JAM multi-step strategy

Carlota Andrés
Jefferson Lab
A.I. for Nuclear Physics Workshop

In collaboration with: J.J. Ethier, W. Melnitchouk, N. Sato

Jefferson Lab Angular Momentum Collaboration

Jefferson Lab
EXPLORING THE NATURE OF MATTER

Motivation

- Traditionally different types of collinear distributions (PDFs, FFs) are obtained from independent analyses.

- Performing **simultaneous** fits of different collinear distributions allows us to:

  - Study the limits in x and Q^2 of collinear factorization

  - Test the universality of PDFs, FFs…

  - Extract the distributions in a **rigorous** way where all the data are studied using the **same** theoretical framework

- In this talk: (first) **simultaneous** analysis of **unpolarized** PDFs and FFs $\longrightarrow$ **Strange** quark distribution

Evolution of JAM

		JAM15	JAM16	JAM17	JAM19	JAM20?
Process	DIS	✓	✗	✓	✓	✓
	DY	✗	✗	✗	✓	✓
	SIA	✗	✓	✓	✓	✓
	SIDIS	✗	✗	✓	✓	✓
						+ More processes
Function	f	✗	✗	✗	✓	✓
	Δf	✓	✗	✓	✗	✓
	D_f^h	✗	✓	✓	✓	✓

Evolution of JAM

		JAM15	JAM16	JAM17	JAM19	JAM20?
Process	DIS	✓	✗	✓	✓	✓
	DY	✗	✗	✗	✓	✓
	SIA	✗	✓	✓	✓	✓
	SIDIS	✗	✗	✓	✓	✓
						+ More processes
Function	f	✗	✗	✗	✓	✓
	Δf	✓	✗	✓	✗	✓
	D_f^h	✗	✓	✓	✓	✓

First **simultaneous** analysis of **unpolarized** PDFs and FFs

Evolution of JAM

		JAM15	JAM16	JAM17	JAM19	JAM20?
Process	DIS	✓	✗	✓	✓	✓
	DY	✗	✗	✗	✓	✓
	SIA	✗	✓	✓	✓	✓
	SIDIS	✗	✗	✓	✓	✓
						+ More processes
Function	f	✗	✗	✗	✓	✓
	Δf	✓	✗	✓	✗	✓
	D_f^h	✗	✓	✓	✓	✓

First **simultaneous** analysis of **unpolarized** PDFs and FFs

Why JAM19?

To study the **strange** quark distribution

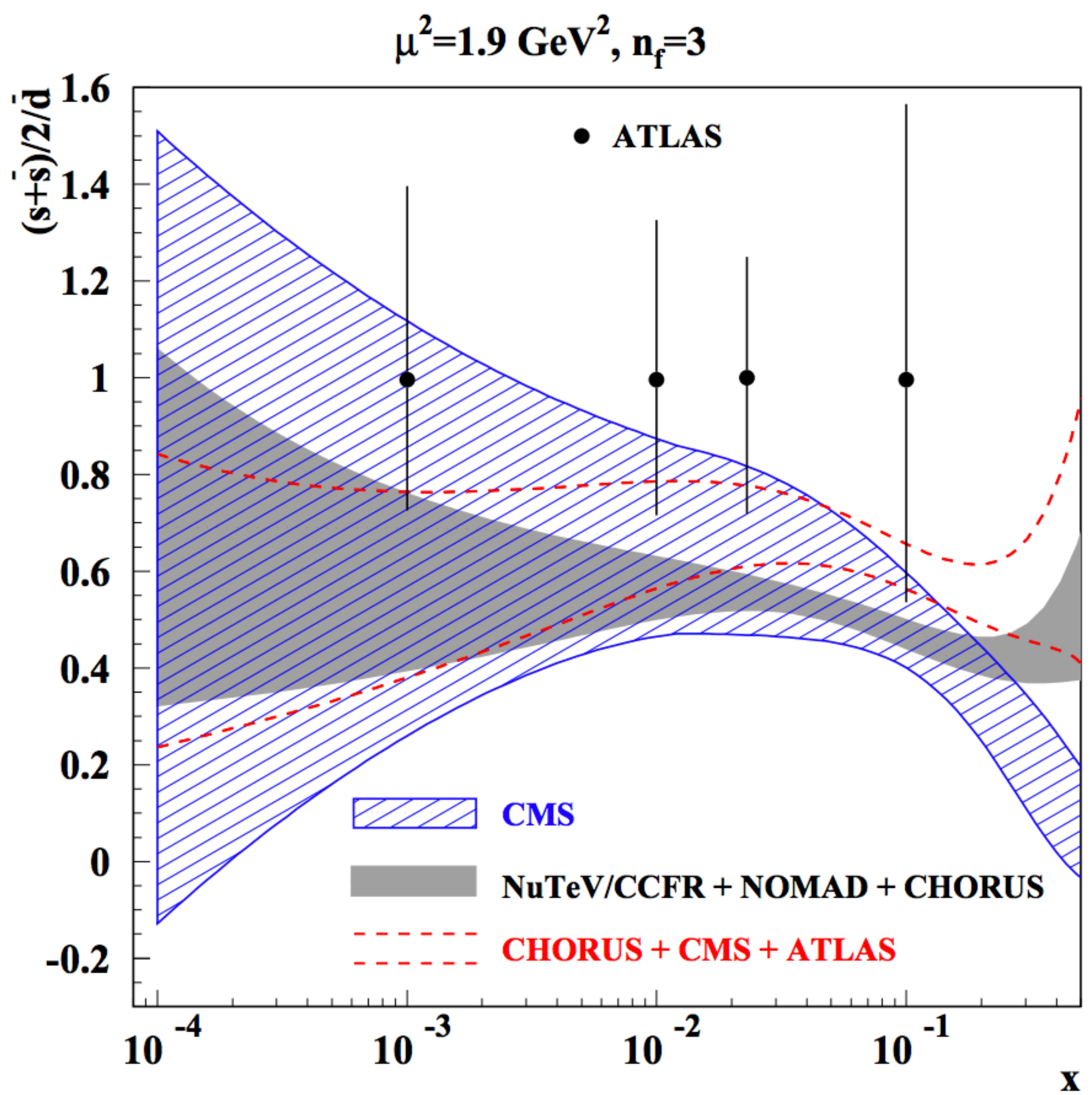
Motivation II

- The strange PDF is less known than the non-strange light flavors
- Traditionally: neutrino-(heavy) nucleus DIS data used to extract the strange PDF.
  - Drawbacks: nuclear effects on PDFs.
- W and Z inclusive production in p-p collisions also sensitive to flavor separation
  - Drawbacks: tension between CMS and ATLAS results?

Motivation II

Alekhin et al. arXiv:1404.6469 [hep-ph]

$r_s =$



First challenge: a wide variety of data

$\text{DIS}: l + (p,d) \to l' + X$

$\text{DY}: l + (p,d) \to l\bar{l} + X$

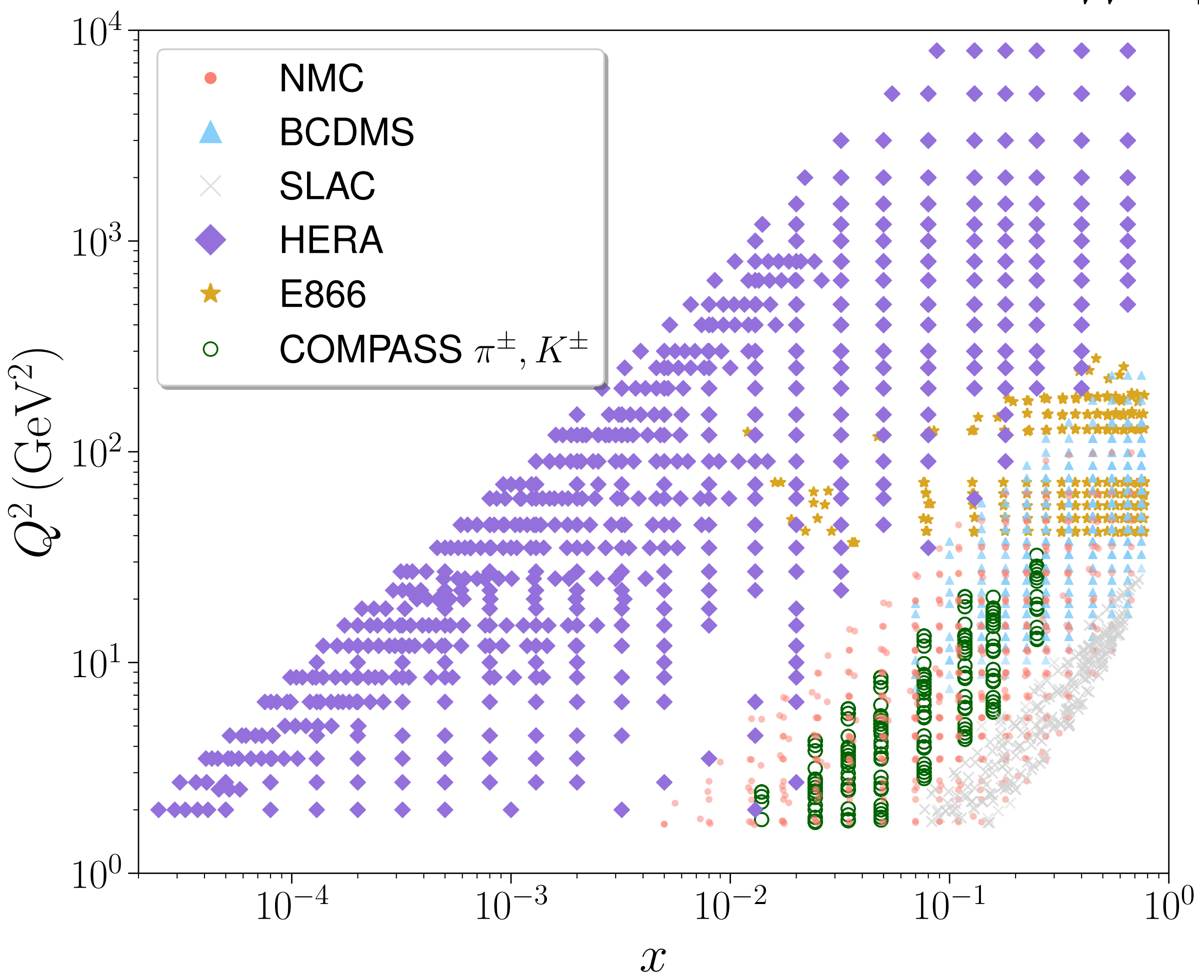
$\text{SIA}: e^+ + e^- \to h + X$

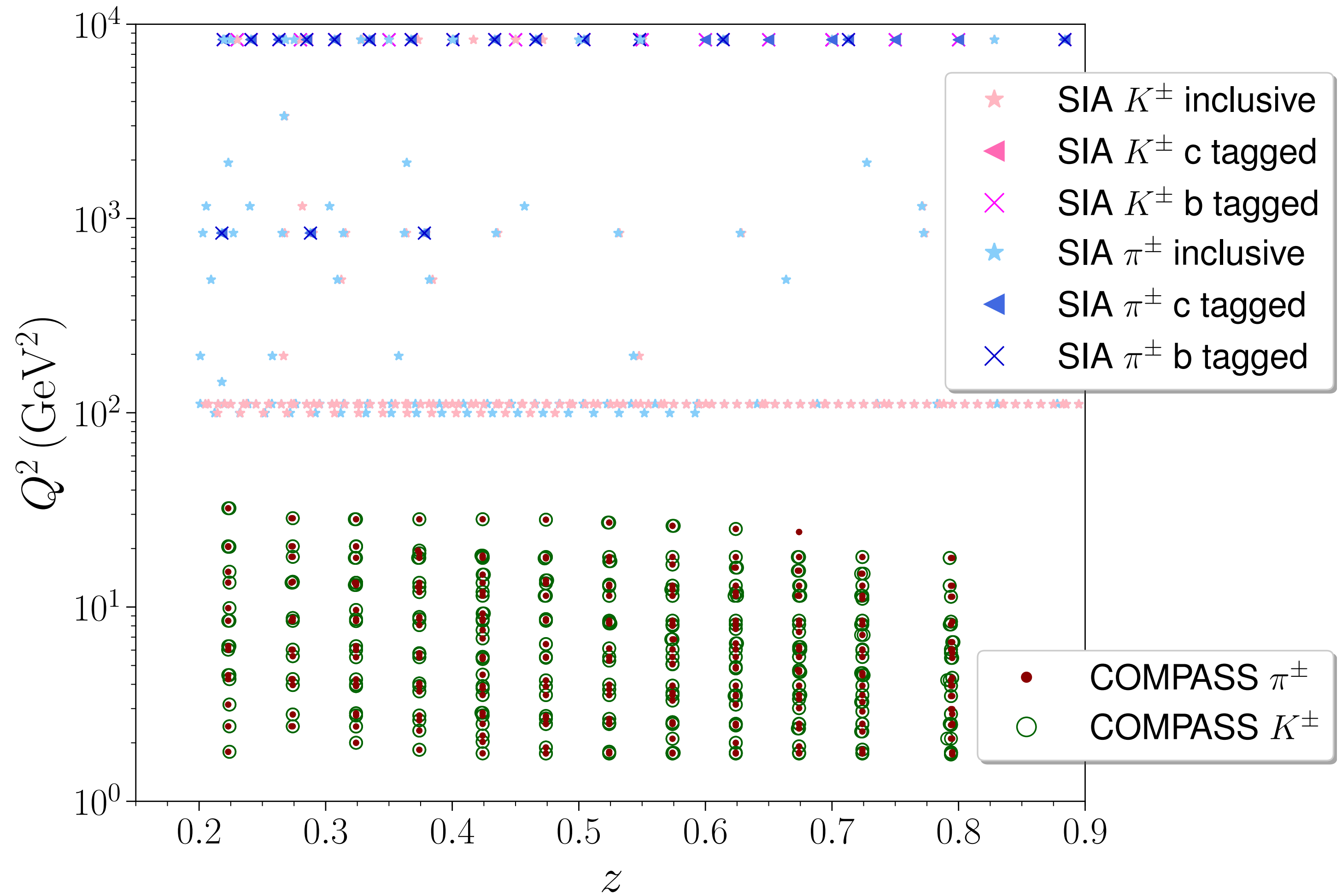
$\text{SIDIS}: l + d \to l' + h + X$

Total: 4366 data points!

$Q^2 > \mathrm{m_c^2}$

$W^2 > 10\,\mathrm{GeV}^2$





Second challenge: MC fits

- Typical PDF parametrization:

$$x\Delta f(x) = Nx^a(1-x)^b(1+c\sqrt{x}+dx)$$

$$\chi^2 = \sum_e^{N_{exp}} \sum_i^{N_{data}} \frac{(D_i^e - T_i)^2}{(\sigma_i^e)^2}$$

- Perform single χ^2-fit: → Multiple local minima!

Parameters difficult to constrain

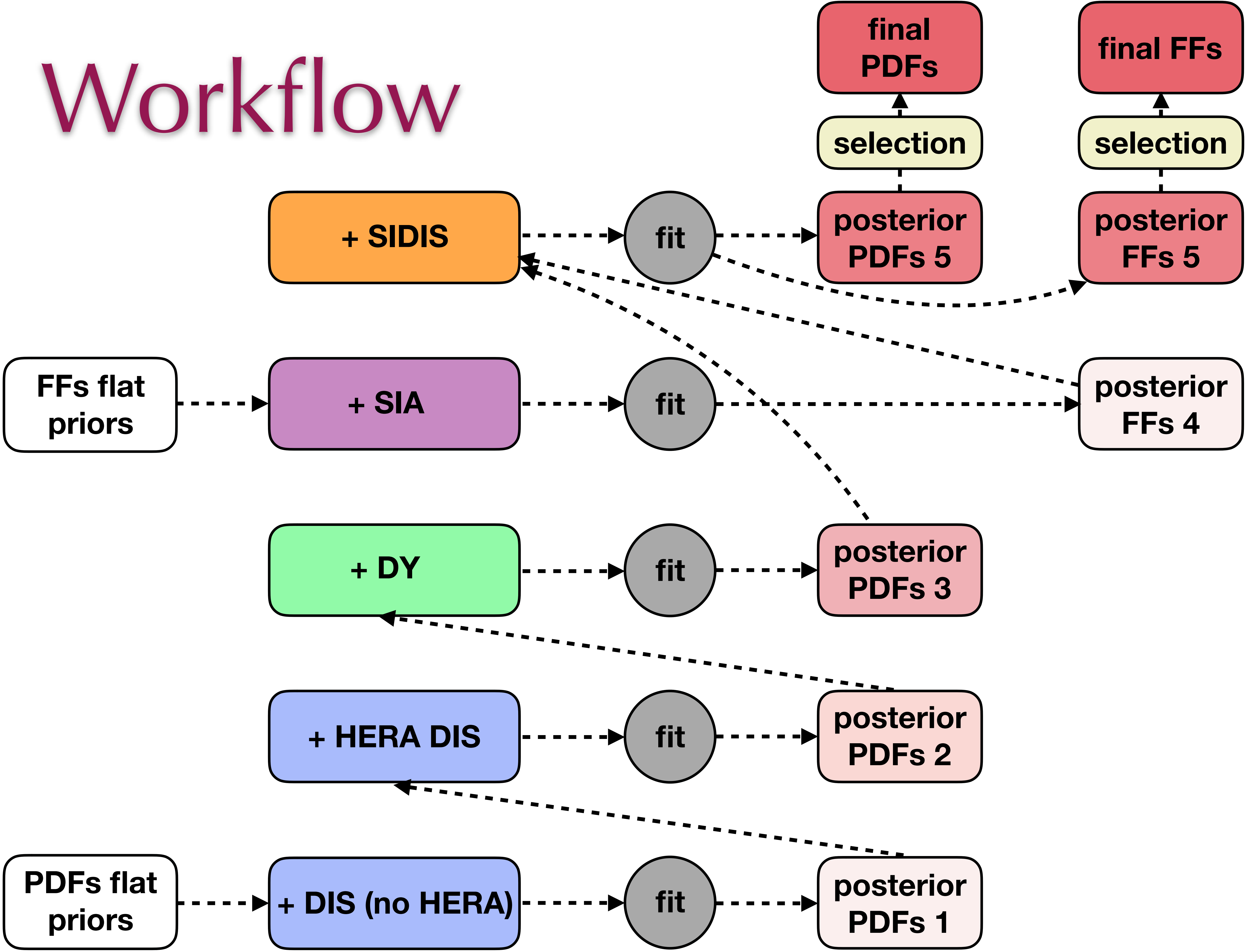
Hessian method for uncertainties → Introduces tolerance criteria

Unsuitable for simultaneous analysis of collinear distributions

- Monte Carlo methods:
  - Allow efficient exploration of the parameter space
  - Uncertainties directly obtained from MC replicas

JAM19 multi-step methodology

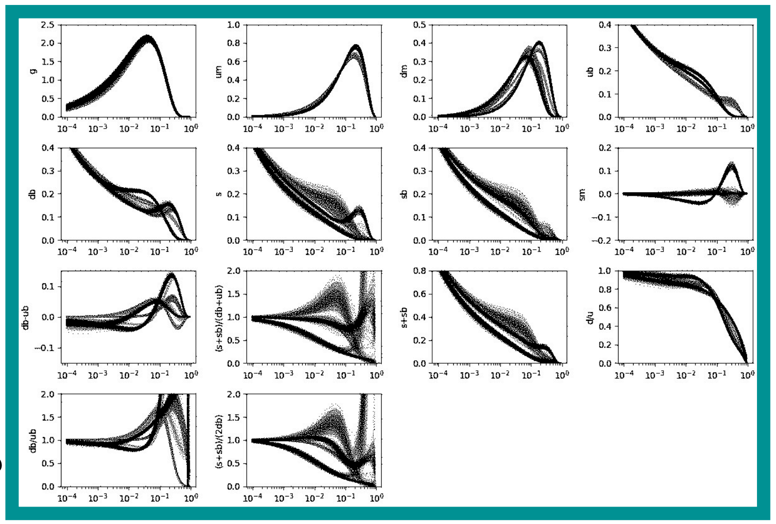
Workflow


final PDFs
final FFs
selection
selection
+ SIDIS
fit
posterior PDFs 5
posterior FFs 5
FFs flat priors
+ SIA
fit
posterior FFs 4
+ DY
fit
posterior PDFs 3
+ HERA DIS
fit
posterior PDFs 2
PDFs flat priors
+ DIS (no HERA)
fit
posterior PDFs 1

