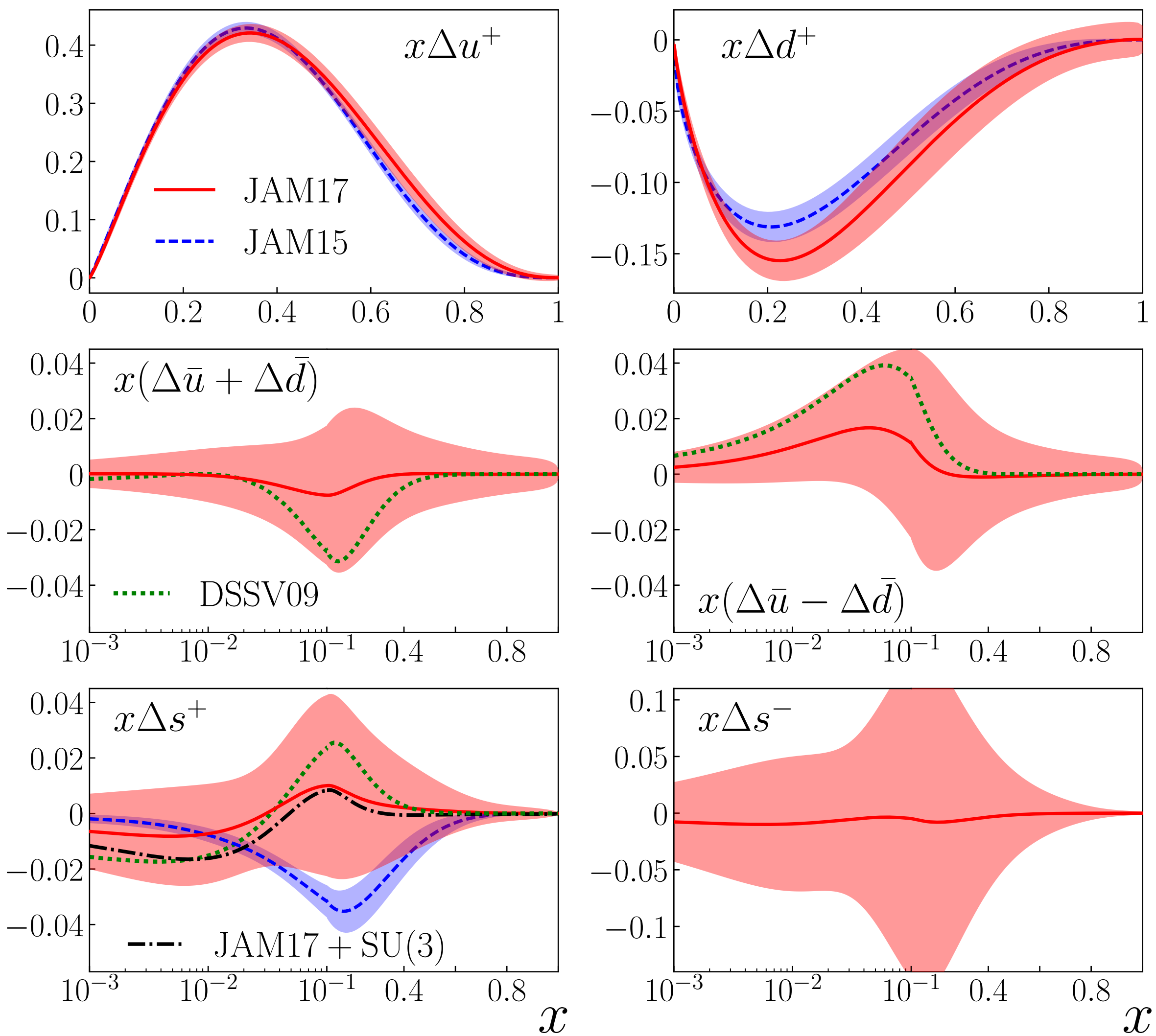
$\longrightarrow$ favored FFs well constrained; unfavored not as well…

$\longrightarrow$ nontrivial shape of $s \to K$ fragmentation

$\longrightarrow$ hard $g \to K$ fragmentation?

Previous JAM analyses

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



→ no assumption of SU(3) symmetry

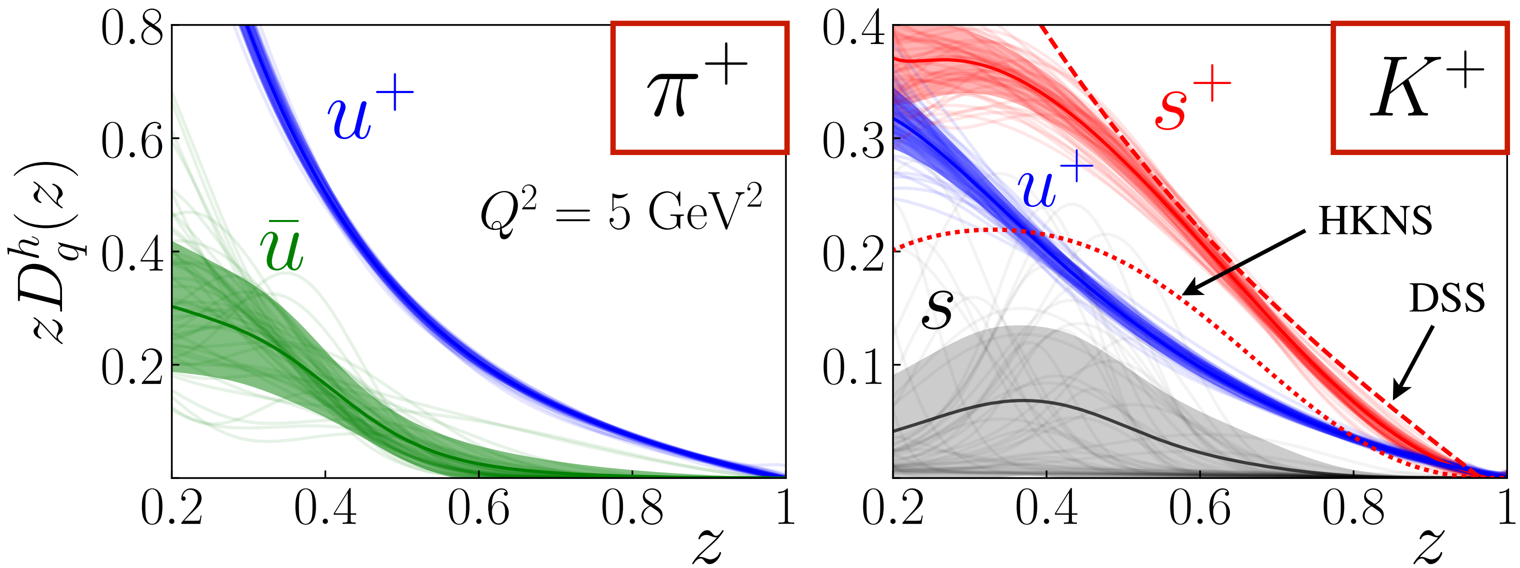
→ Δs slightly > 0 at high x, consistent with zero

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

Ethier, Sato, WM
PRL 119, 132001 (2017)

Previous JAM analyses

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



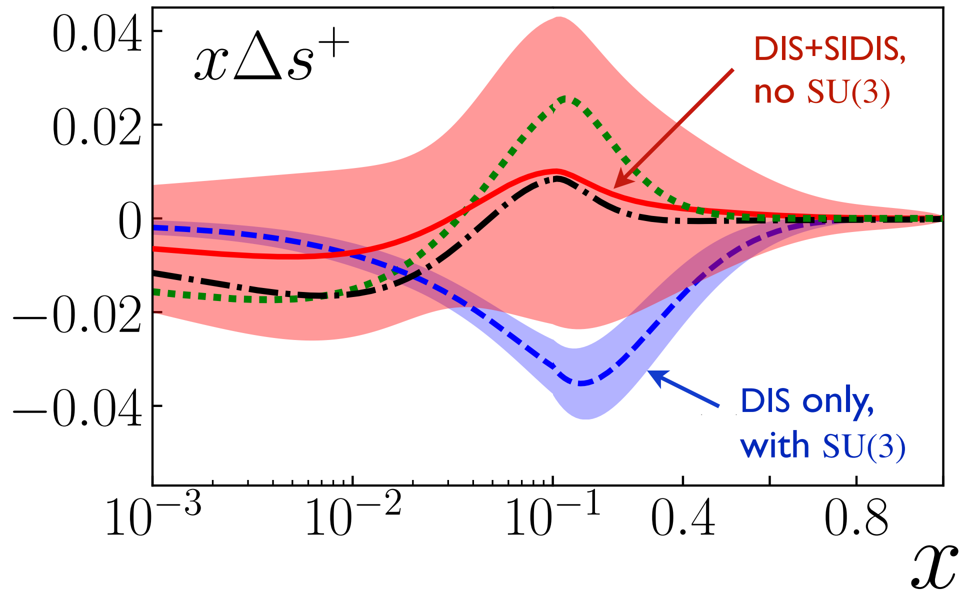
Ethier, Sato, WM (2017)

$\longrightarrow$ some constraint from SIDIS on unfavored FFs $(e.g.\ \ s \to K^+)$, but uncertainties still large

$\longrightarrow$ new result more consistent with DSS at moderate z

Previous JAM analyses

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



Ethier, Sato, WM (2017)

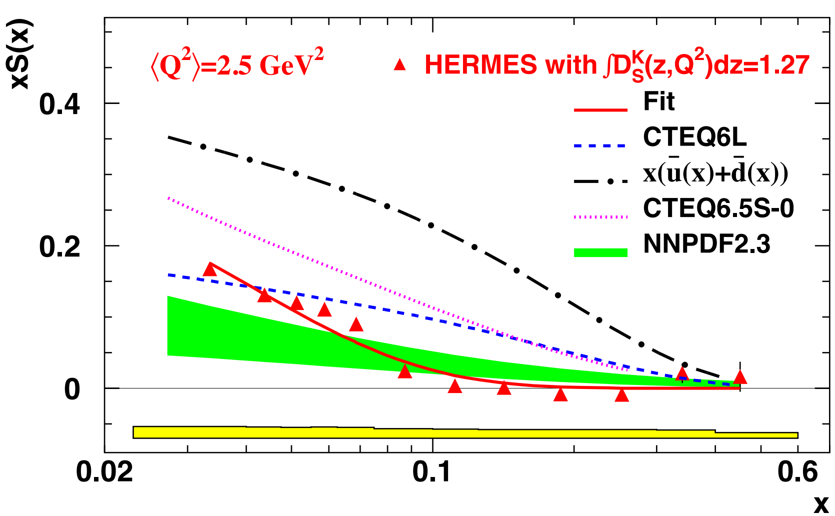
$\longrightarrow$ weak sensitivity to Δs^+ from DIS data & evolution
— $\mathrm{SU(3)}$ pulls Δs^+ to generate moment ~ -0.1
— negative peak at $x \sim 0.1$ induced by fixing $b \sim 6-8$

$\longrightarrow$ less negative $\Delta s = -0.03(10)$ gives larger total helicity $\Delta\Sigma = 0.36(9)$

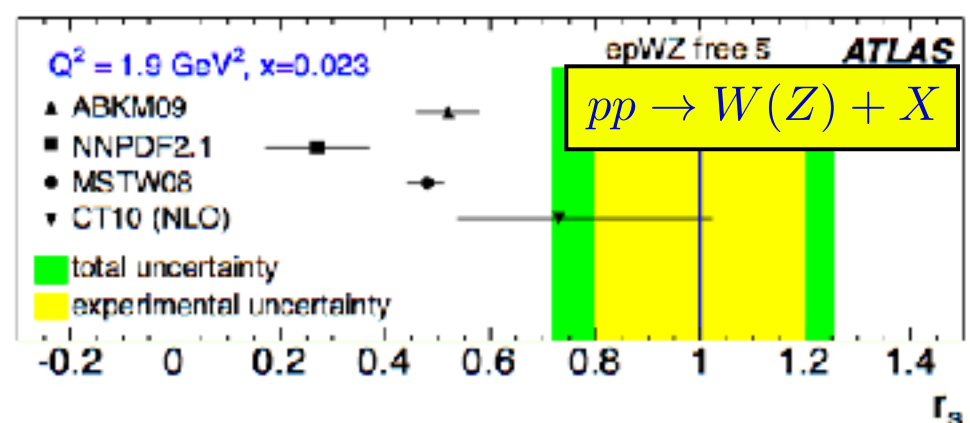
Strange quarks

- 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 $\kappa = \dfrac{s+\bar{s}}{\bar{u}+\bar{d}} \sim 0.2 - 0.5$