JAM19: multi-step fitting

PDFs

$x\,f(x)$

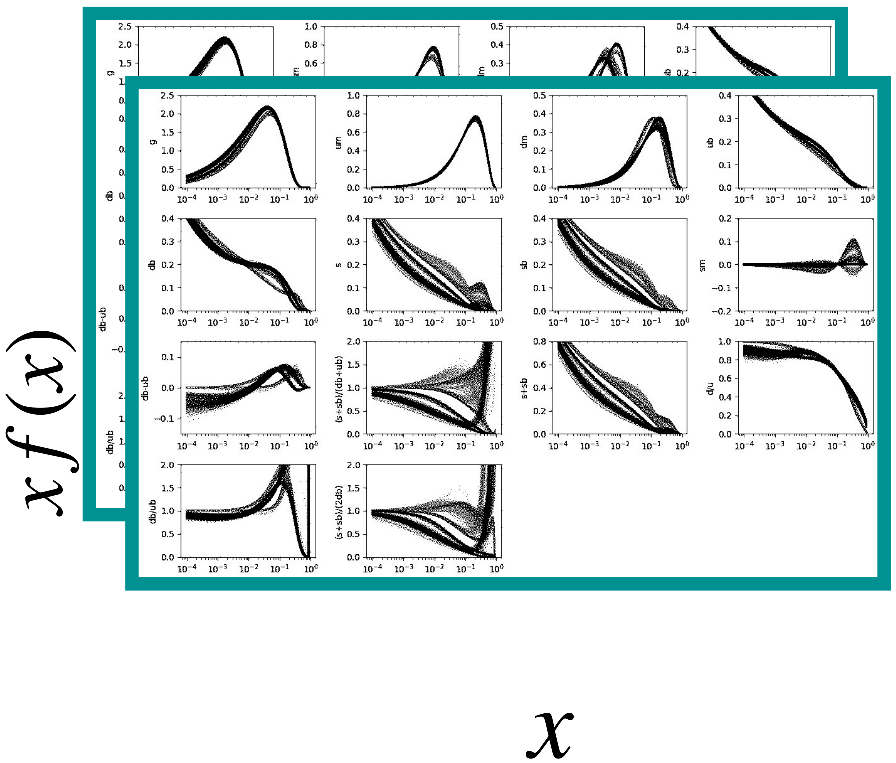


x

\+ DIS data

JAM19: multi-step fitting

PDFs



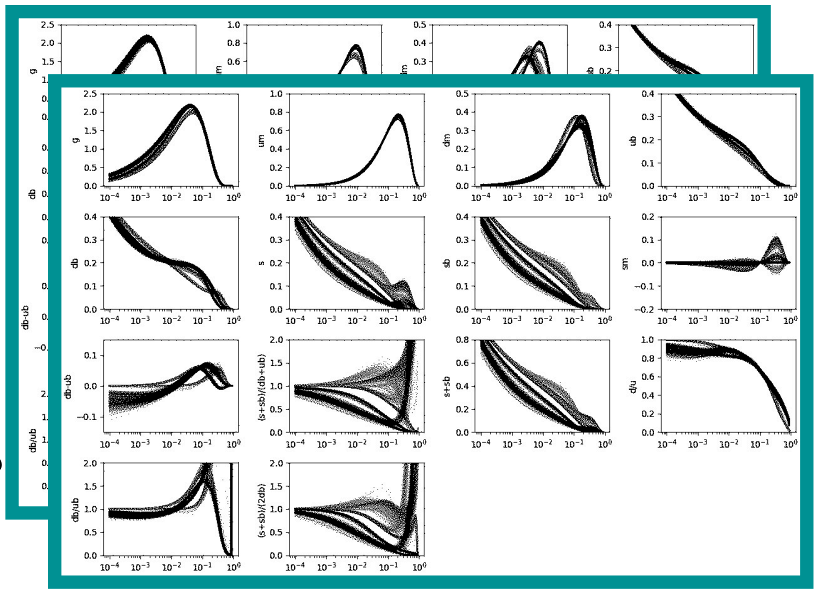
+ DIS data

+ DIS + DY data

JAM19: multi-step fitting

PDFs

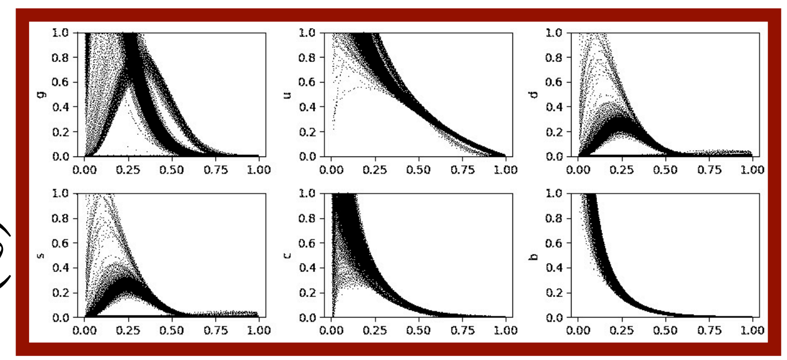
$x\,f(x)$



x

PION FF

$z\,D(z)$



z

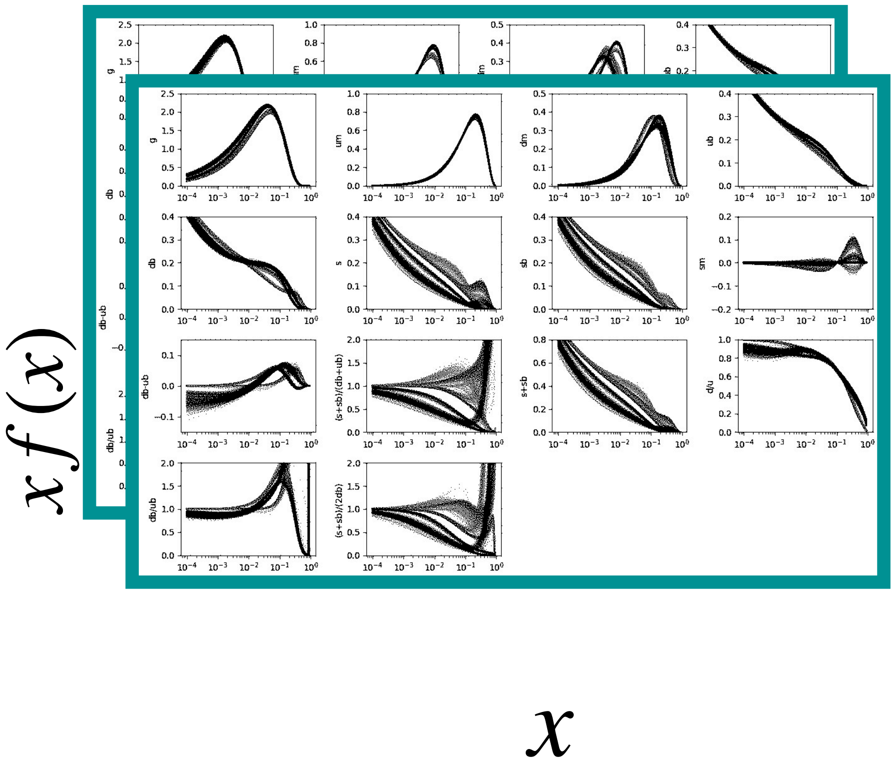
+ DIS data

+ DIS + DY data

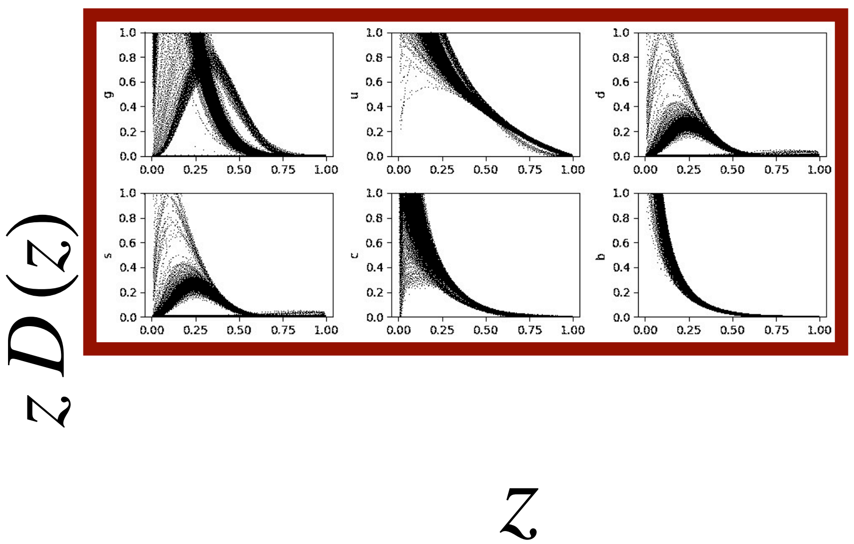
+ SIA pion data

JAM19: multi-step fitting

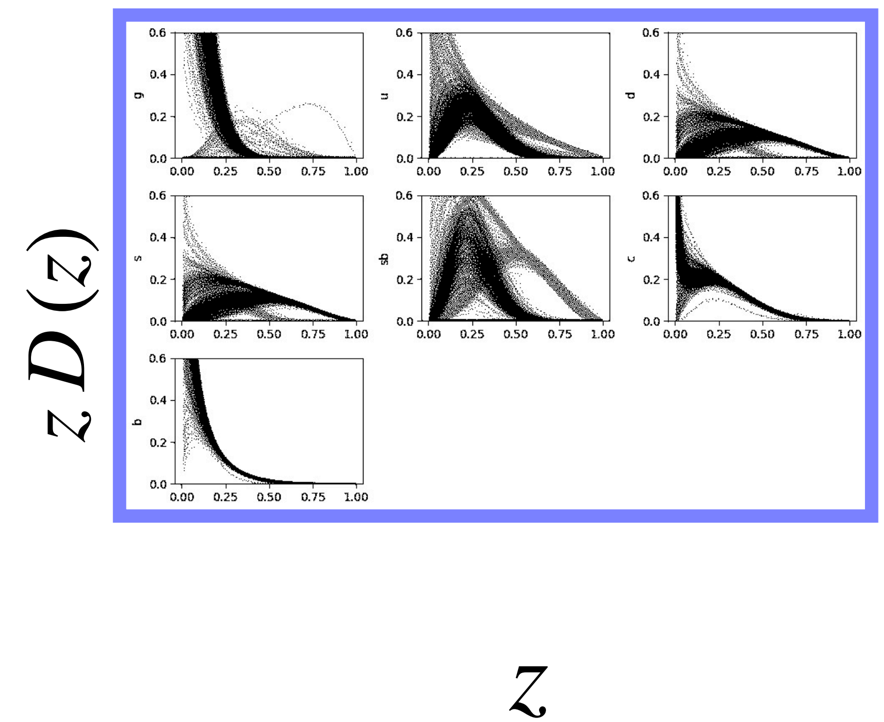
PDFs



PION FF



KAON FF



+ DIS data

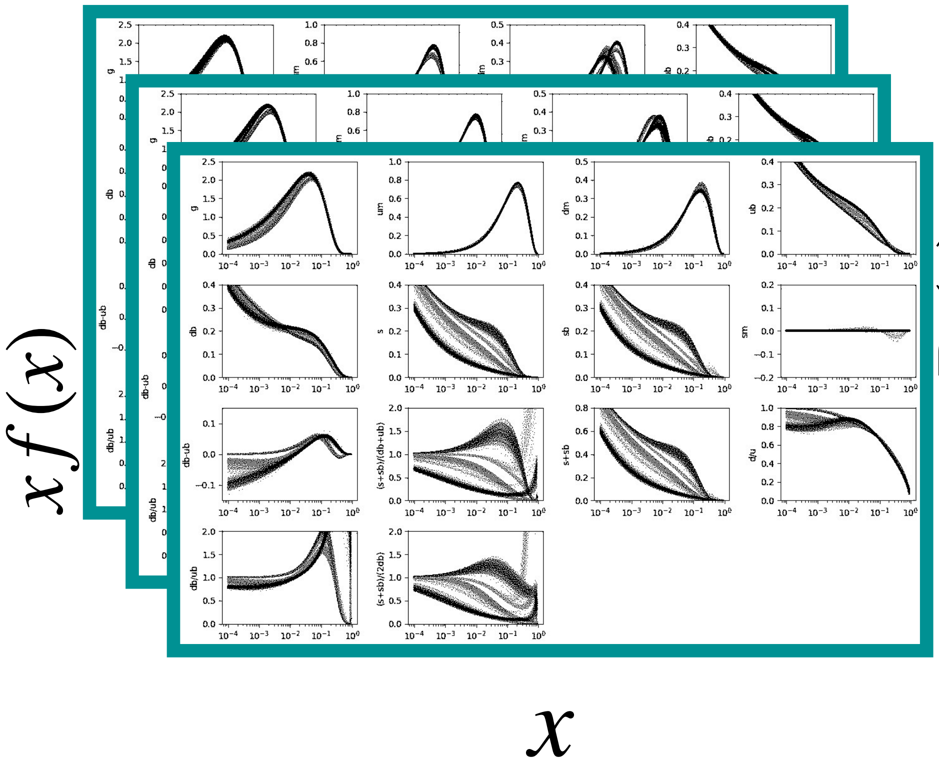
+ DIS + DY data

+ SIA pion data

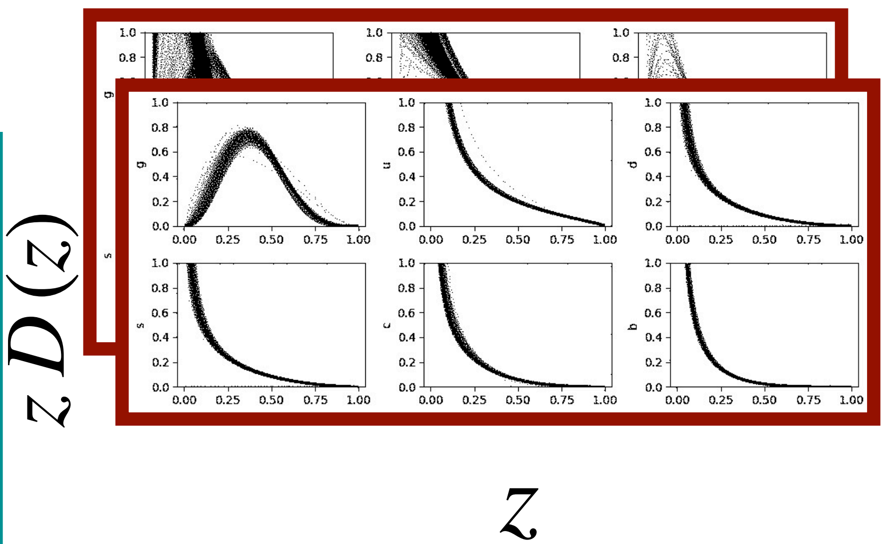
+ SIA kaon data

JAM19: multi-step fitting

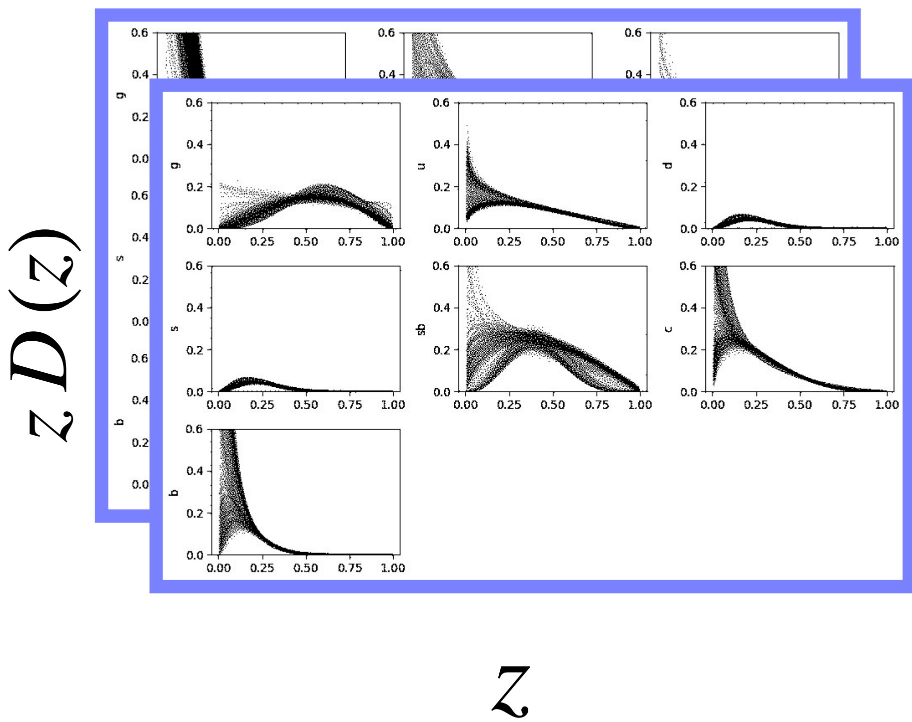
PDFs



PION FF



KAON FF



+ DIS data

+ DIS + DY data

+ SIDIS data

+ SIA pion data

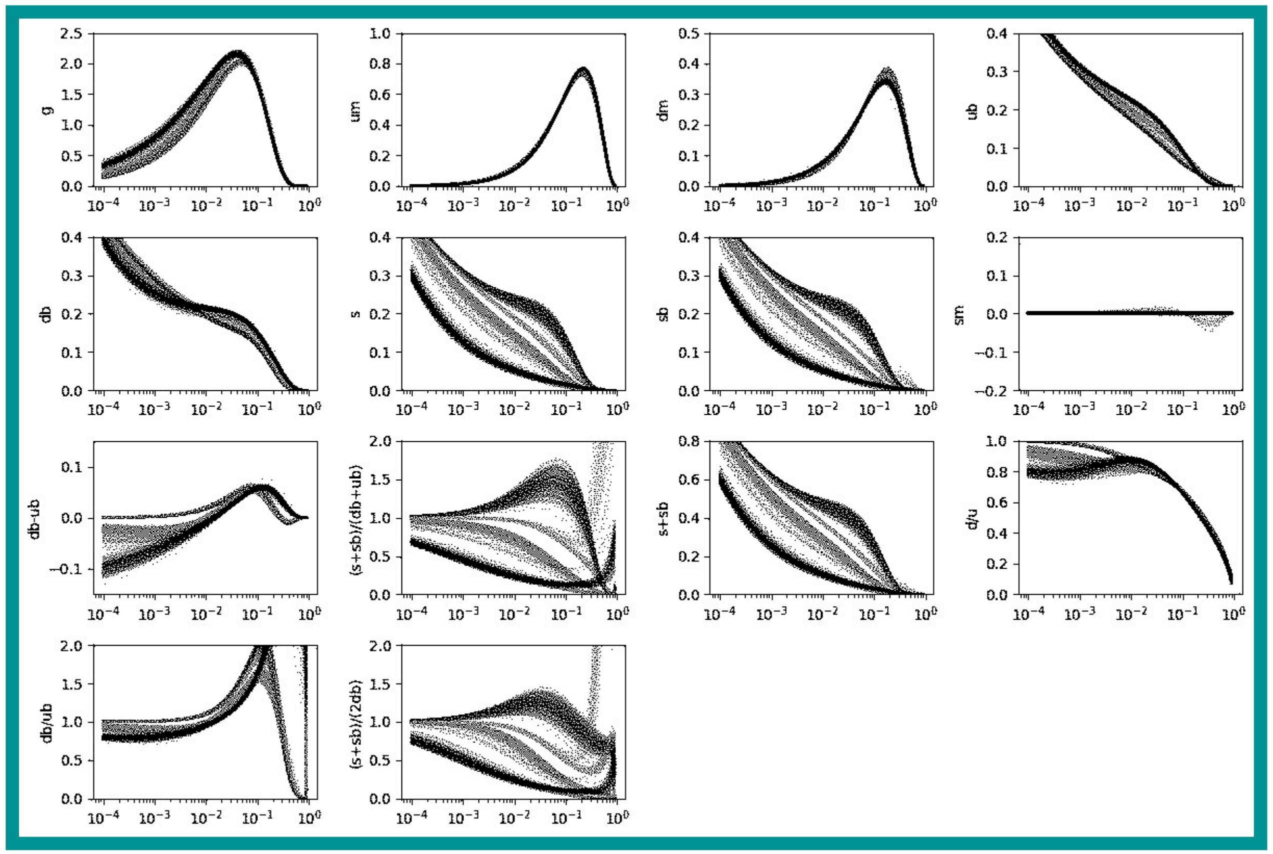
+ SIDIS pion data

+ SIA kaon data

+ SIDIS kaon data

Discrminating multiple solutions

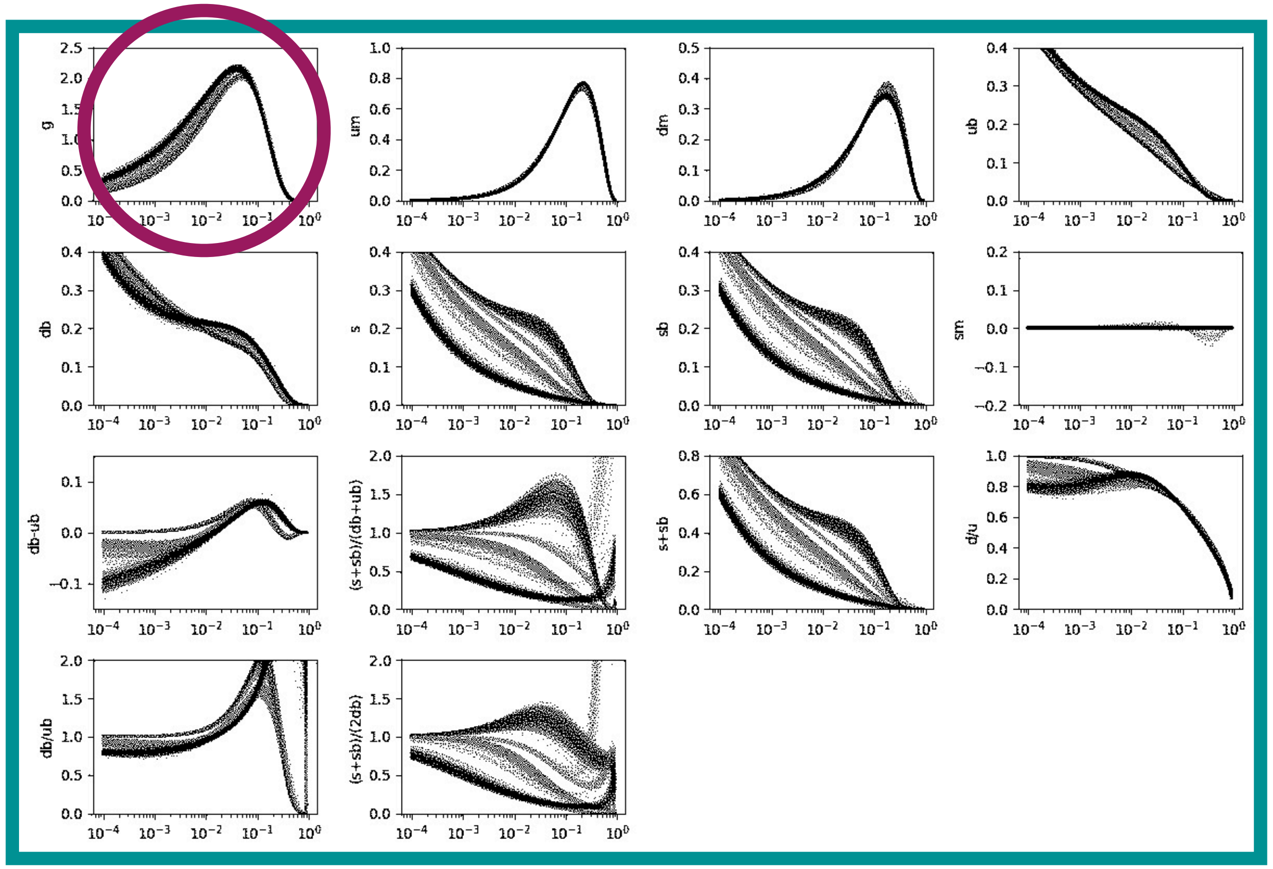
$x\,f(x)$



x

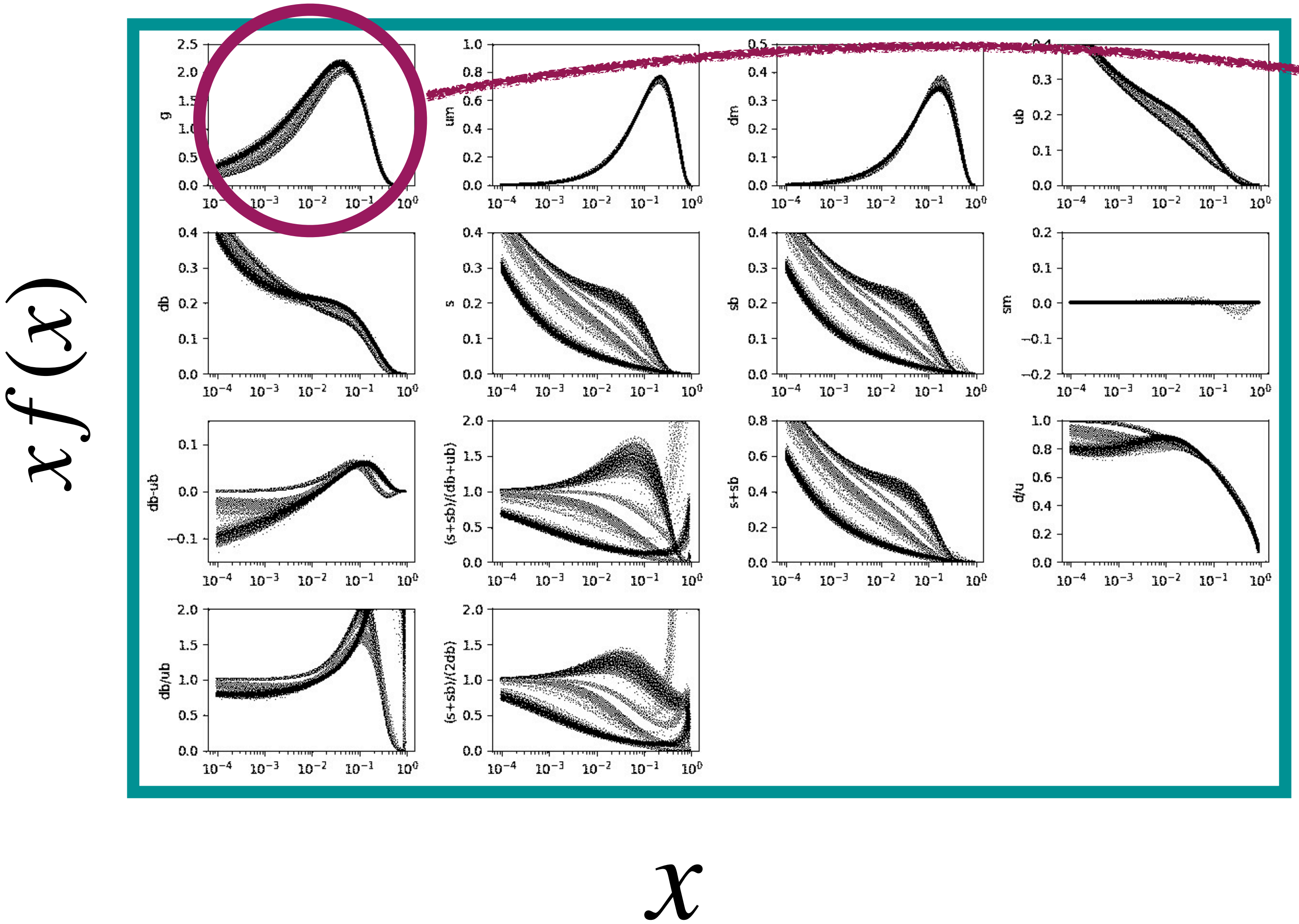
Discrminating multiple solutions

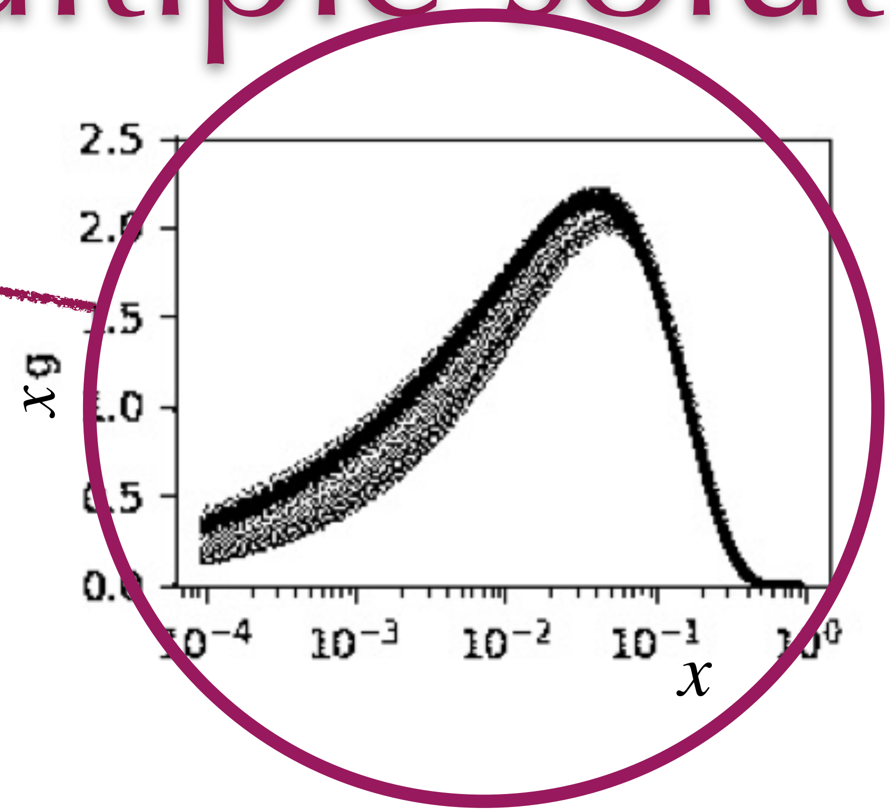
$x\,f(x)$



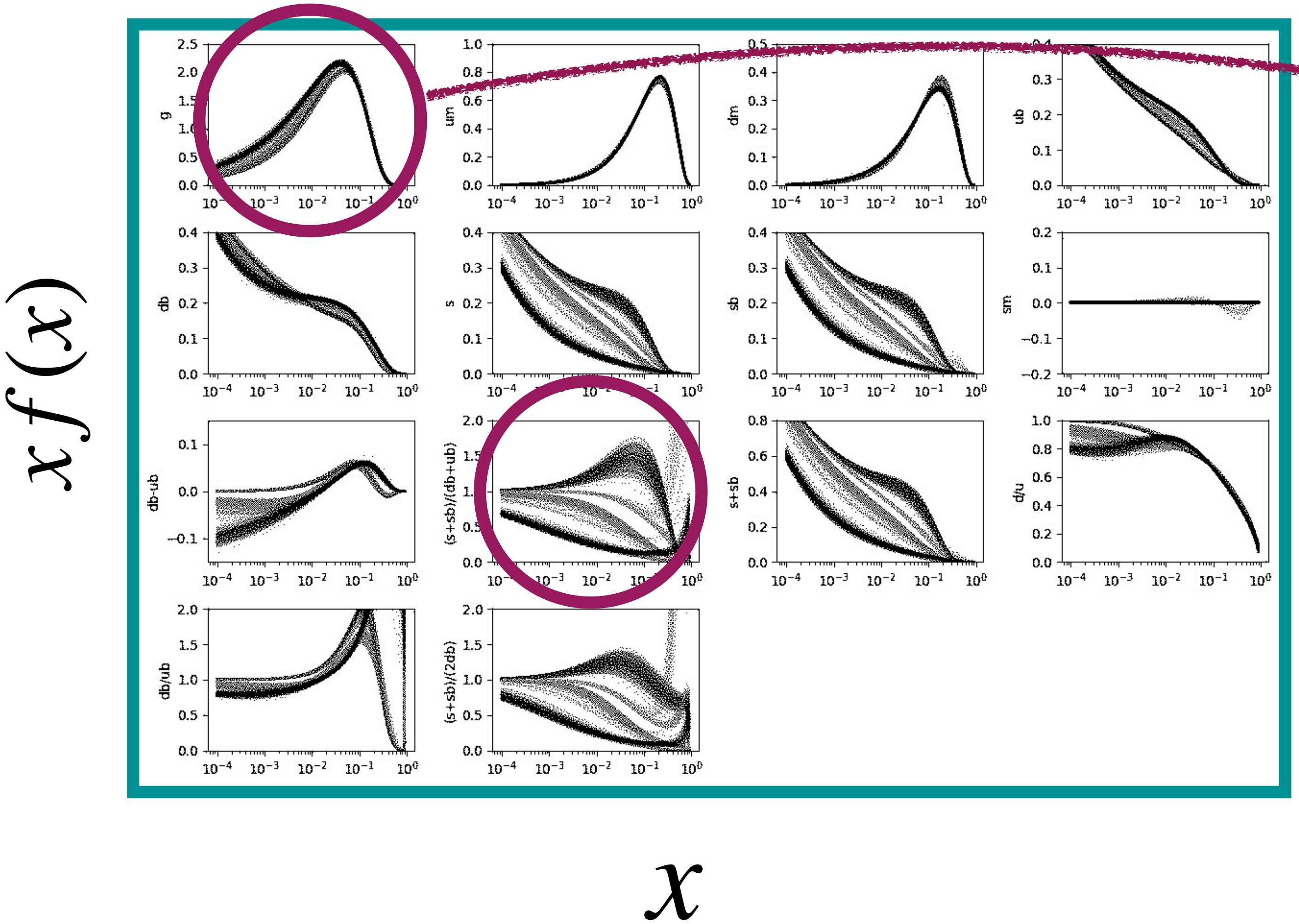
x

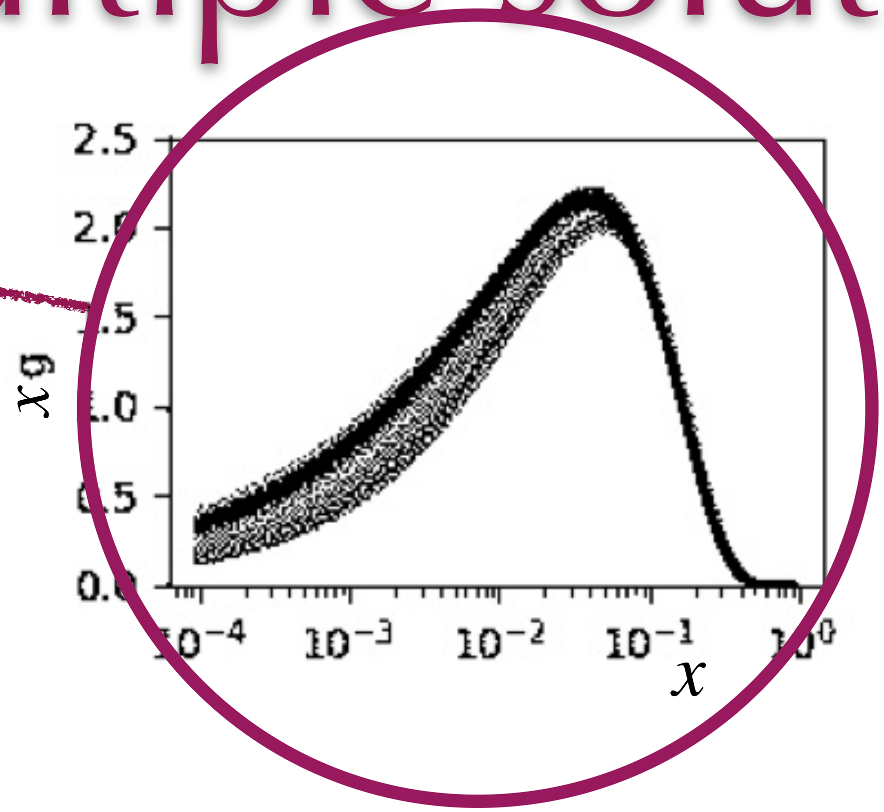
Discrminating multiple solutions



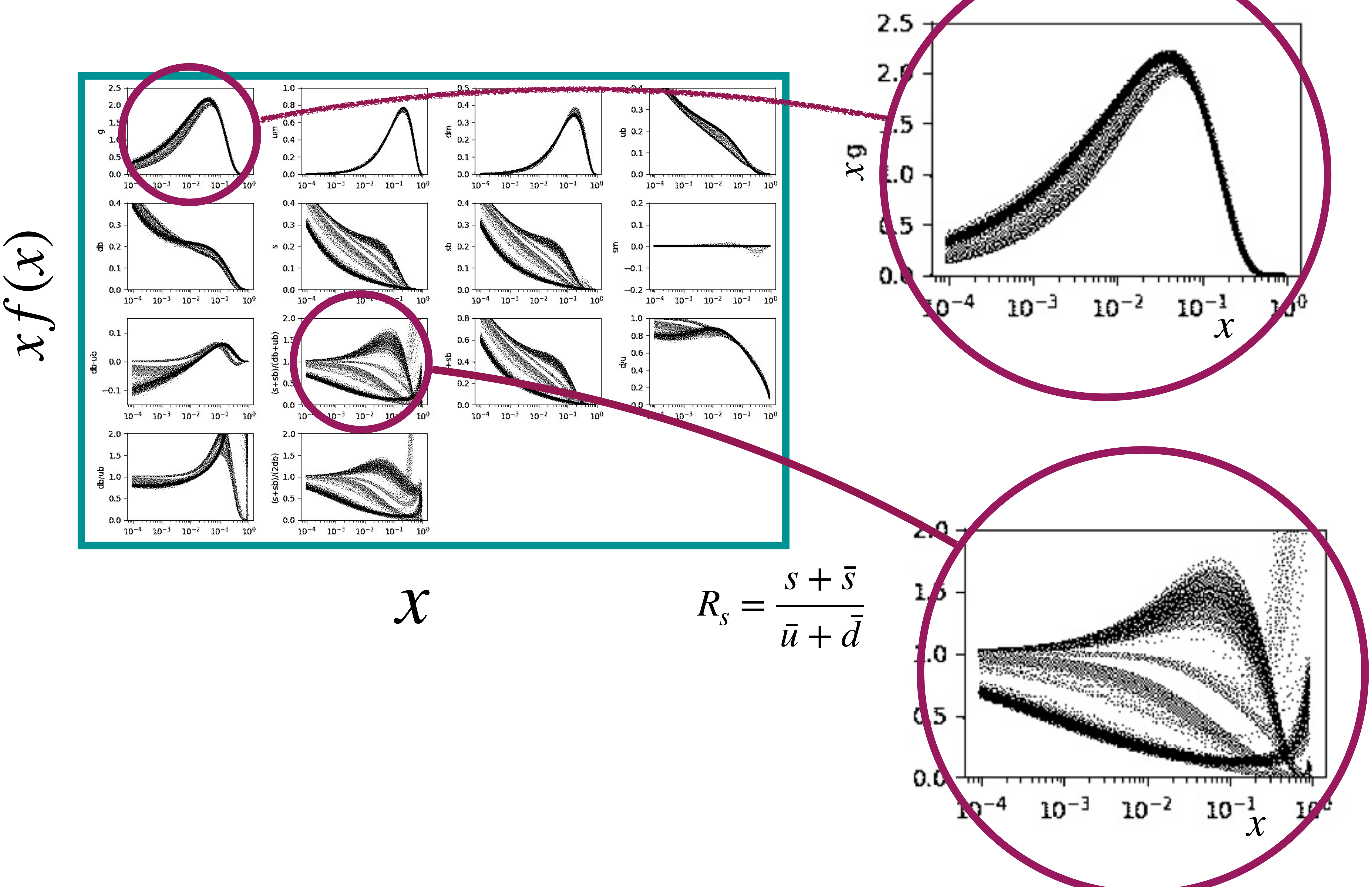


Discrminating multiple solutions





Discrminating multiple solutions

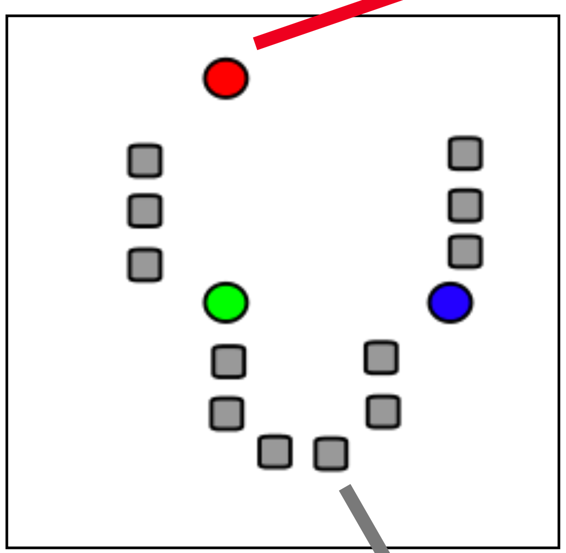


$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

k-means clustering

E.g. $f(x) = x^{\alpha}(1-x)^{\beta}$

(α^*, β^*) : centroid



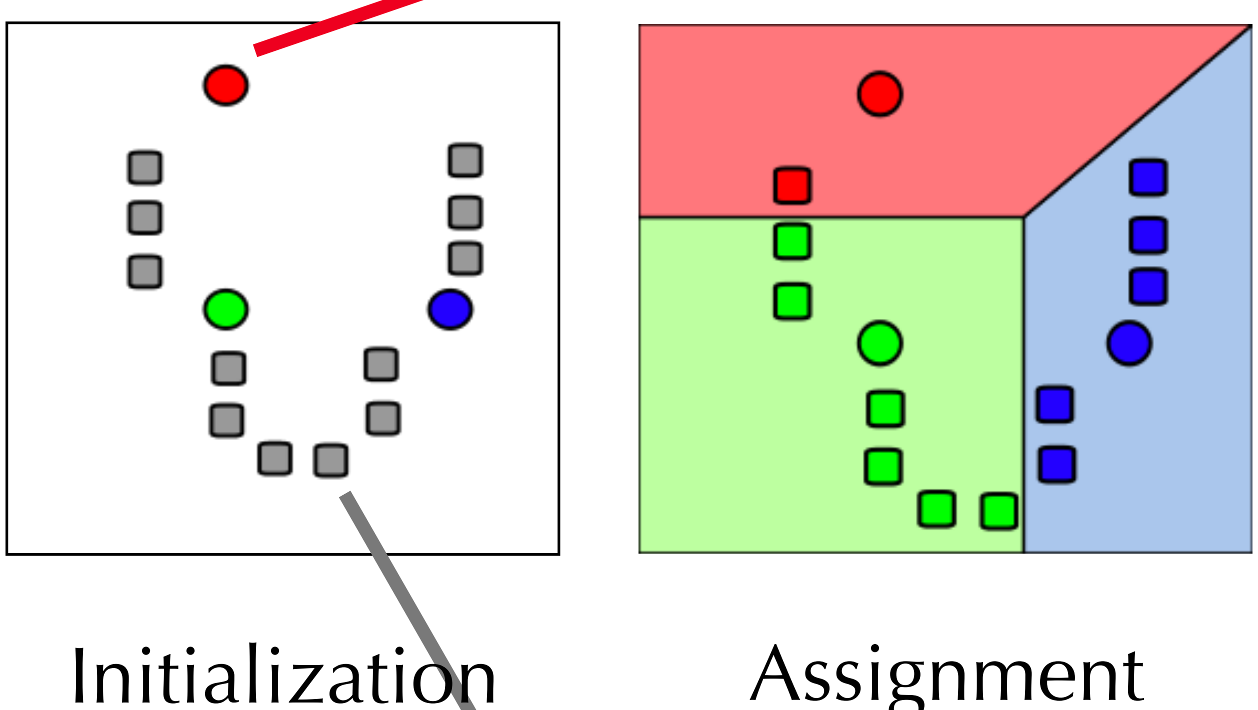
Initialization

(α_i, β_i) : replica

k-means clustering

E.g. $f(x) = x^{\alpha}(1-x)^{\beta}$

(α^*, β^*) : centroid

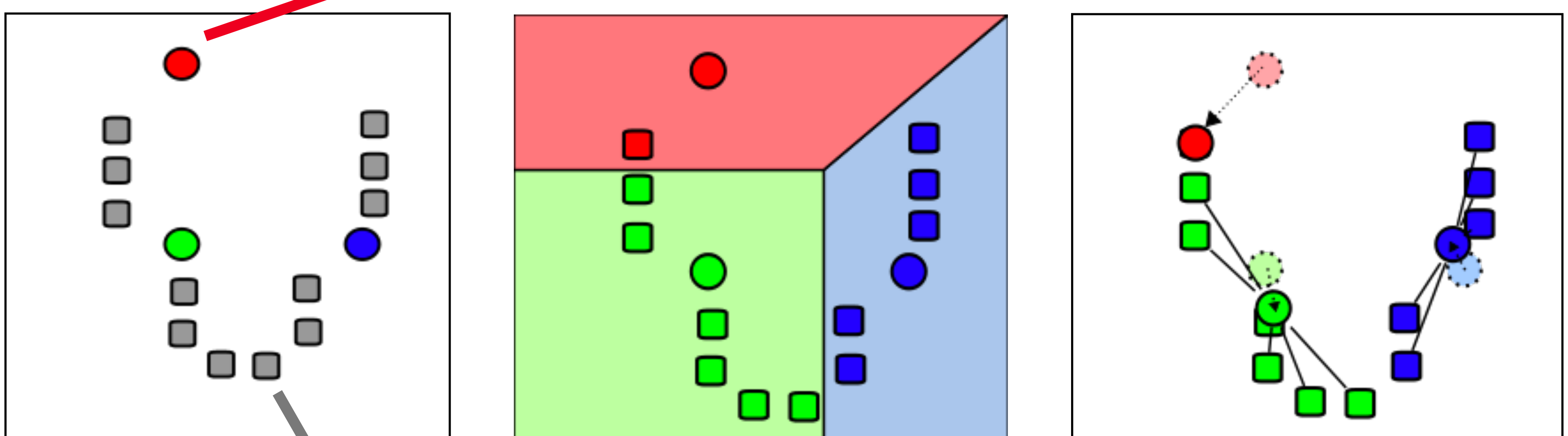


(α_i, β_i) : replica

k-means clustering

E.g. $f(x) = x^{\alpha}(1-x)^{\beta}$

(α^*, β^*) : centroid



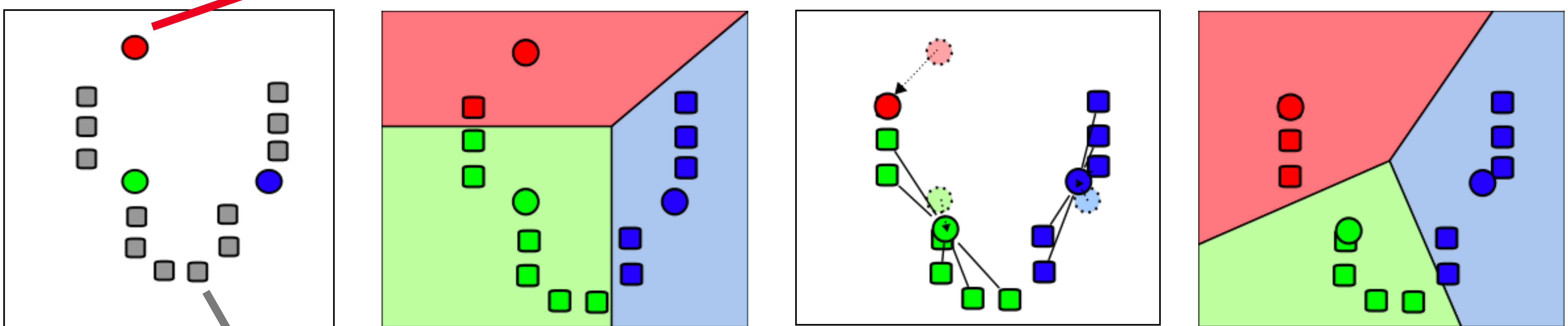
Initialization Assignment Update

(α_i, β_i) : replica

k-means clustering

E.g. $f(x) = x^{\alpha}(1-x)^{\beta}$

(α^*, β^*) : centroid



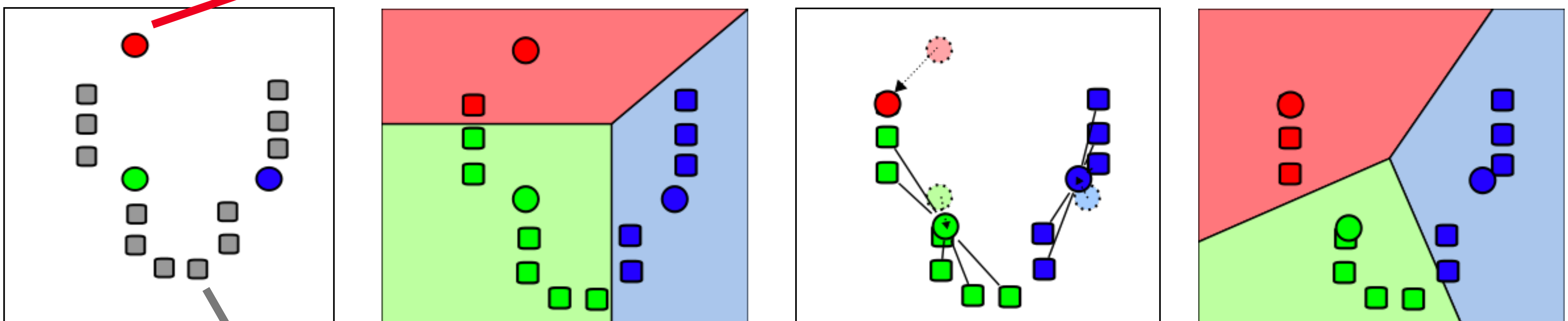
Initialization Assignment Update Assignment