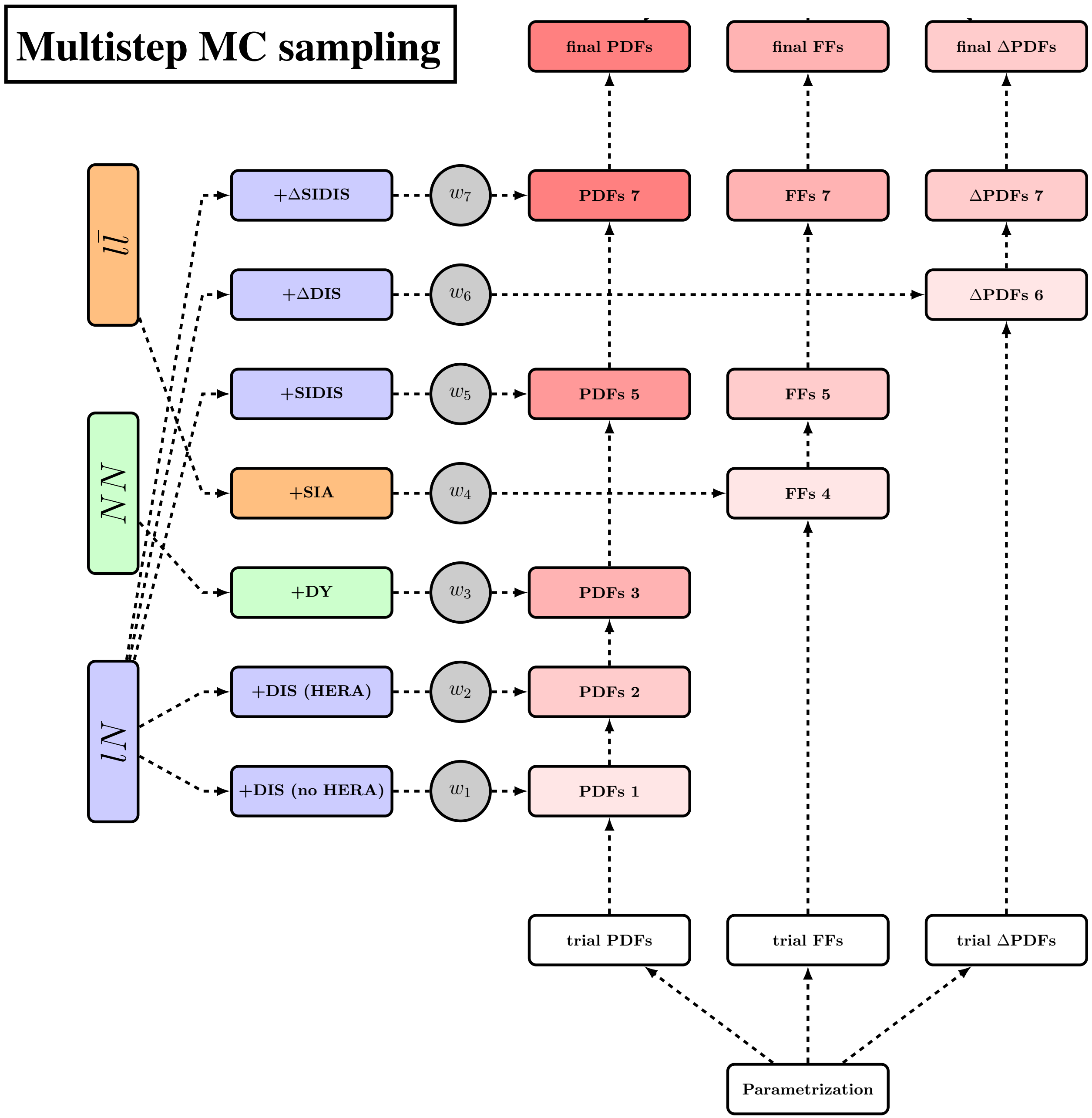
HERMES, PRD 89 (2014) 097101



ATLAS PRL 109 (2012) 012001

$$r_s = (s+\bar{s})/2\bar{d} = 1.00^{+0.25}_{-0.28}$$

JAM 2019 analysis



JAM 2019 analysis

- "Inverse problem" leads to multiple solutions
  $\longrightarrow$ explore possible correlations between solutions by identifying *clusters* of solutions

- Use *k*-means clustering algorithm to identify regions in parameter space where solutions cluster

e.g.

1. *k* initial "means" (in this case *k*=3) are randomly generated within the data domain (shown in color).

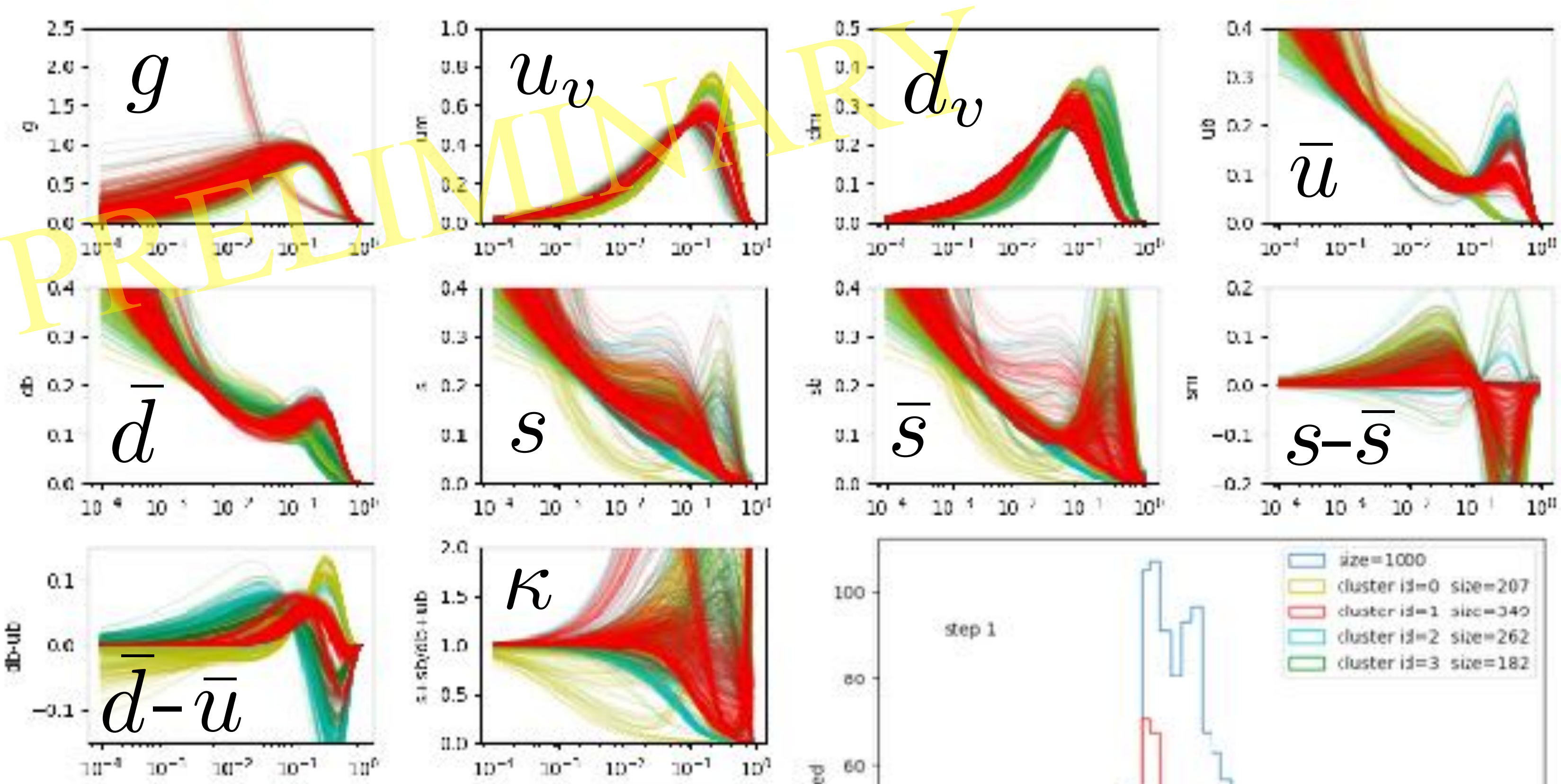
2. *k* clusters are created by associating every observation with the nearest mean. The partitions here represent the Voronoi diagram generated by the means.

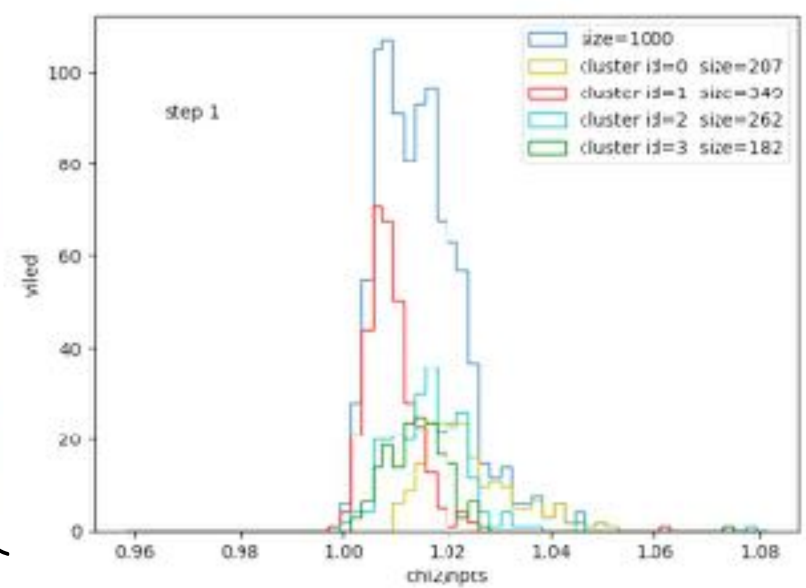
3. The centroid of each of the *k* clusters becomes the new mean.