(α_i, β_i) : replica

k-means clustering

E.g. $f(x) = x^{\alpha}(1-x)^{\beta}$

(α^*, β^*) : centroid

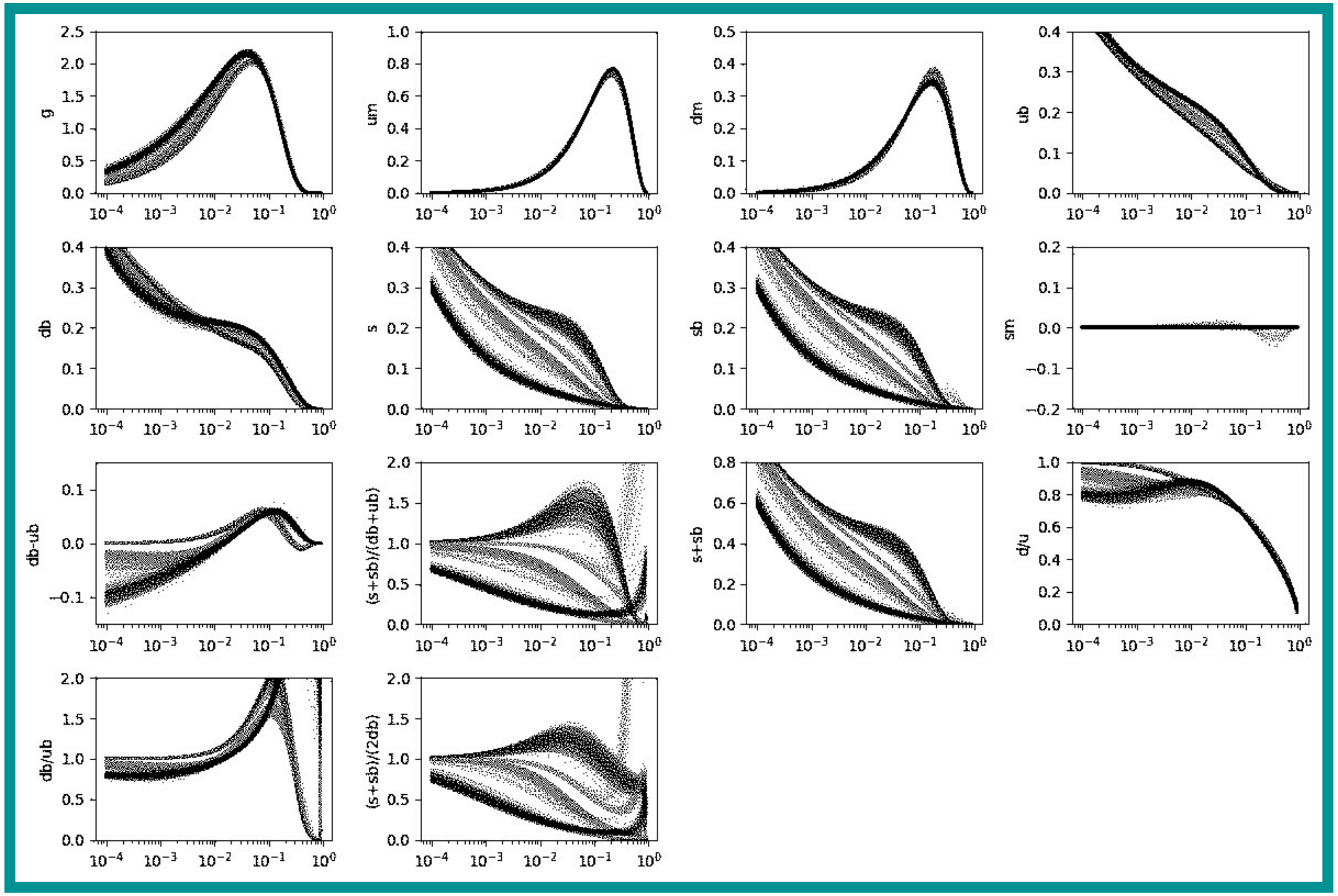


Initialization Assignment Update Assignment

Repeat until convergence

(α_i, β_i) : replica

Discriminating multiple solutions

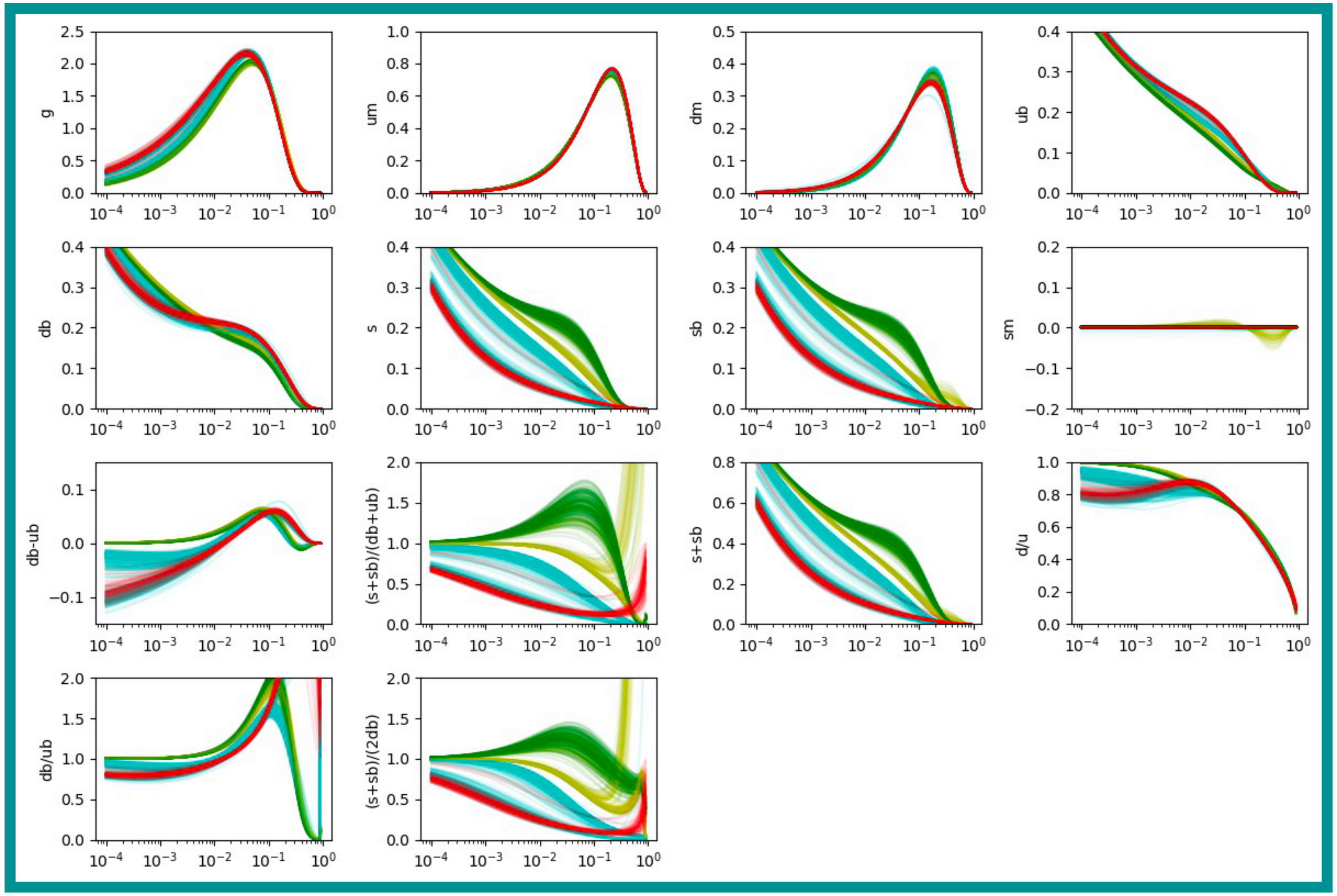


+ DIS data

+ DIS + DY data

+ SIDIS data

Discriminating multiple solutions



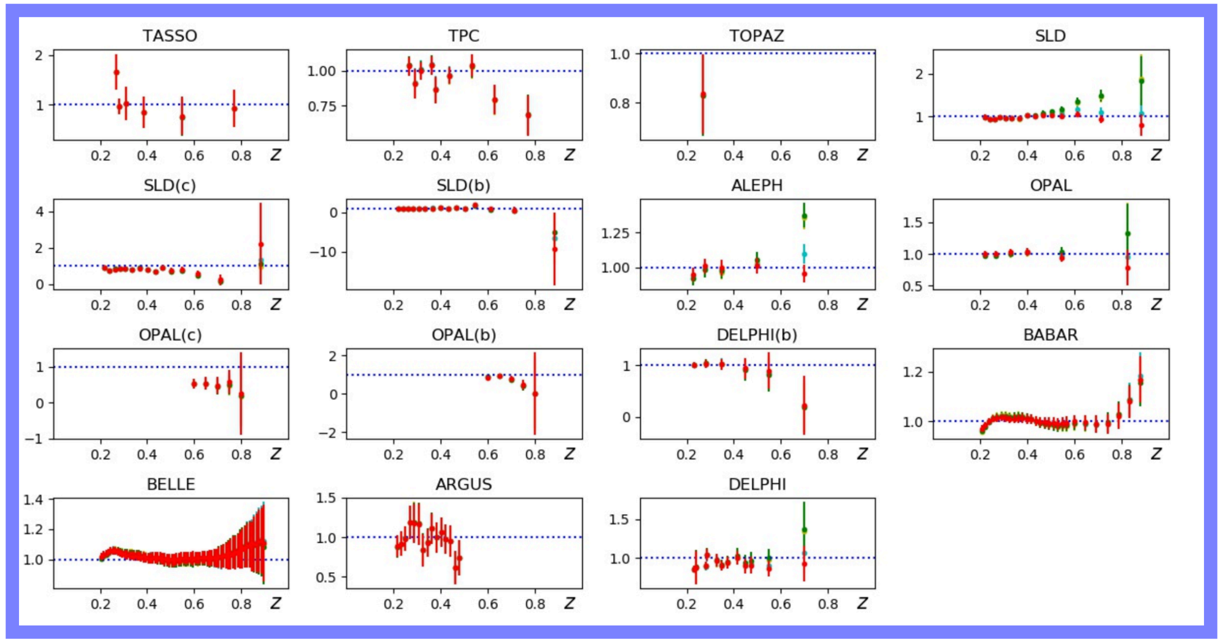
+ DIS data

+ DIS + DY data

+ SIDIS data

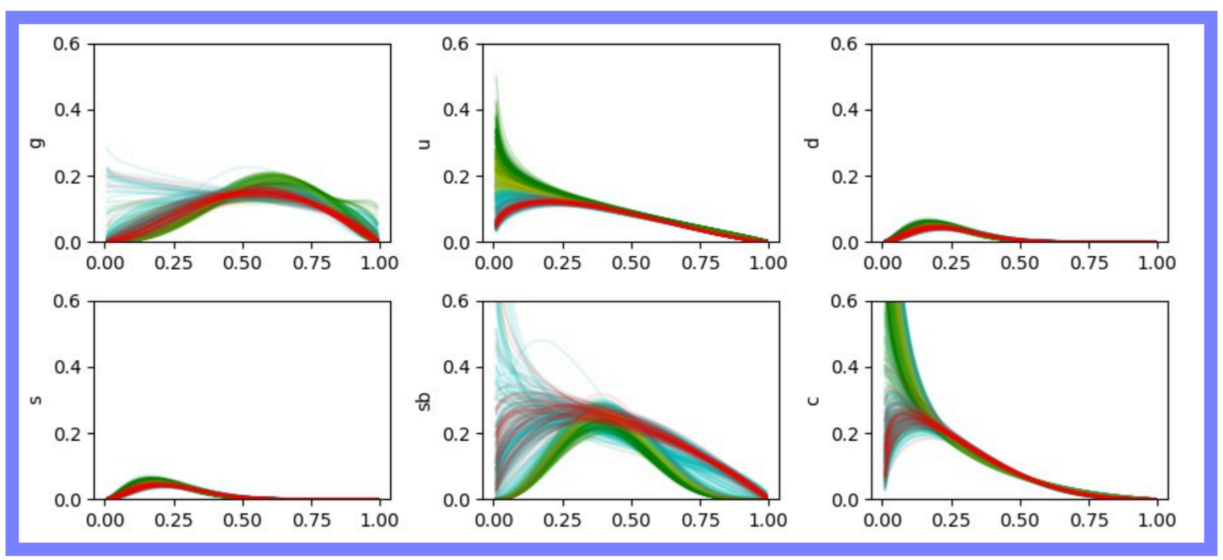
SIA K+/K- data

Data/Theory



z

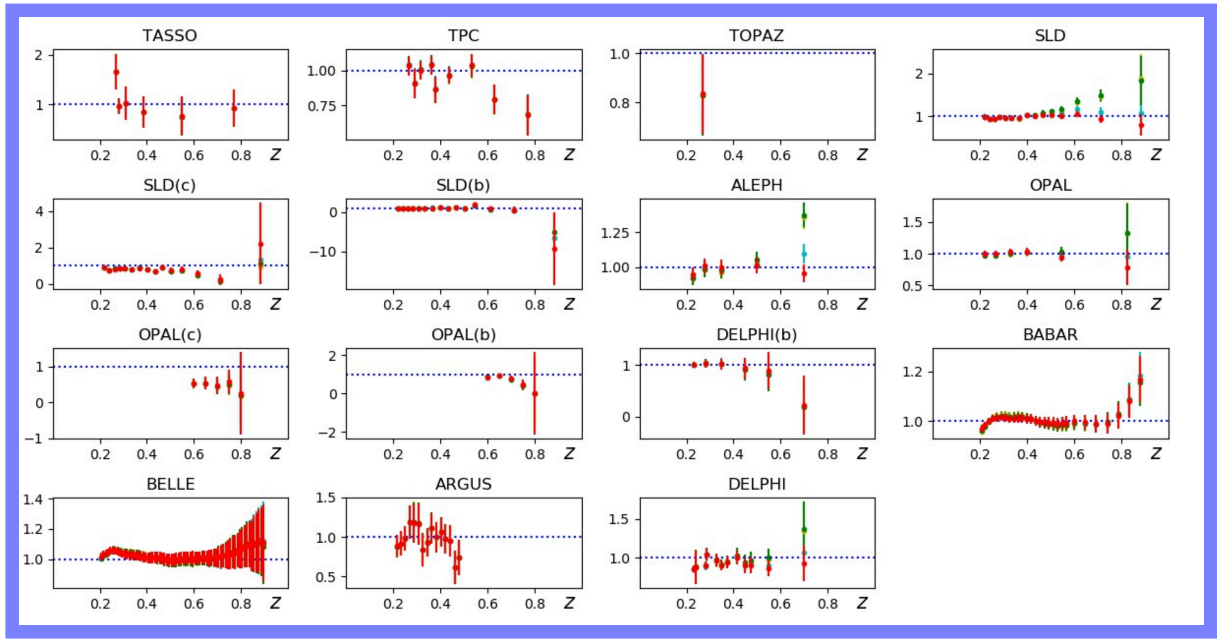
$zD_q^{K^+}$