4. Steps 2 and 3 are repeated until convergence has been reached.

JAM 2019 analysis

- fixed-target DIS only



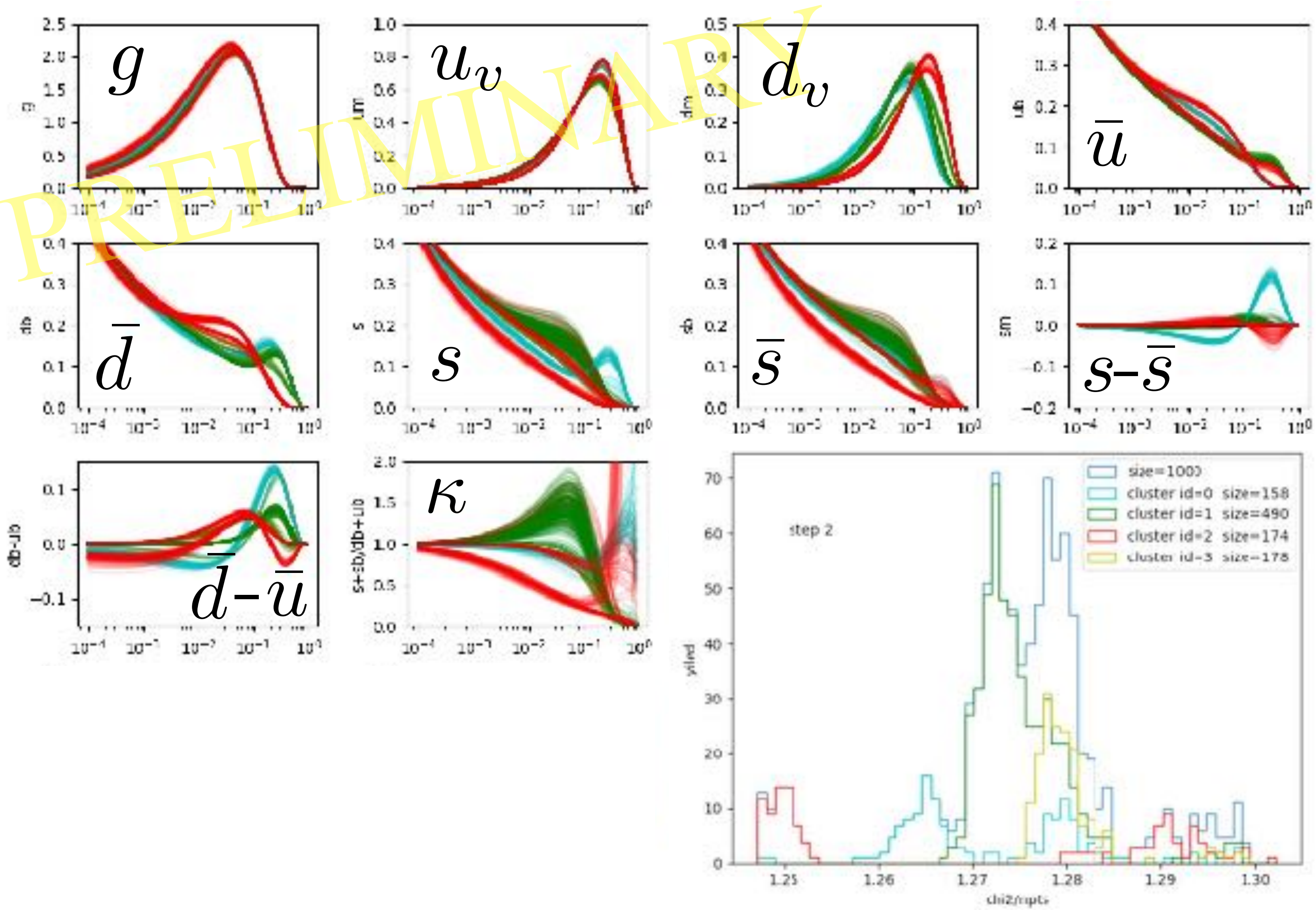


red cyan green yellow

$\chi^2 \downarrow \longleftrightarrow \chi^2 \uparrow$

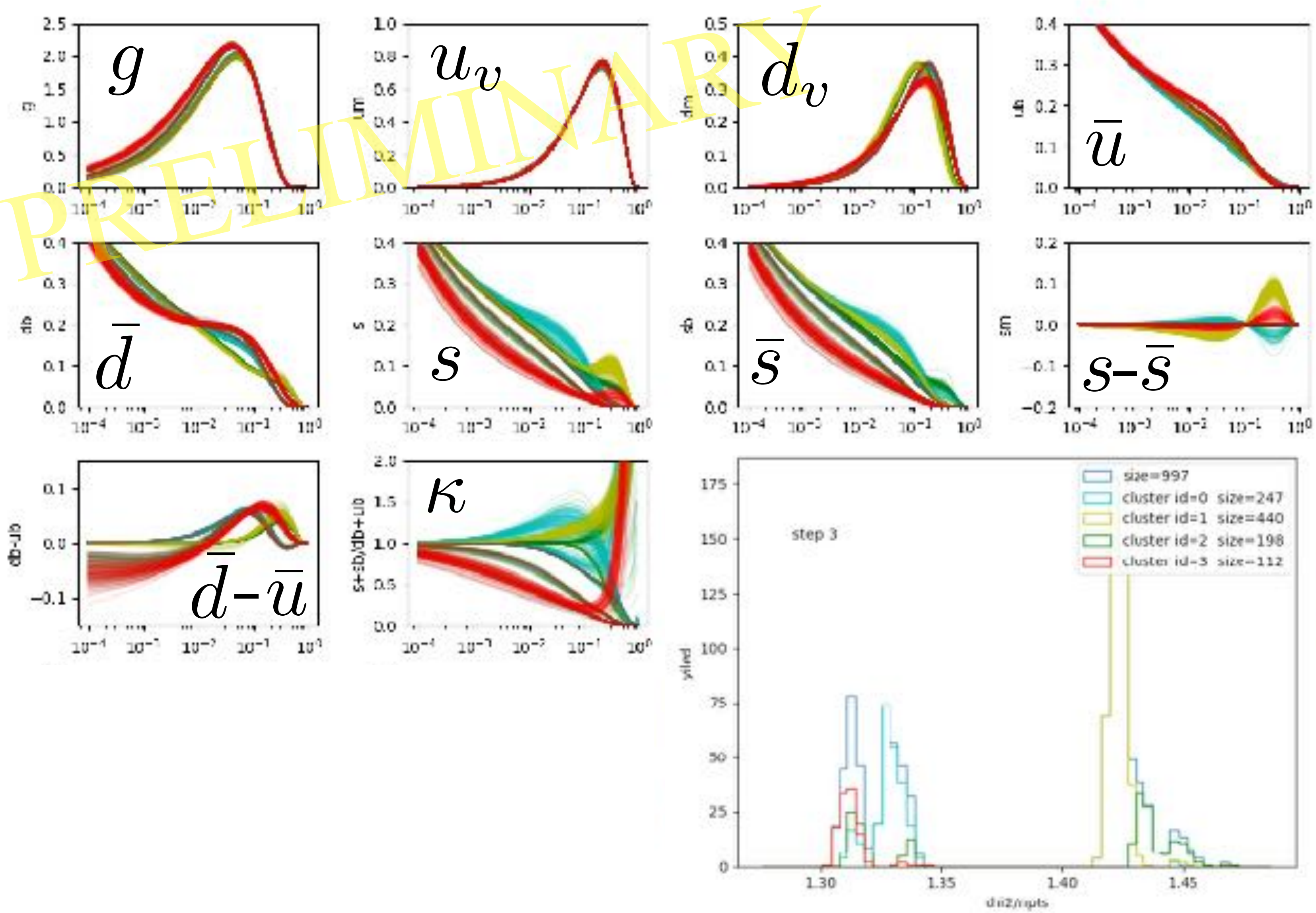
JAM 2019 analysis

- fixed-target DIS + HERA



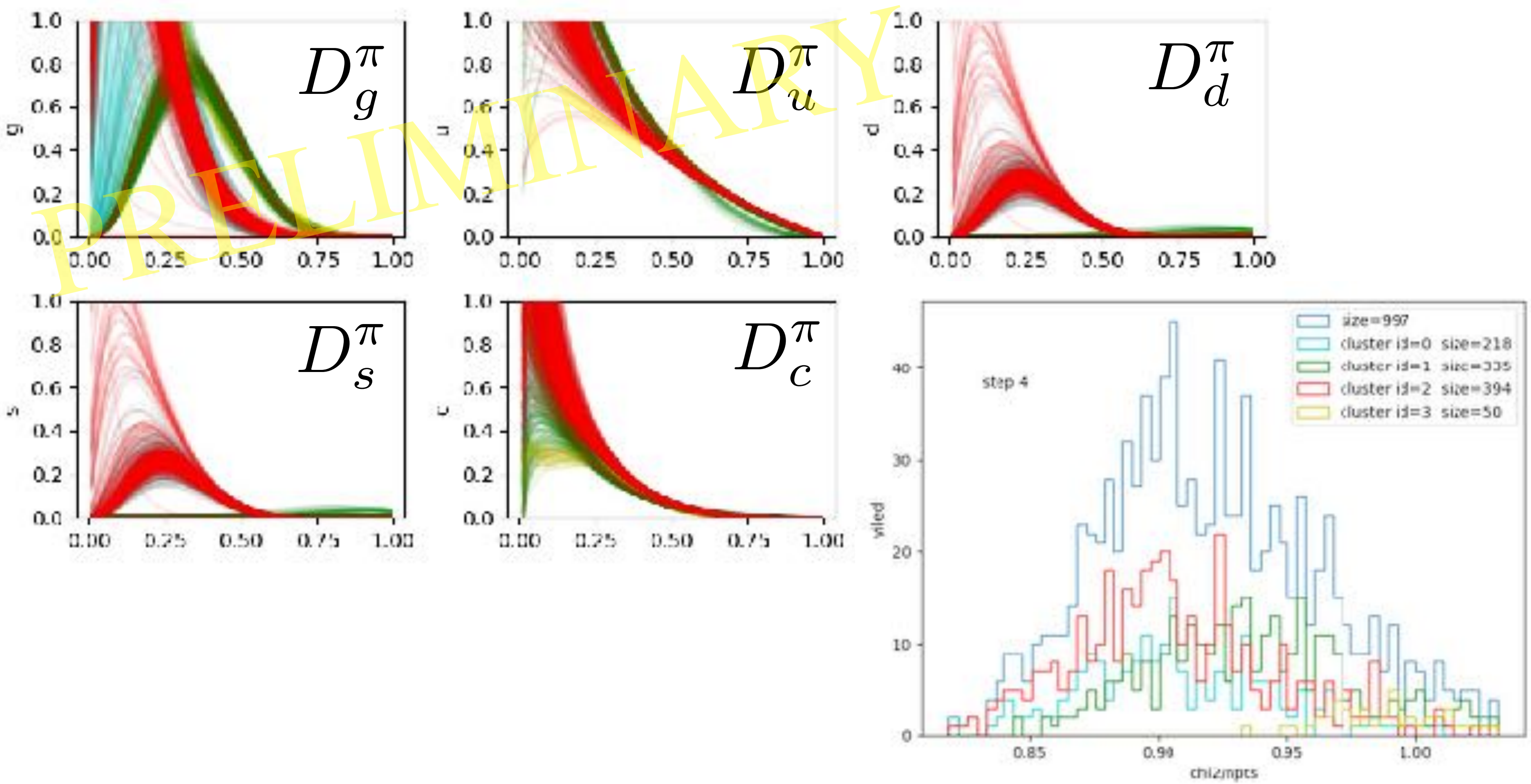
JAM 2019 analysis