z

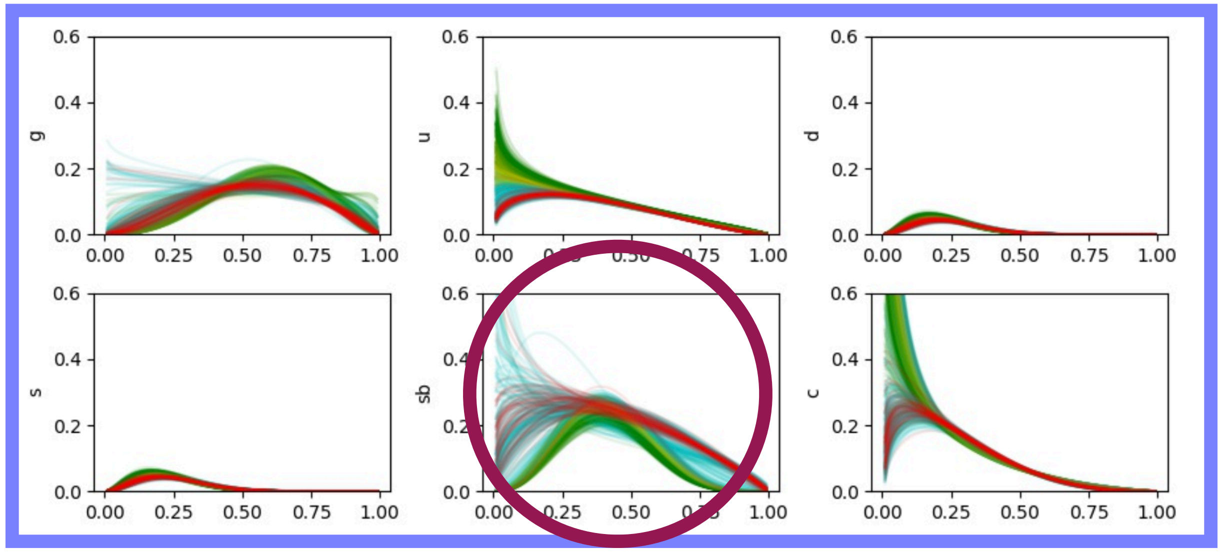
SIA K+/K- data

Data/Theory



z

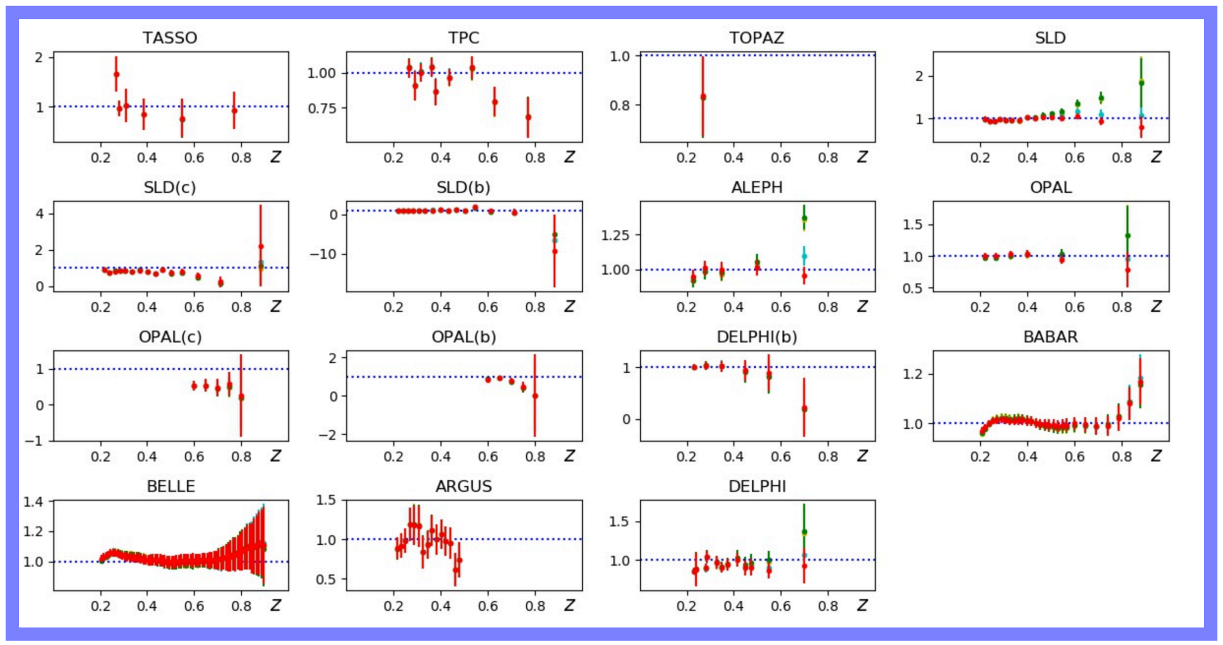
$zD_q^{K^+}$



z

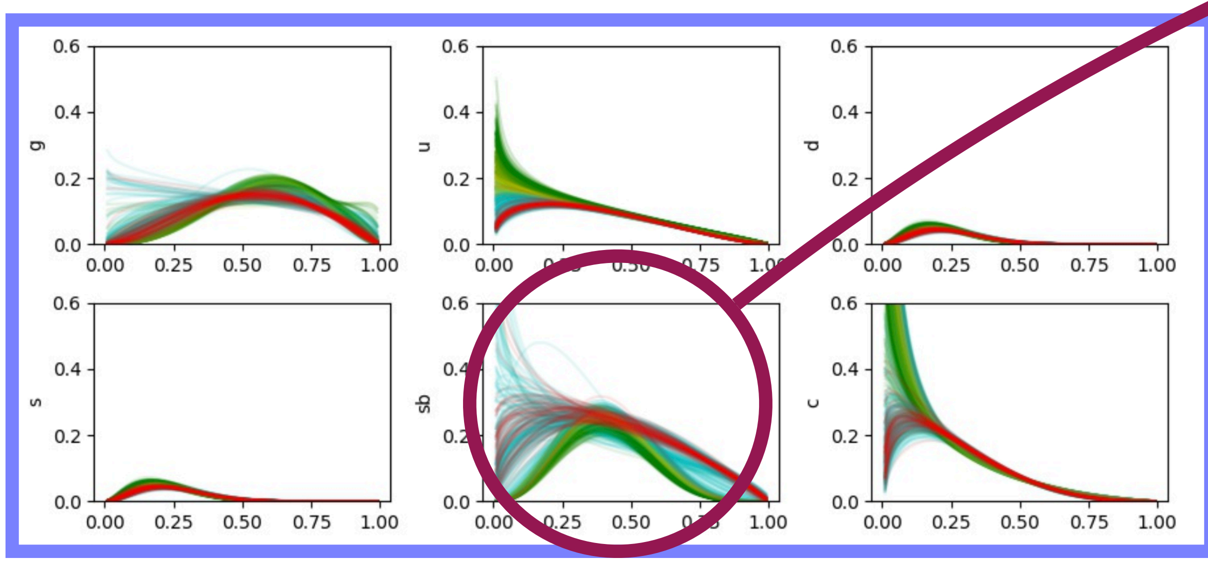
SIA K+/K- data

Data/Theory



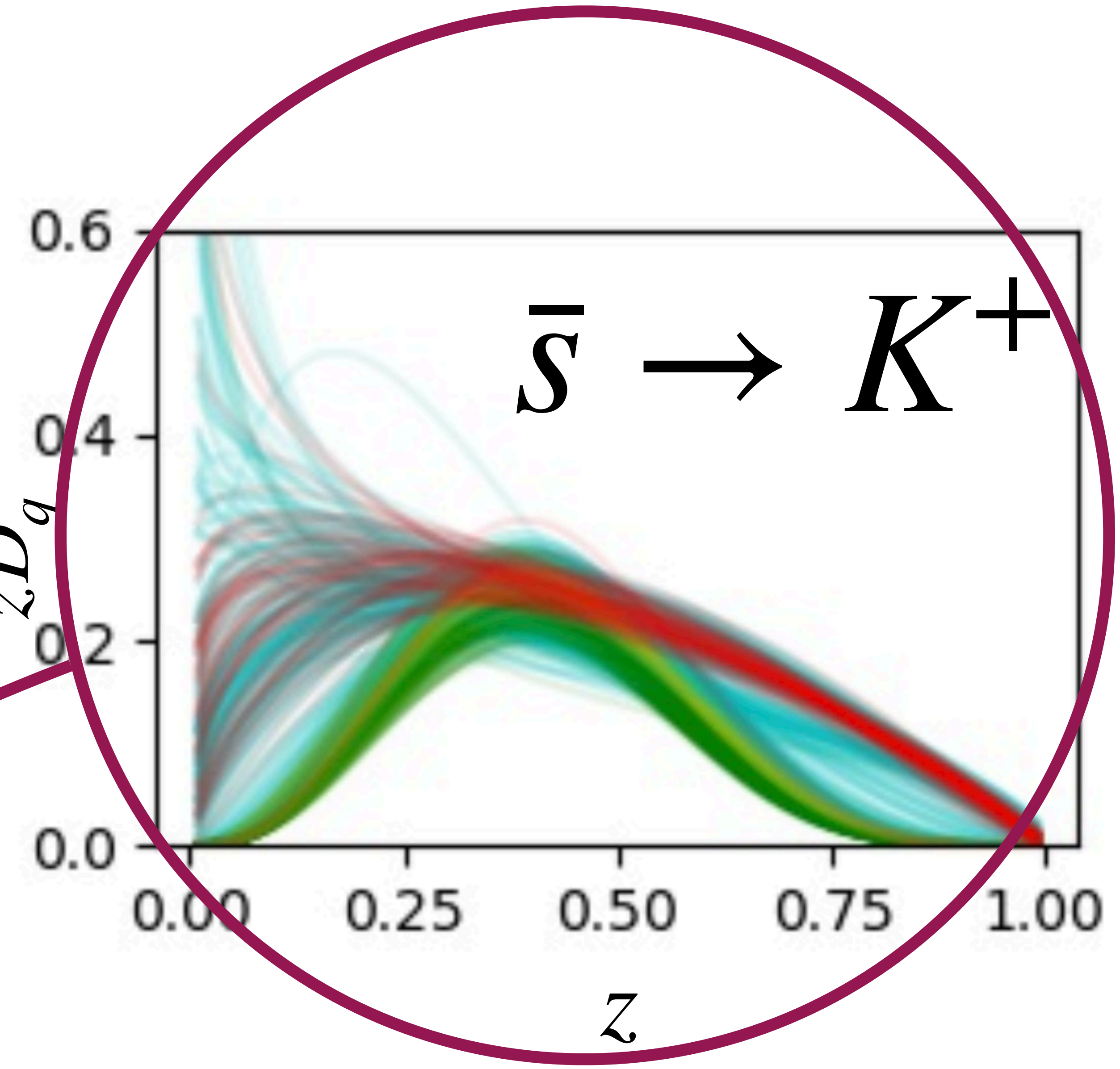
z

$zD_q^{K^+}$



z

$zD_q^{K^+}$



z

SIDIS K-

Unfavored solutions

Large $s(x)$

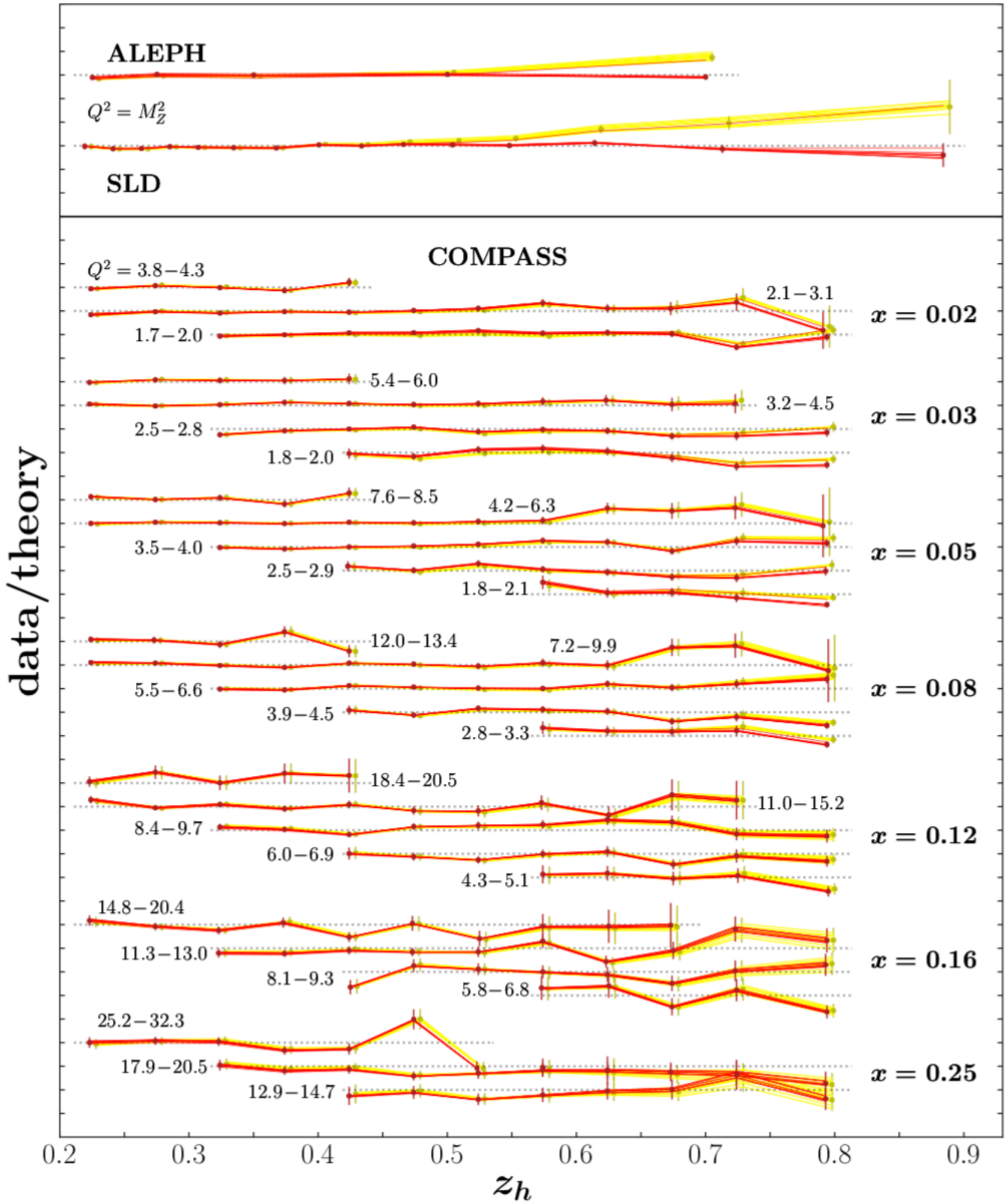
Small $D_{s^\pm}^{K^\pm}(z)$

Favored solutions

Large $D_{s^\pm}^{K^\pm}(z)$

Small $s(x)$

$\chi^2_{\mathrm{SLD}} = 4.10$	$\chi^2_{\mathrm{SLD}} = 1.38$
$\chi^2_{\mathrm{ALEPH}} = 4.62$	$\chi^2_{\mathrm{ALEPH}} = 0.34$