- fixed-target DIS + HERA + DY



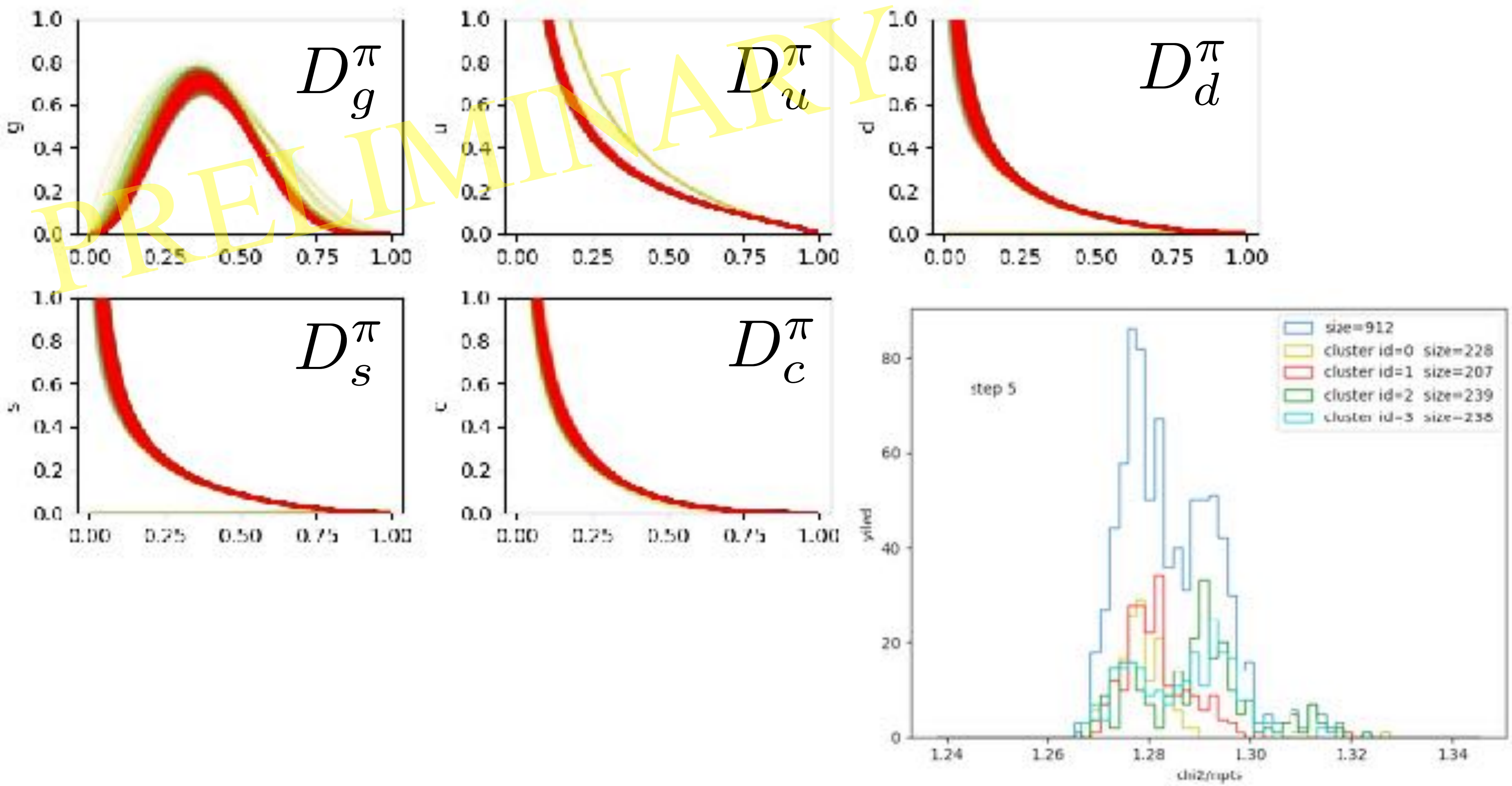
JAM 2019 analysis

- SIA (pion)



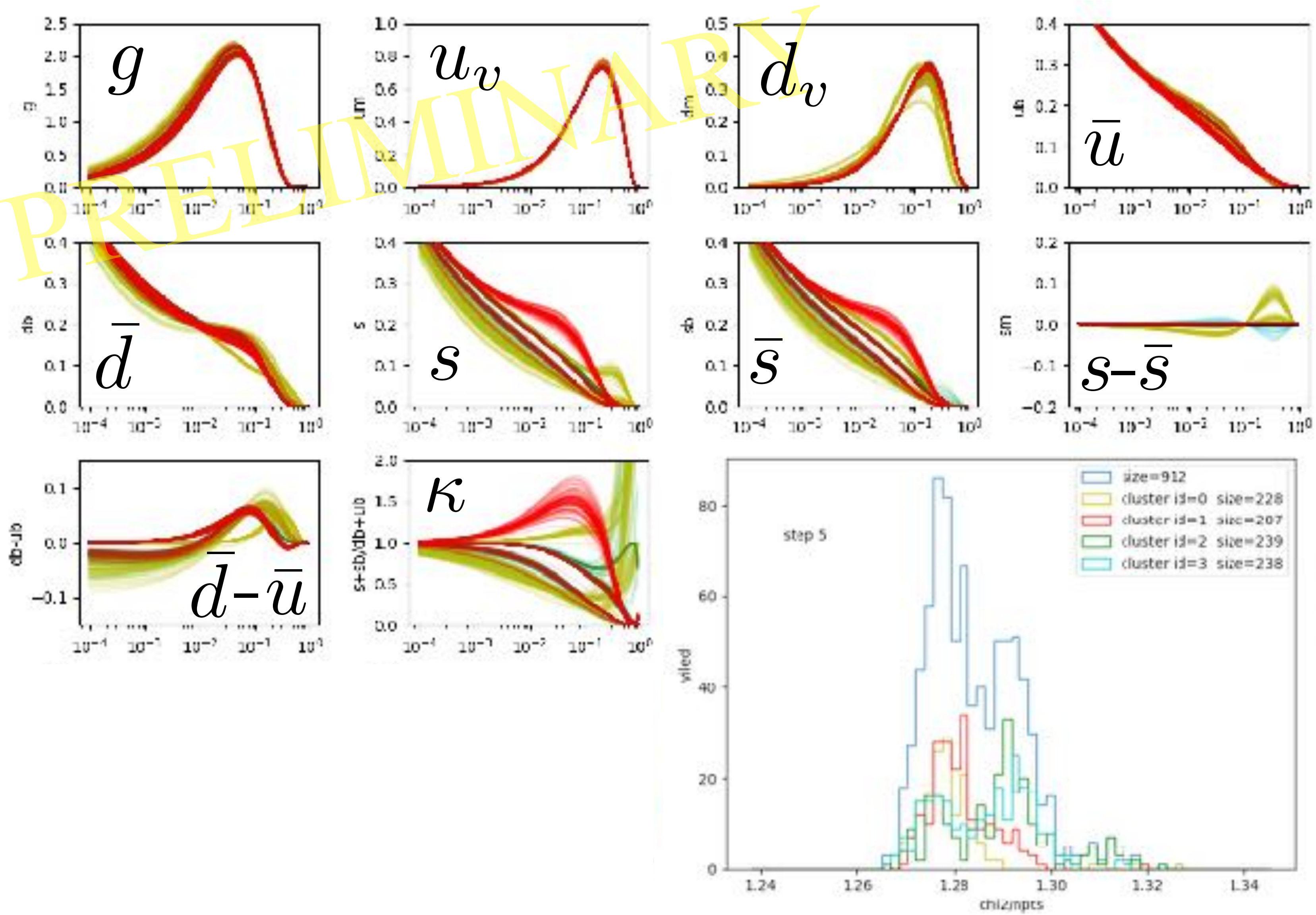
JAM 2019 analysis

- SIA (pion) + SIDIS (pion)