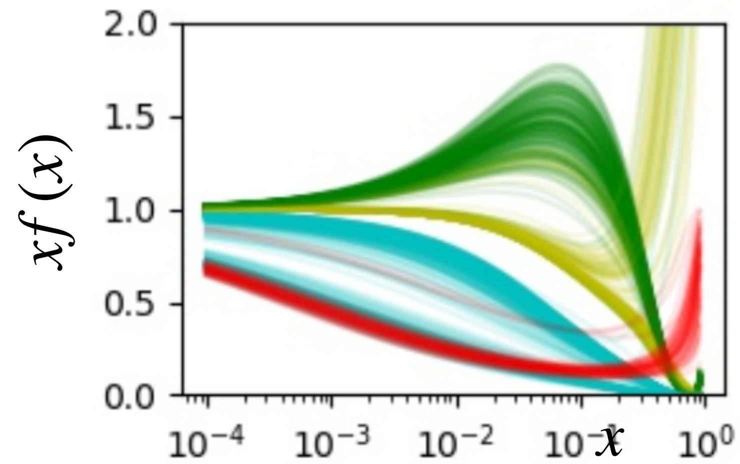


SIA

SIDIS

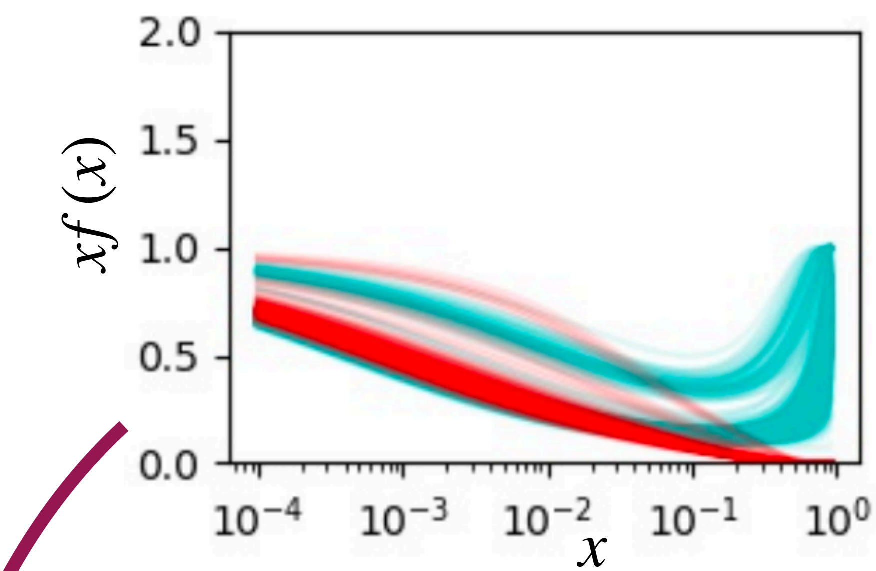
JAM19: Selection criteria

- Apply k-means clustering

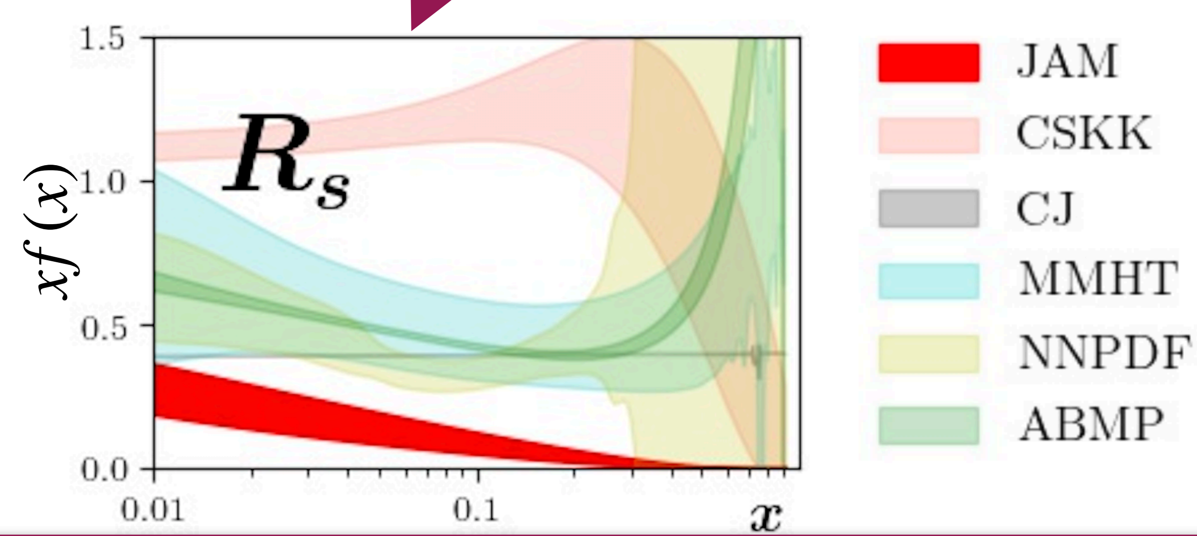


- Classify clusters by increasing χ^2 order in 'extended' reduced

$$\frac{\chi^2}{N_{\text{tot}}} + \sum_{exp} \frac{\chi^2_{\text{exp}}}{N_{\text{exp}}}$$



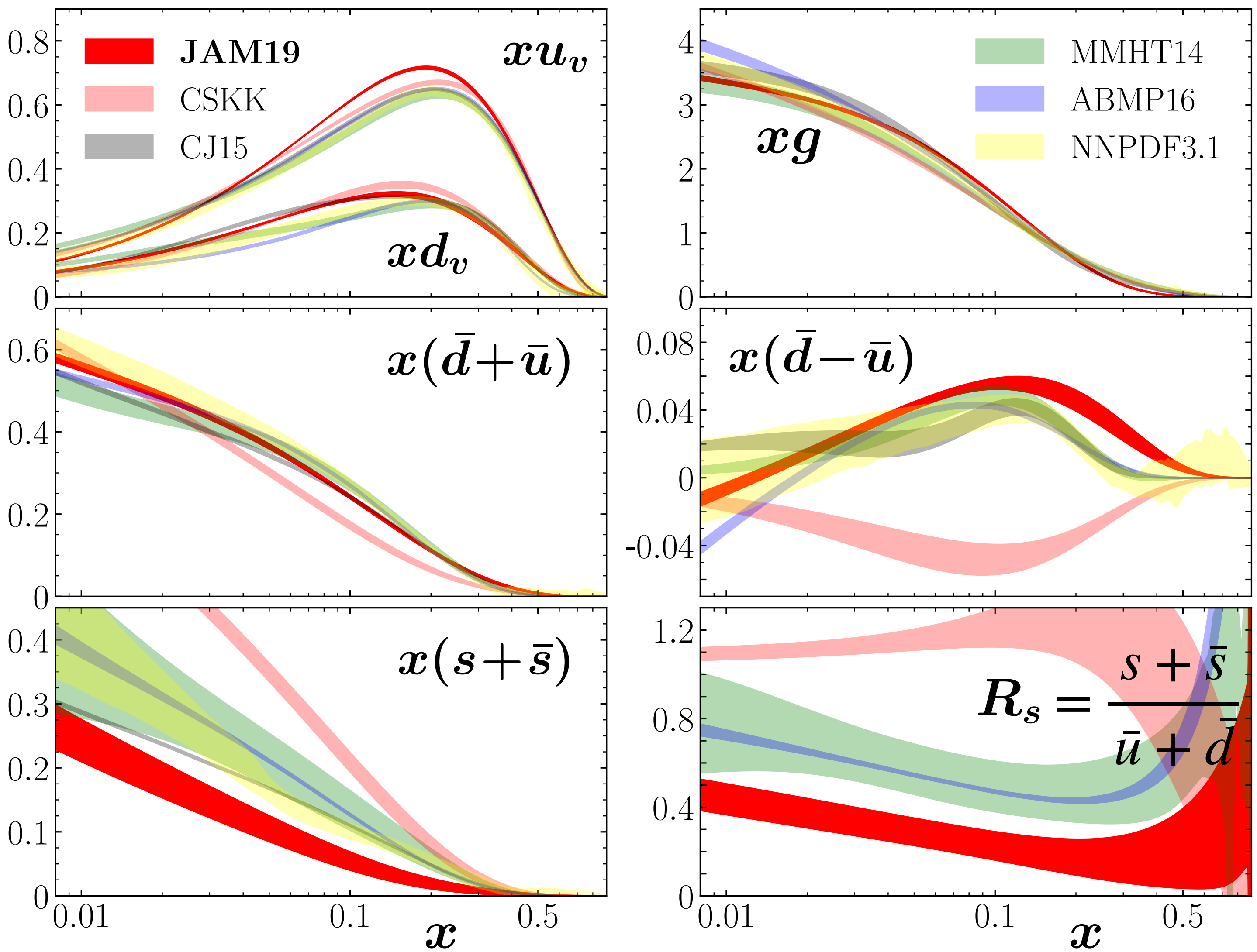
- Perform a new sampling with flat priors around the best cluster



PDF results

JAM19 PDFs

arXiv:1905.03788 [hep-ph]



$\mathrm{DIS}(p, d)$

$\mathrm{DY}(pp, pd)$

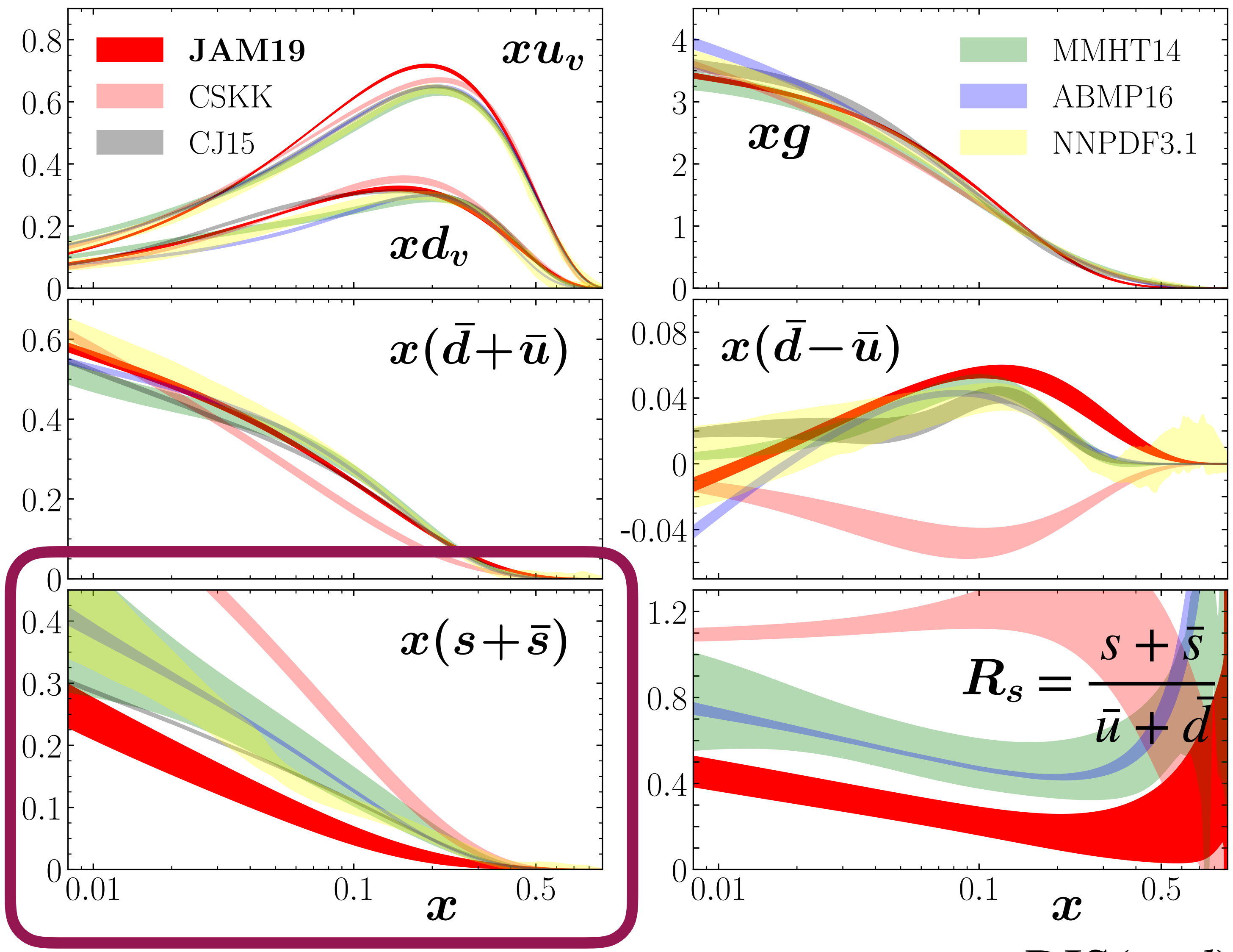
$\mathrm{SIA}(\pi^\pm, K^\pm)$

$\mathrm{SIDIS}(\pi^\pm, K^\pm)$

Q = 2 GeV

JAM19 PDFs

arXiv:1905.03788 [hep-ph]

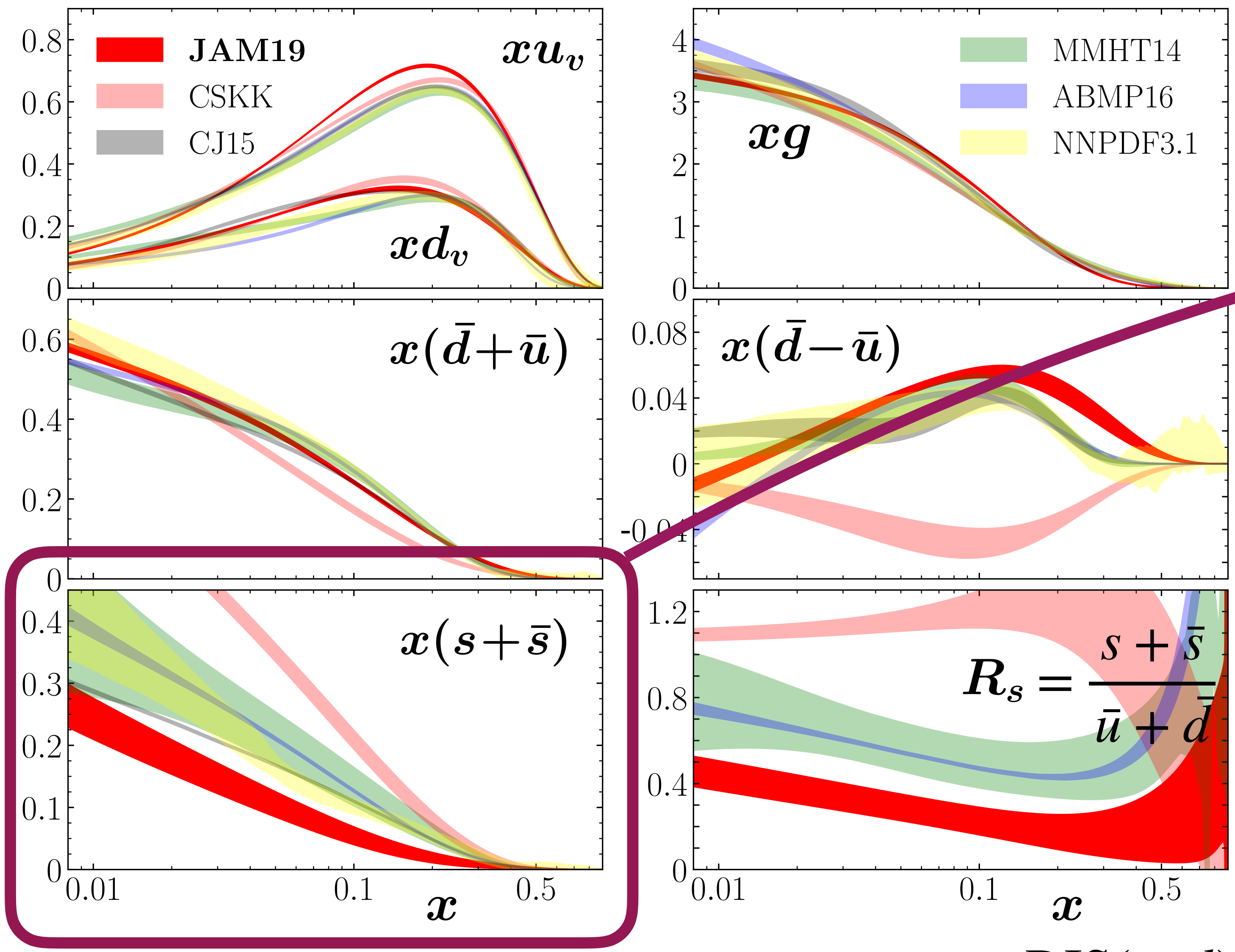


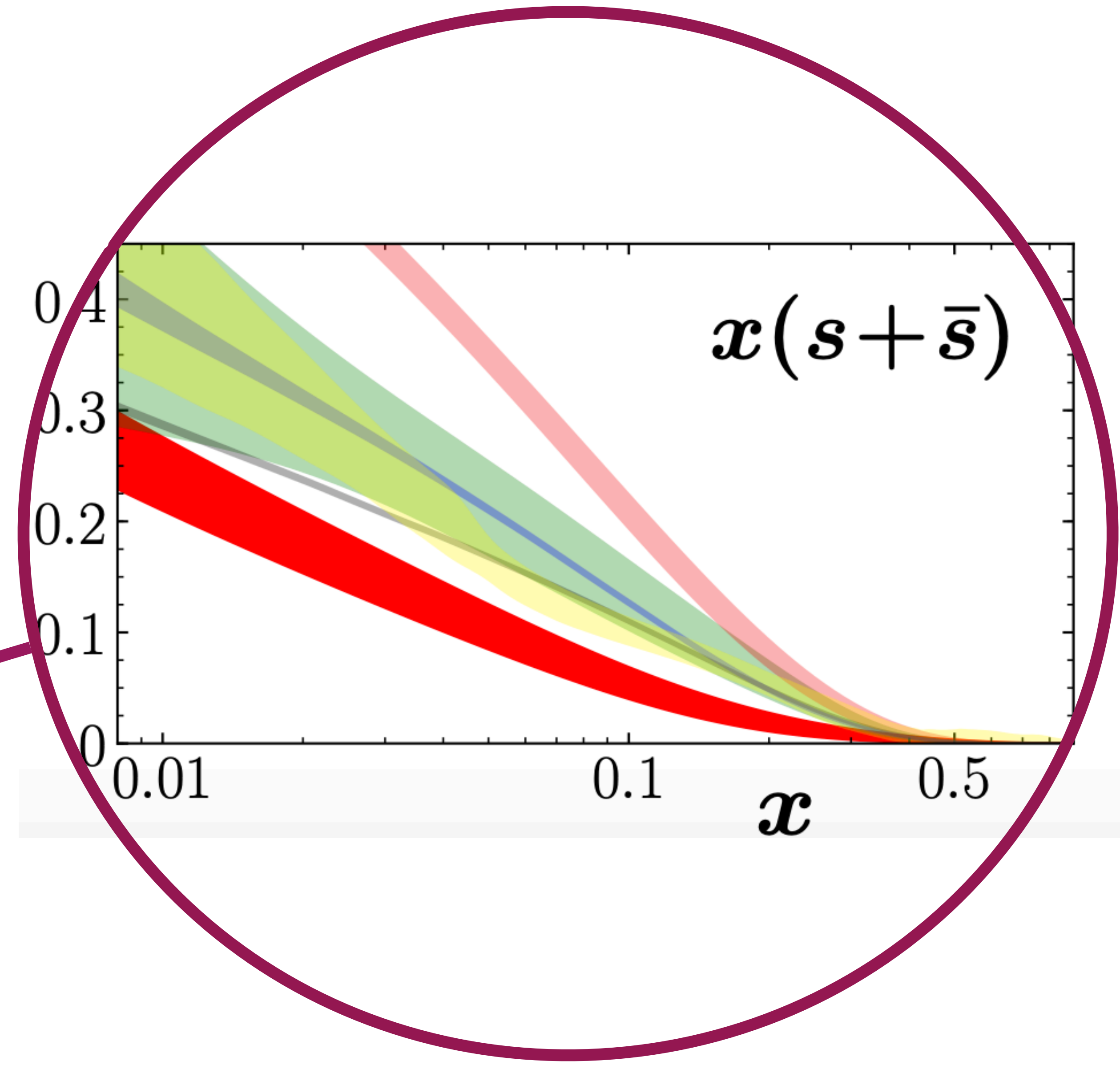
DIS(p, d)
DY(pp, pd)
SIA$(\pi^{\pm}, K^{\pm})$
SIDIS$(\pi^{\pm}, K^{\pm})$

Q = 2 GeV

JAM19 PDFs

arXiv:1905.03788 [hep-ph]





Strong strange suppression

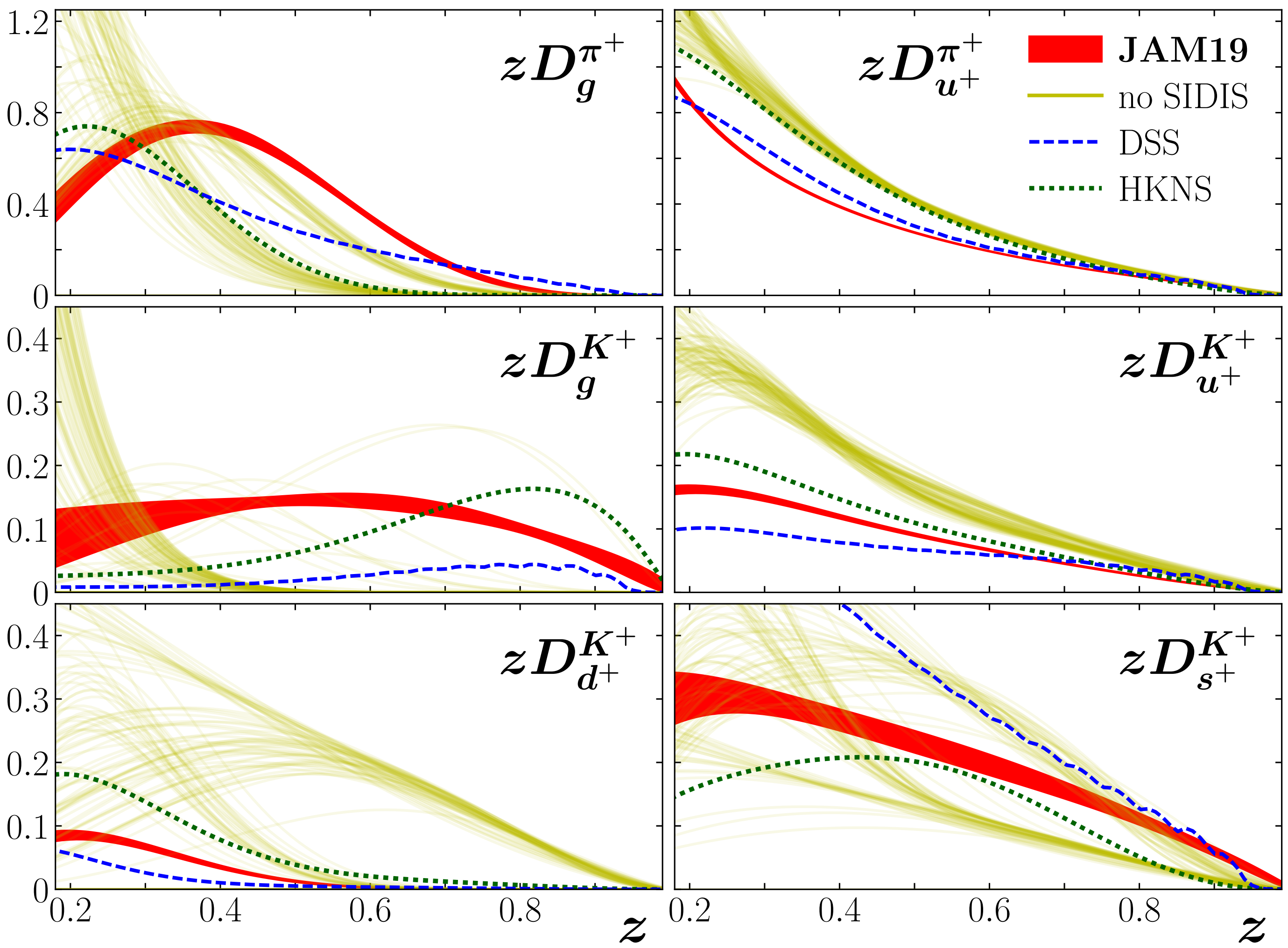
Q = 2 GeV

DIS(p, d)
DY(pp, pd)
SIA($\pi^{\pm}, K^{\pm}$)
SIDIS($\pi^{\pm}, K^{\pm}$)

FF results

JAM19: FF

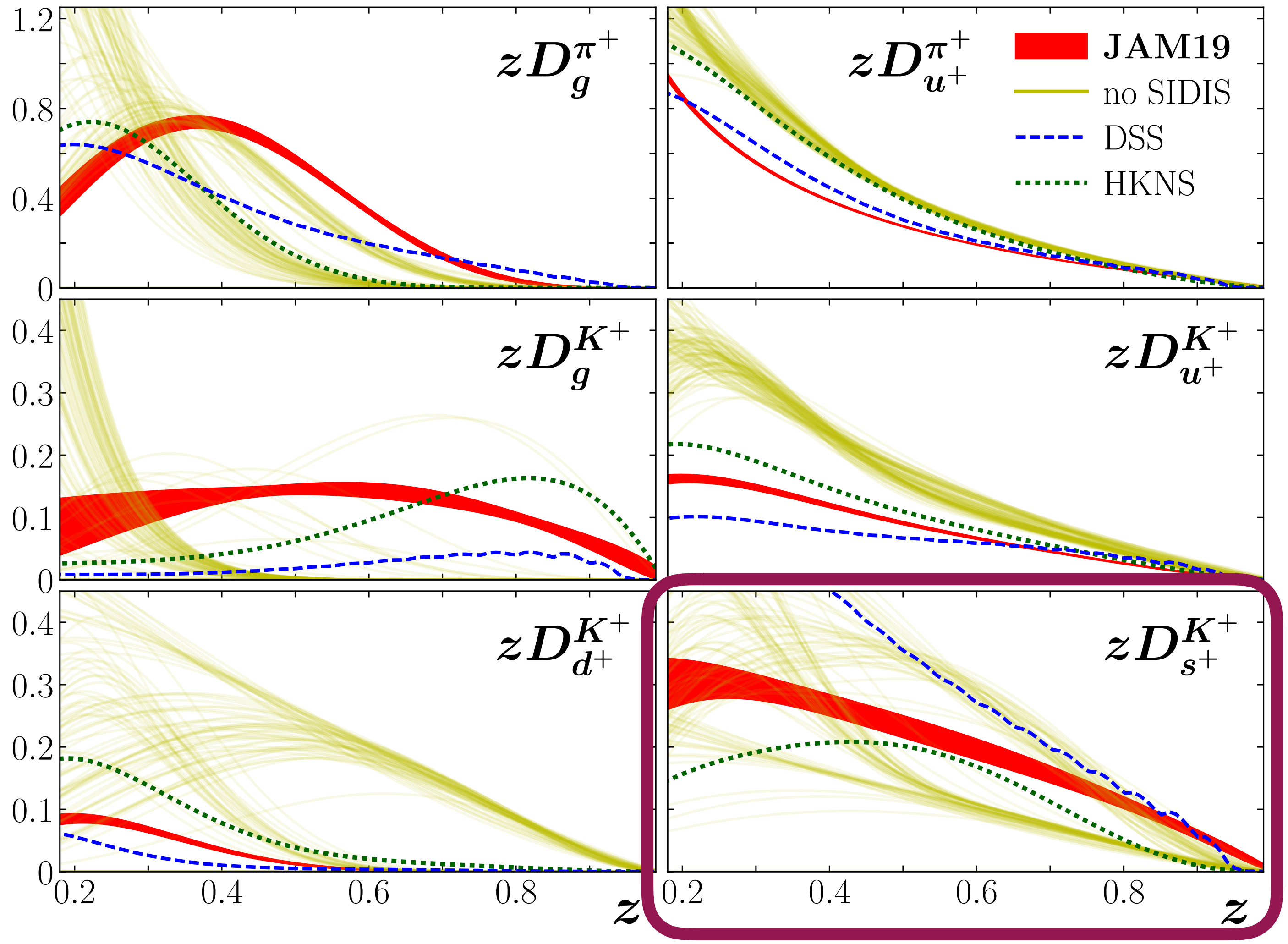
arXiv:1905.03788 [hep-ph]



Q = m$_c$

JAM19: FF

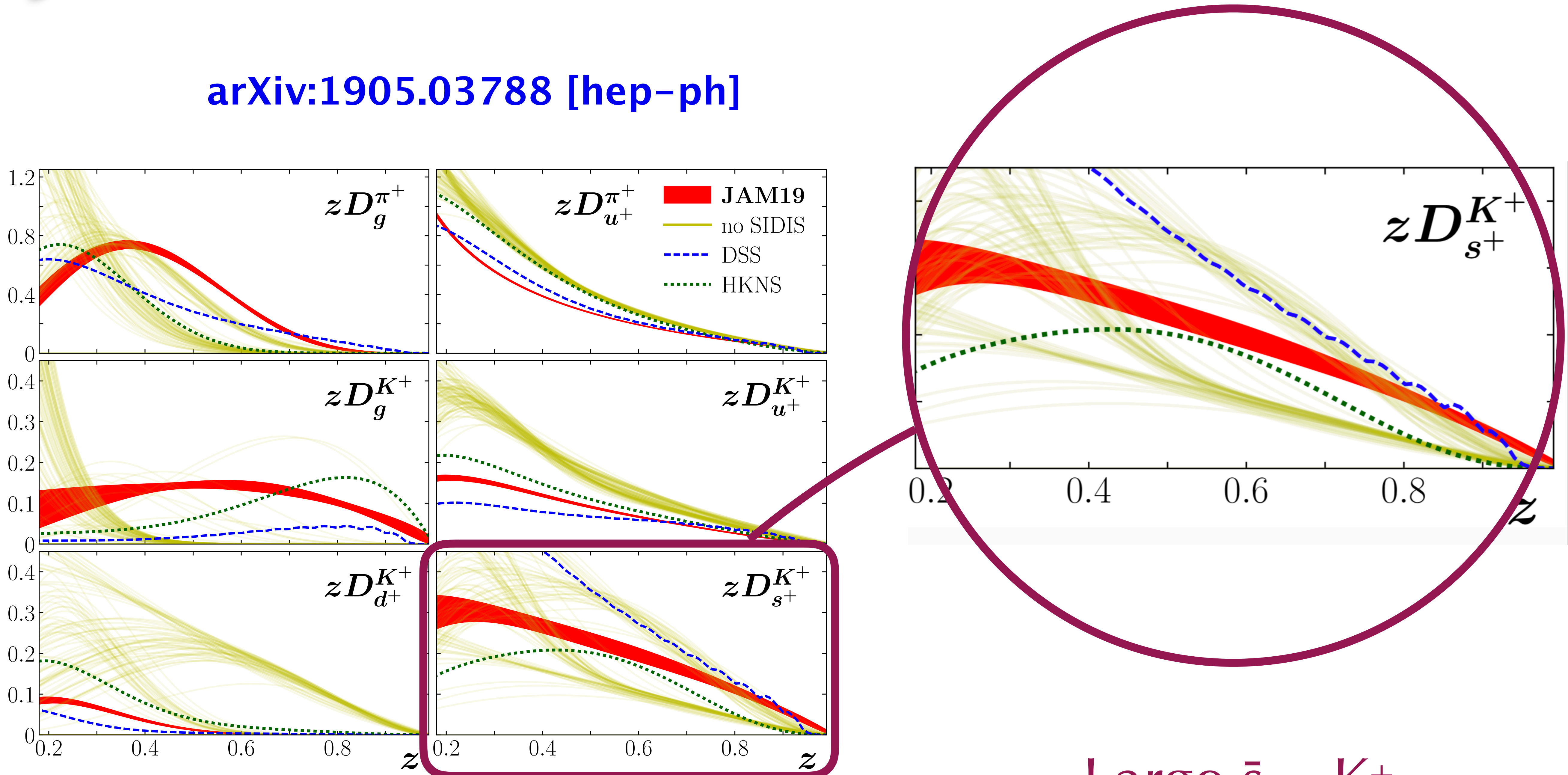
arXiv:1905.03788 [hep-ph]



$$Q = m_c$$

JAM19: FF

arXiv:1905.03788 [hep-ph]



Large $\bar{s} \rightarrow K^+$

$Q = m_c$

Summary

- **Fitting** several (or all) non-perturbative collinear distributions **simultaneously is very challenging**

- **MC** statistical methods are important for a **robust extraction**

  (Crucial for future global TMDs, GPDs analysis)

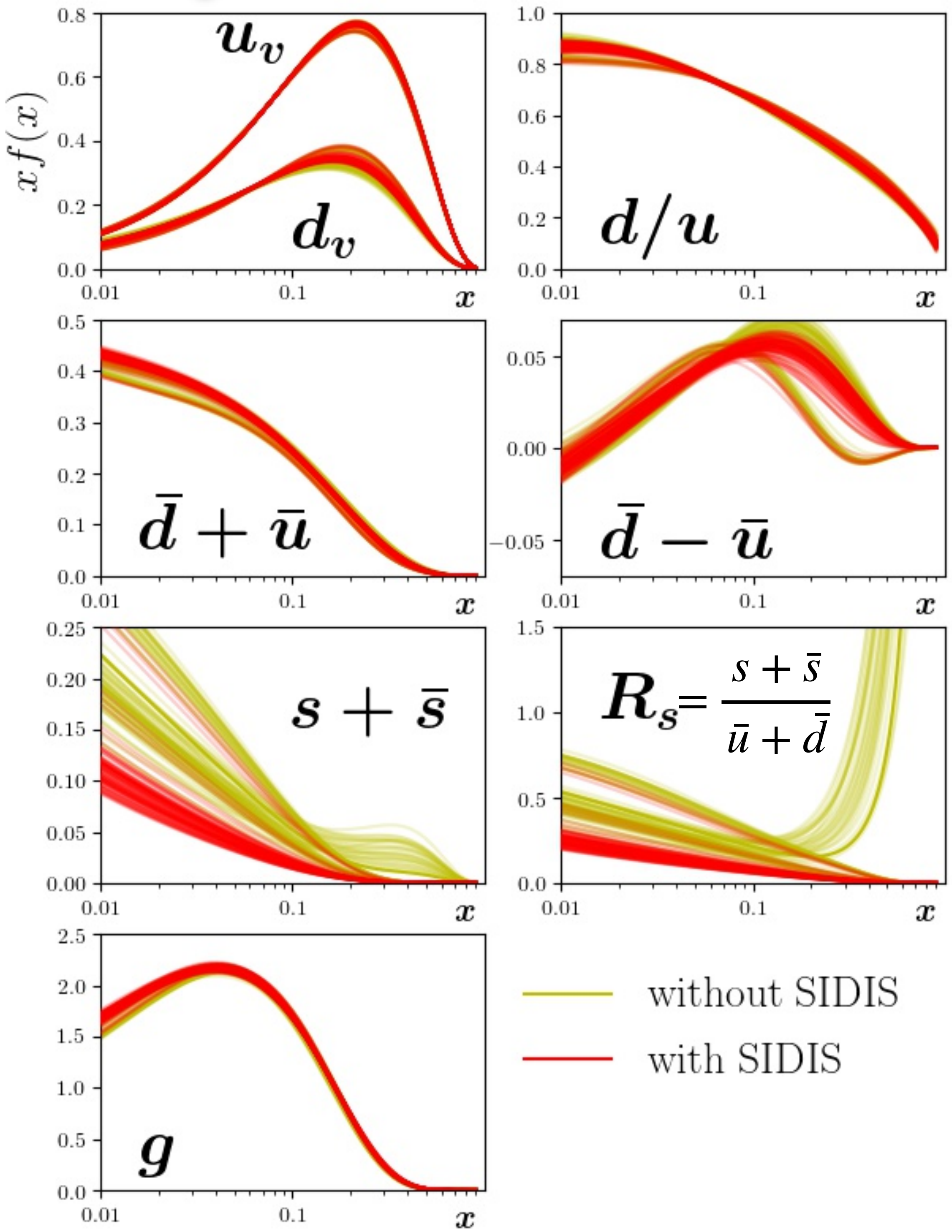
- **New methodology** needed: MC **multi-steps fit**, k-means clustering, 'extended' reduced χ^2

- First simultaneous fit of unpolarized PDFs and FFs: s**trange** PDF **strongly suppressed**