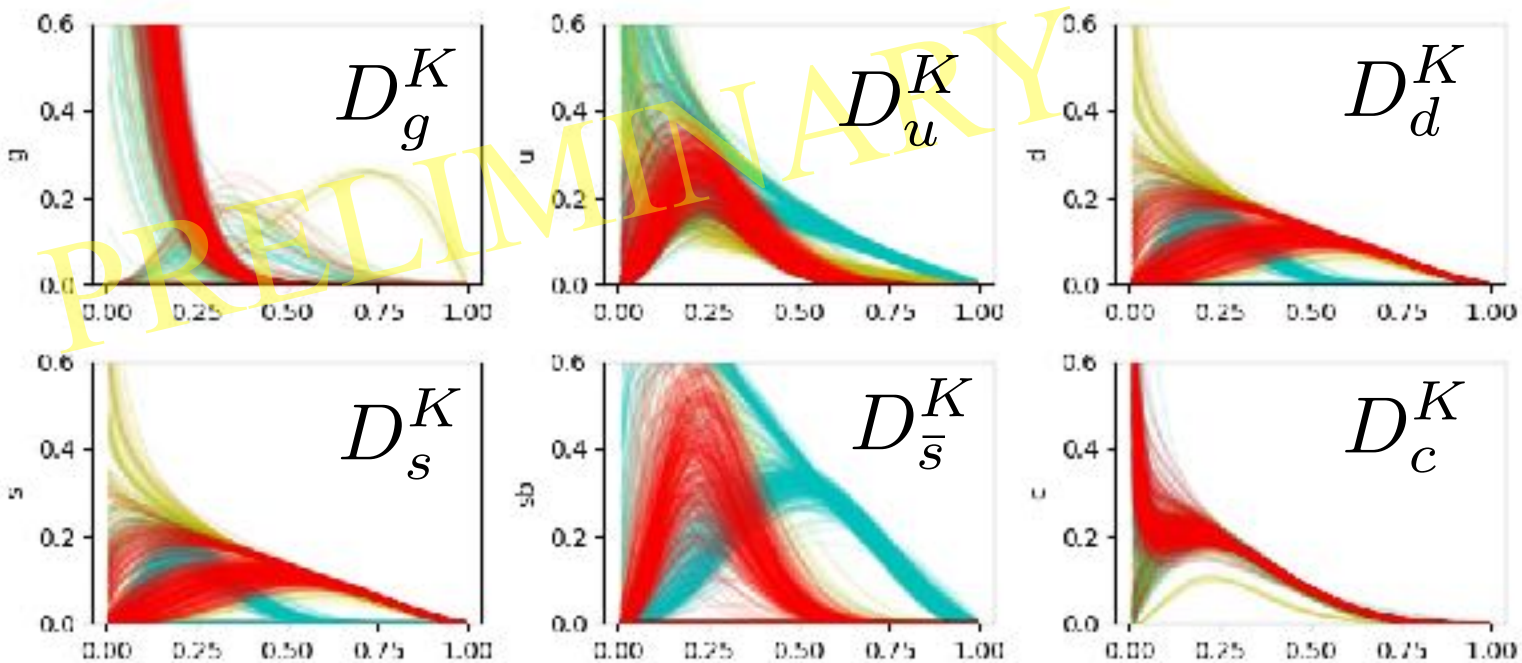
JAM 2019 analysis

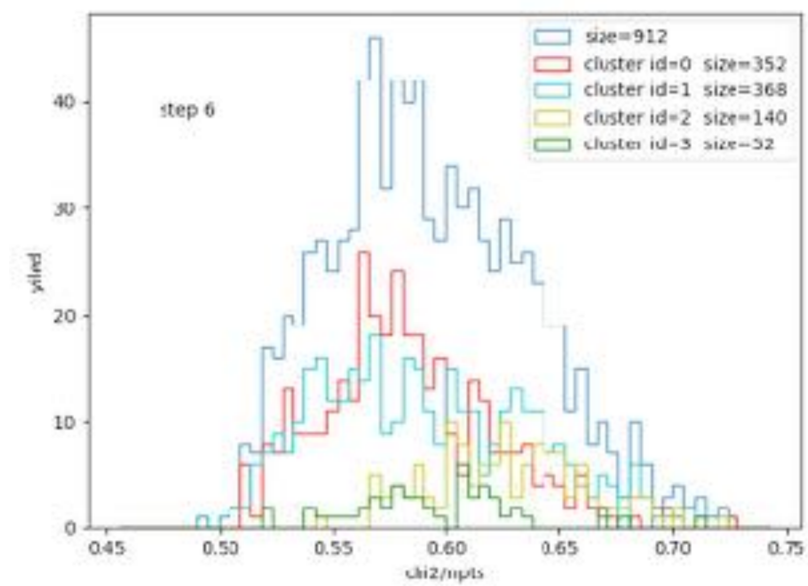
- fixed-target DIS + HERA + DY + SIA (pion) + SIDIS (pion)



JAM 2019 analysis

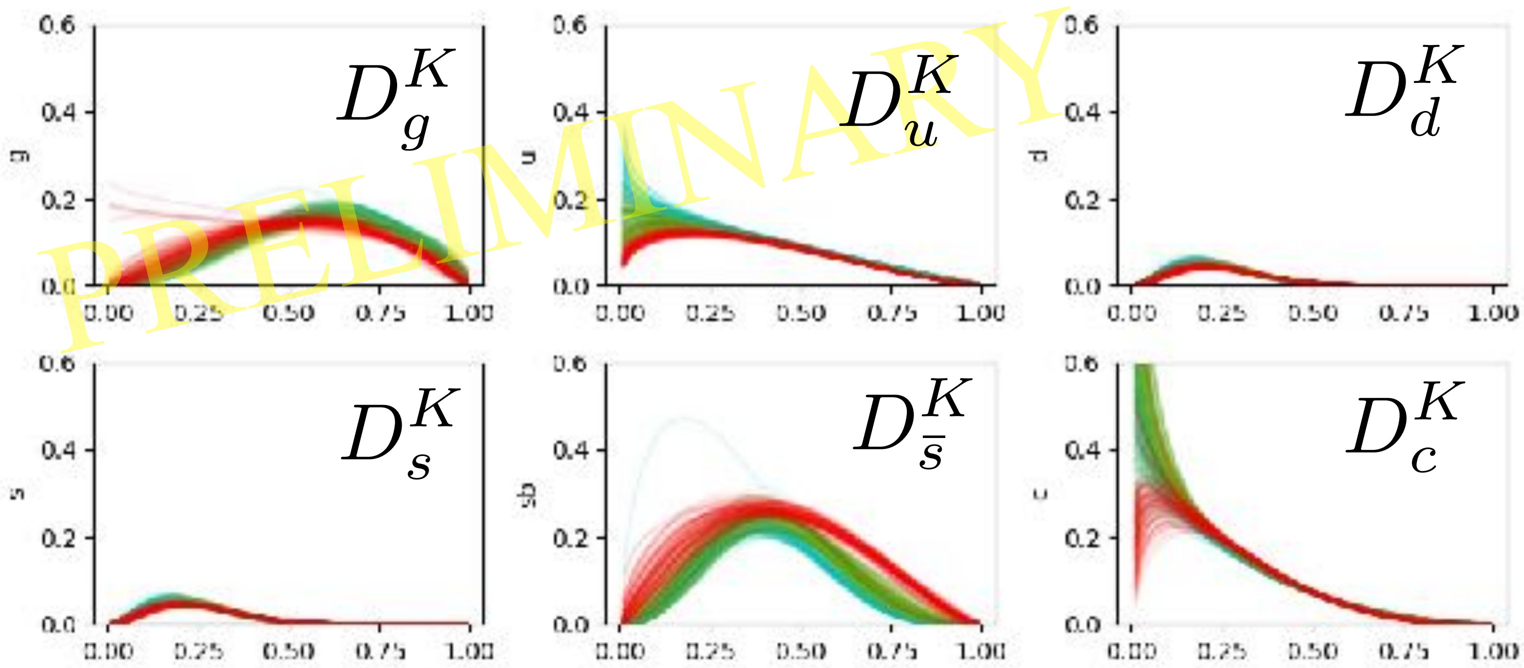
- SIA (kaon)

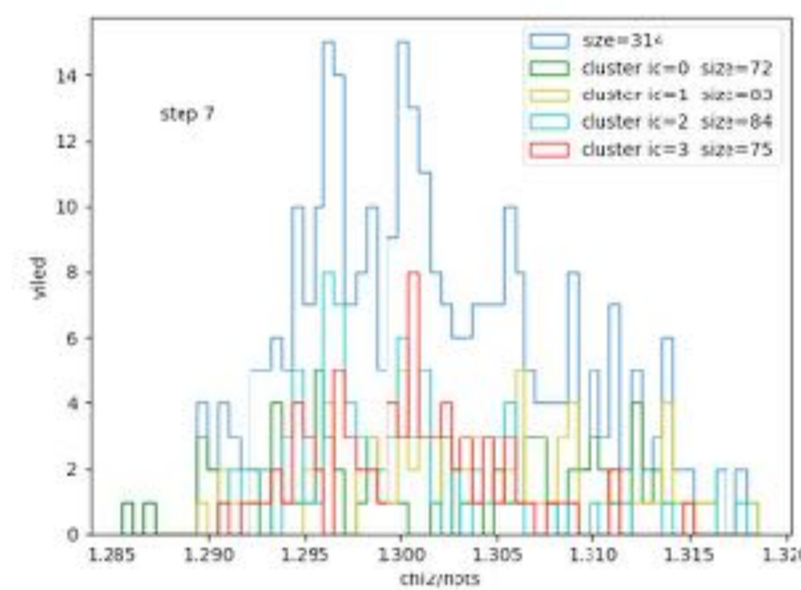




JAM 2019 analysis

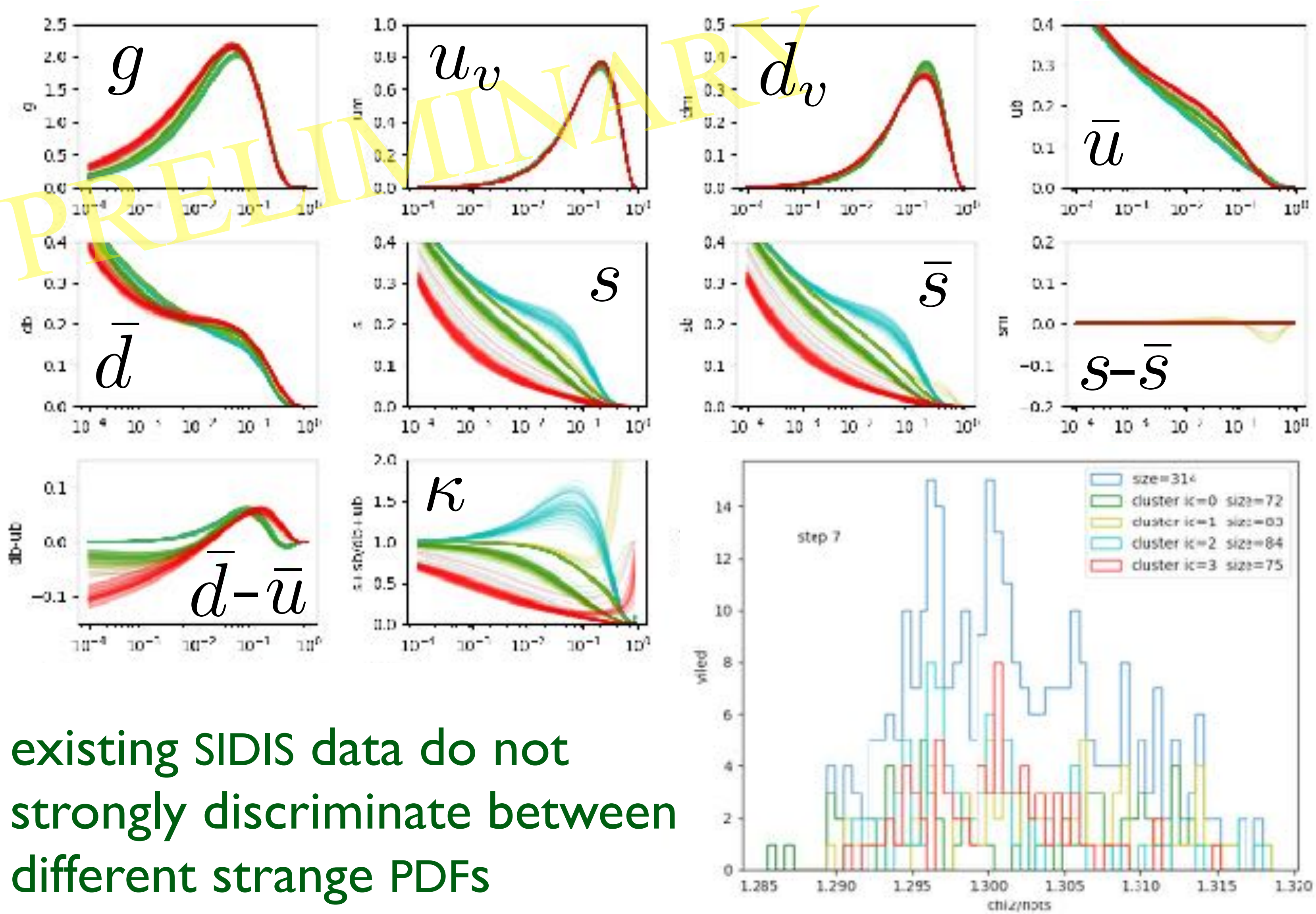
- SIA (kaon) + SIDIS (kaon) + …





JAM 2019 analysis

- fixed-target DIS + HERA + DY + SIA (pion) + SIDIS (pion)



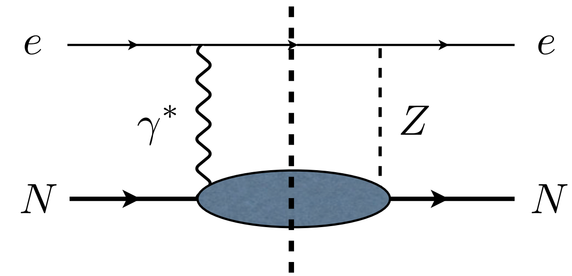
→ existing SIDIS data do not strongly discriminate between different strange PDFs

→ need more precise SIDIS data (JLab12, EIC)

Strange quarks from PVDIS

- Parity-violating DIS allows strange contribution to be isolated, when combined with e.m. p and n DIS data at low/intermediate x

→ at leading order



$$F_2^{\gamma p} = \frac{4}{9}x(u+\bar{u}) + \frac{1}{9}x(d+\bar{d}+s+\bar{s}) + \cdots$$

$$F_2^{\gamma n} = \frac{4}{9}x(d+\bar{d}) + \frac{1}{9}x(u+\bar{u}+s+\bar{s}) + \cdots$$

$$F_2^{\gamma Z,p} = \left(\frac{1}{3} - \frac{8}{9}\sin^2\theta_W\right)x(u+\bar{u}) + \left(\frac{1}{6} - \frac{2}{9}\sin^2\theta_W\right)(d+\bar{d}+s+\bar{s}) + \cdots$$

$$\approx \frac{1}{9}x(u+\bar{u}+d+\bar{d}+s+\bar{s}) + \cdots \quad \text{for } \sin^2\theta_W \approx 1/4$$

→ 3 equations with 3 unknowns;
<u>order of magnitude greater sensitivity</u> of γZ to strange PDF

→ V x A term also sensitive to $s - \bar{s}$

$$F_3^{\gamma Z,p} = \frac{2}{3}(u-\bar{u}) + \frac{1}{3}(d-\bar{d}+s-\bar{s}) + \cdots$$

Outlook

- New paradigm in global analysis — *simultaneous* determination of collinear distributions using MC sampling of parameter space

- *Short-term*: "universal" QCD analysis of all observables sensitive to collinear (unpolarized & polarized) PDFs and FFs

- *Longer-term*: technology developed will be applied to global analysis of transverse momentum dependent (TMD) distributions to map out full 3-d image of hadrons