Backup

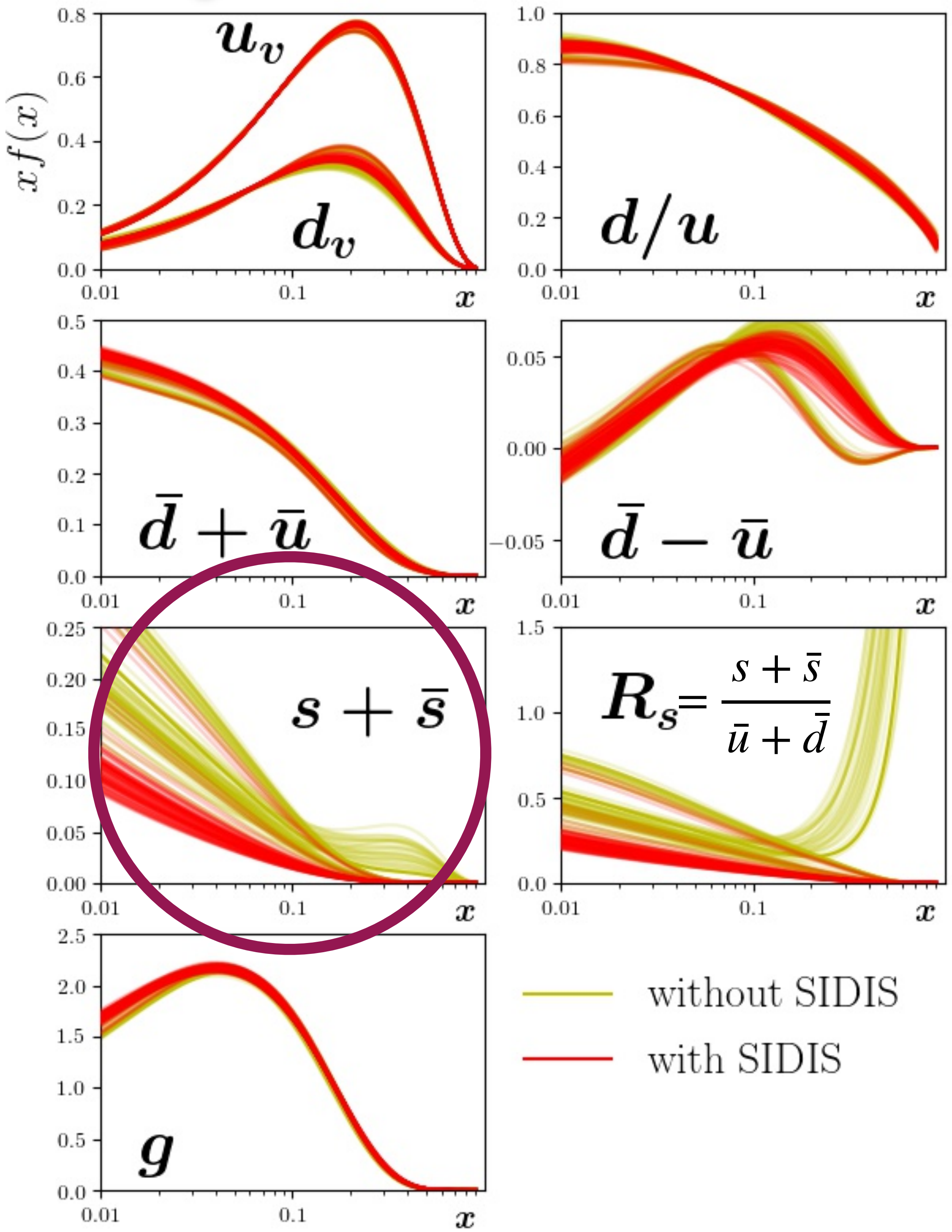
Impact of SIDIS data

Impact of SIDIS data on PDFs



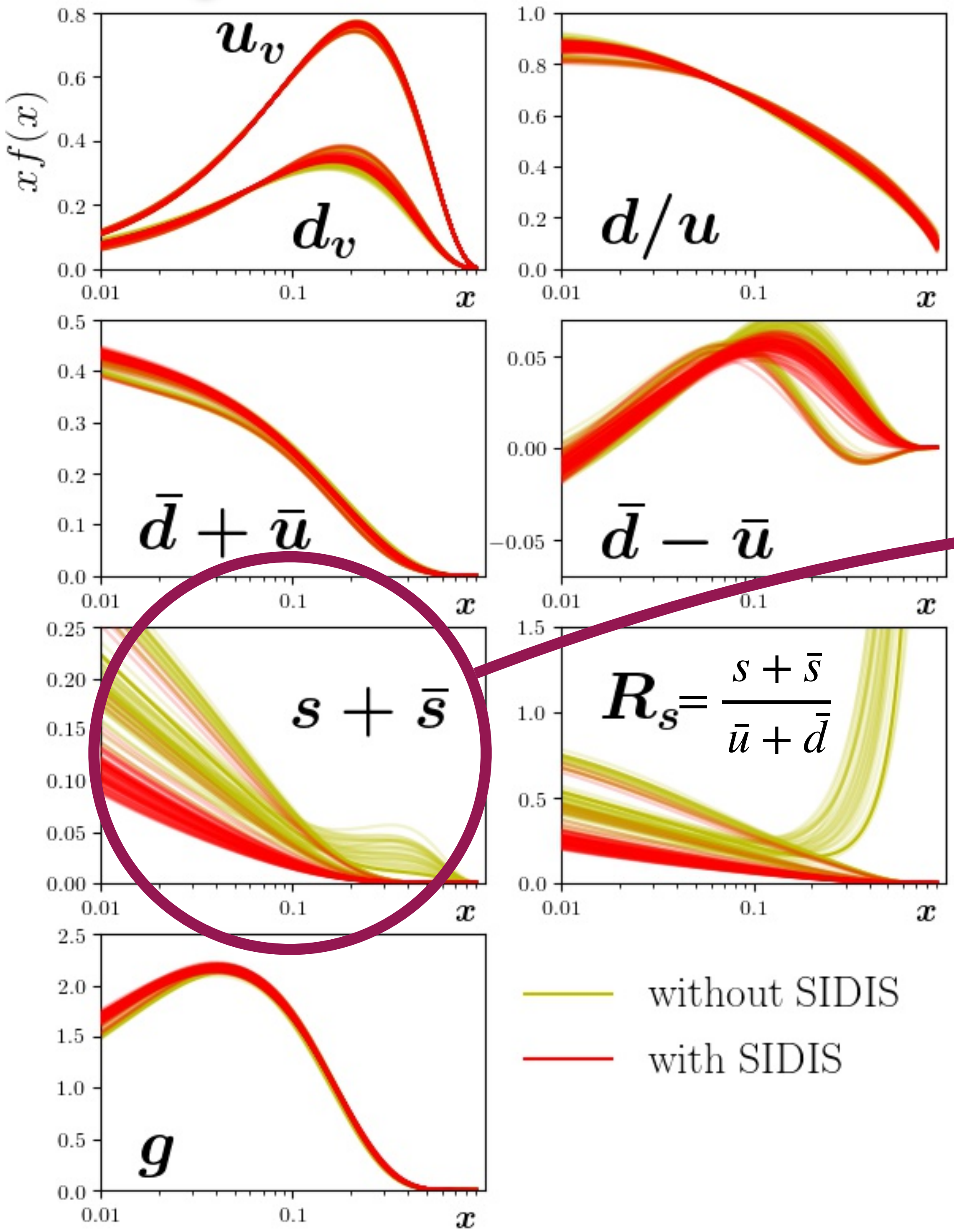
Q = m$_c$

Impact of SIDIS data on PDFs

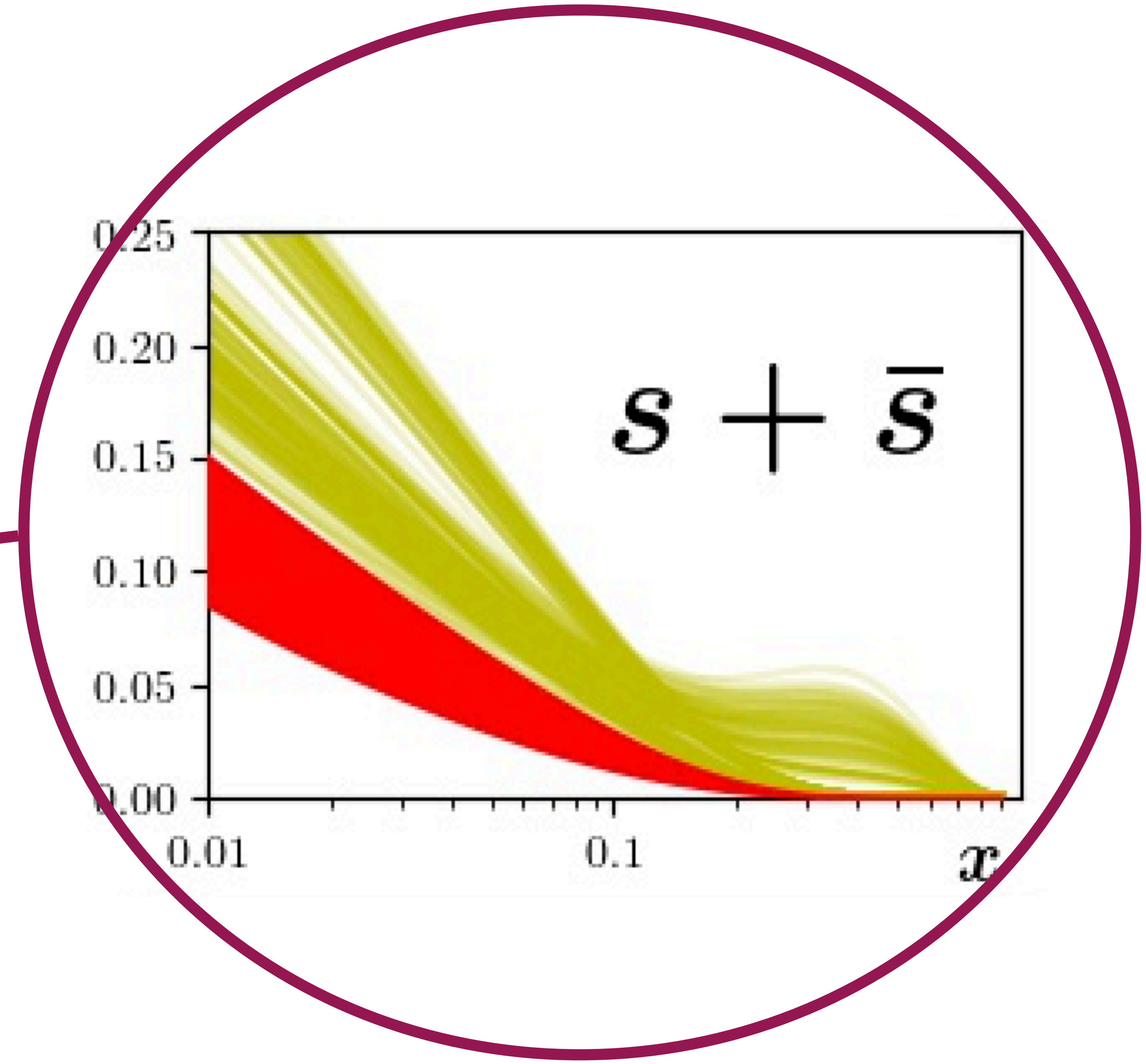


Q = m$_c$

Impact of SIDIS data on PDFs

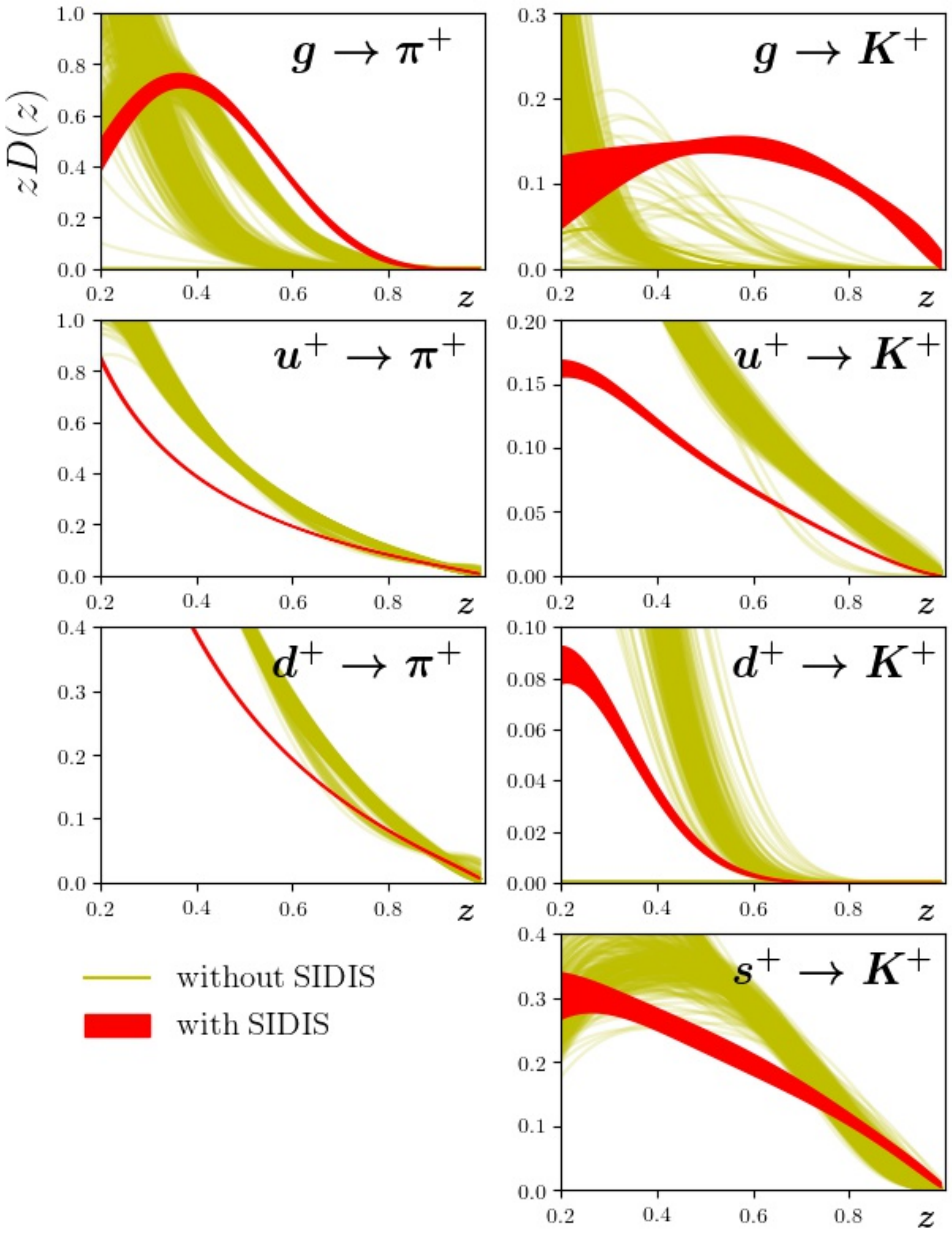


Q = m$_c$



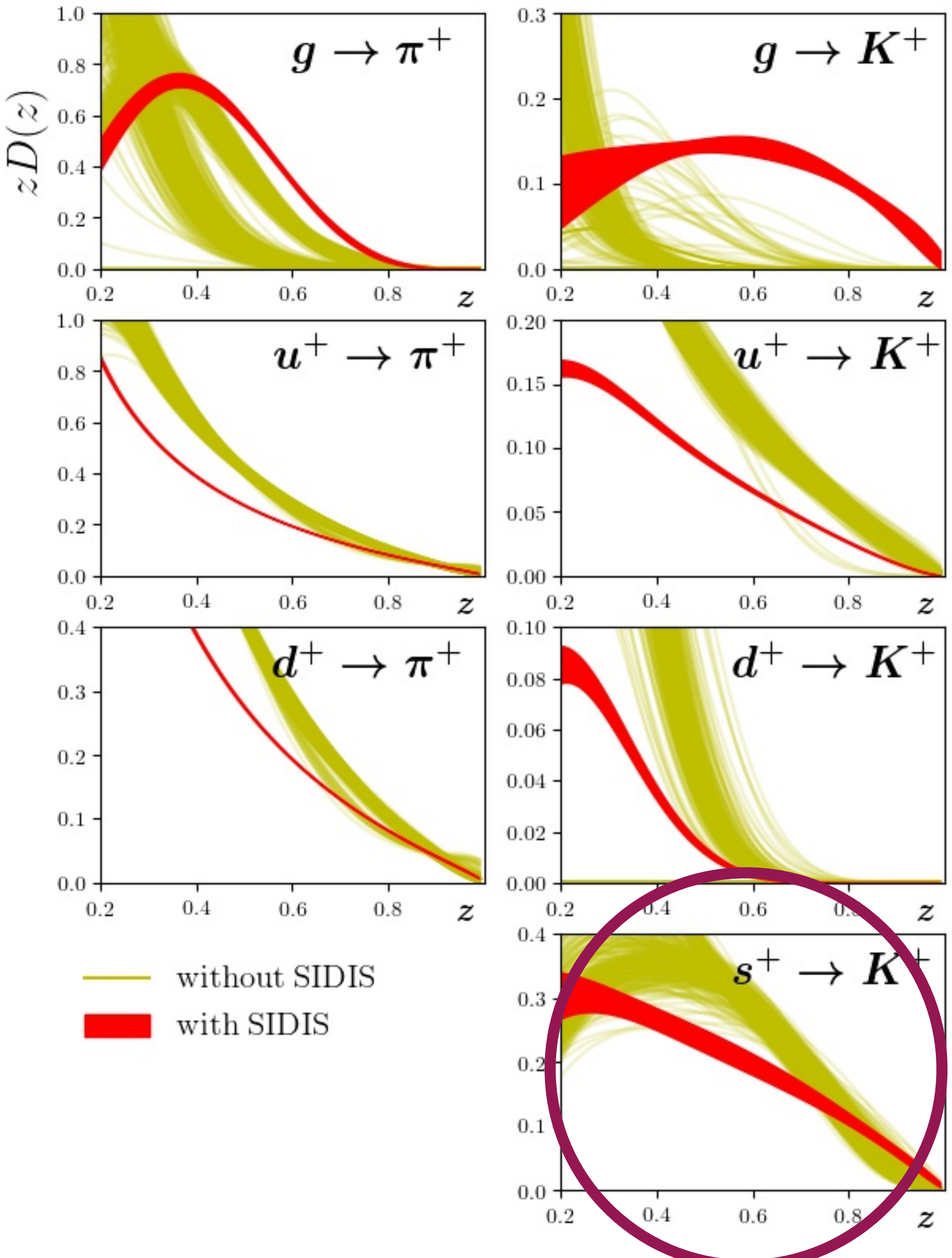
Strong strange suppression

Impact of SIDIS data on FFs



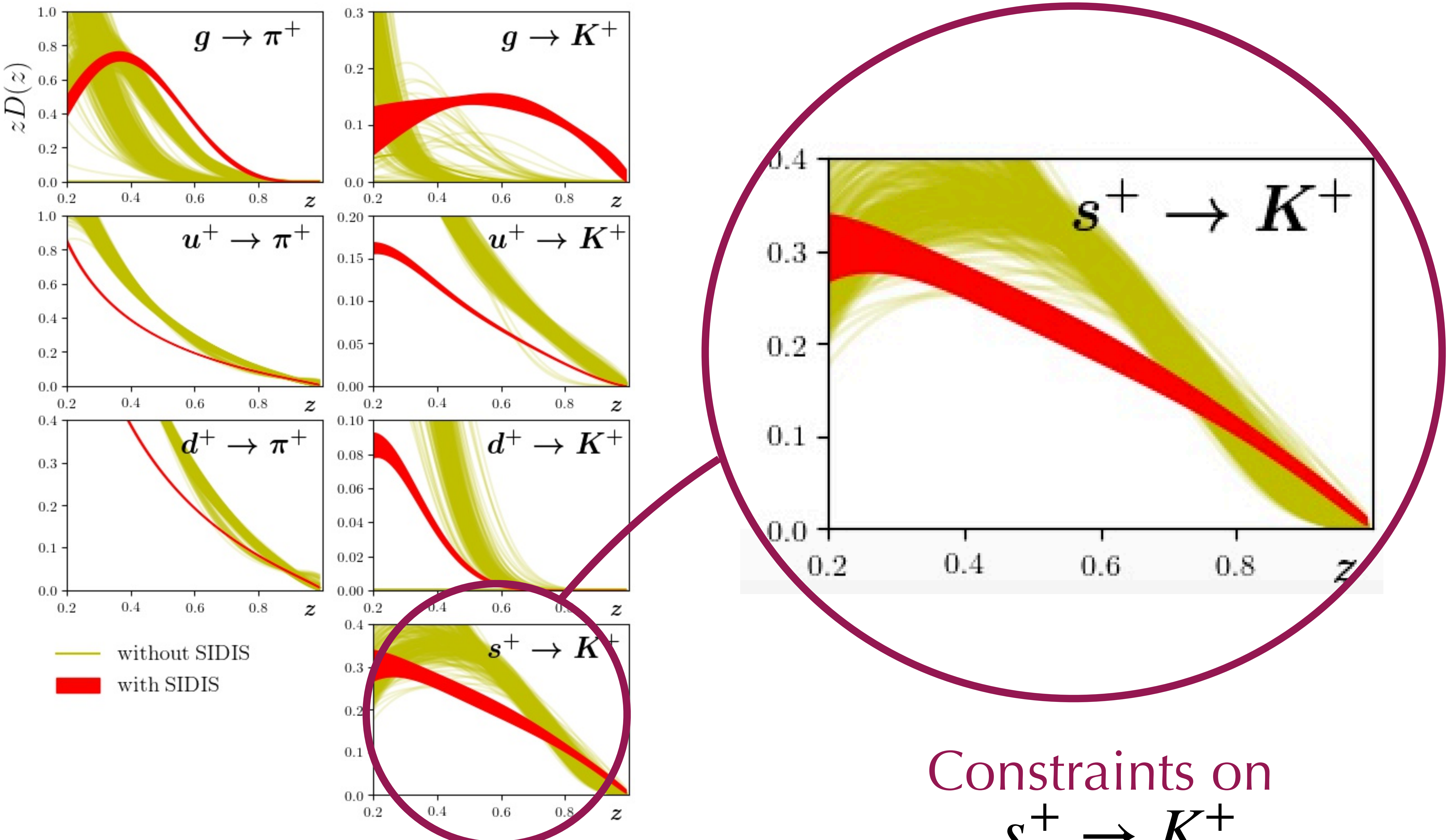
$Q = m_c$

Impact of SIDIS data on FFs



Q = m_c

Impact of SIDIS data on FFs



$\mathbf{Q = m_c}$

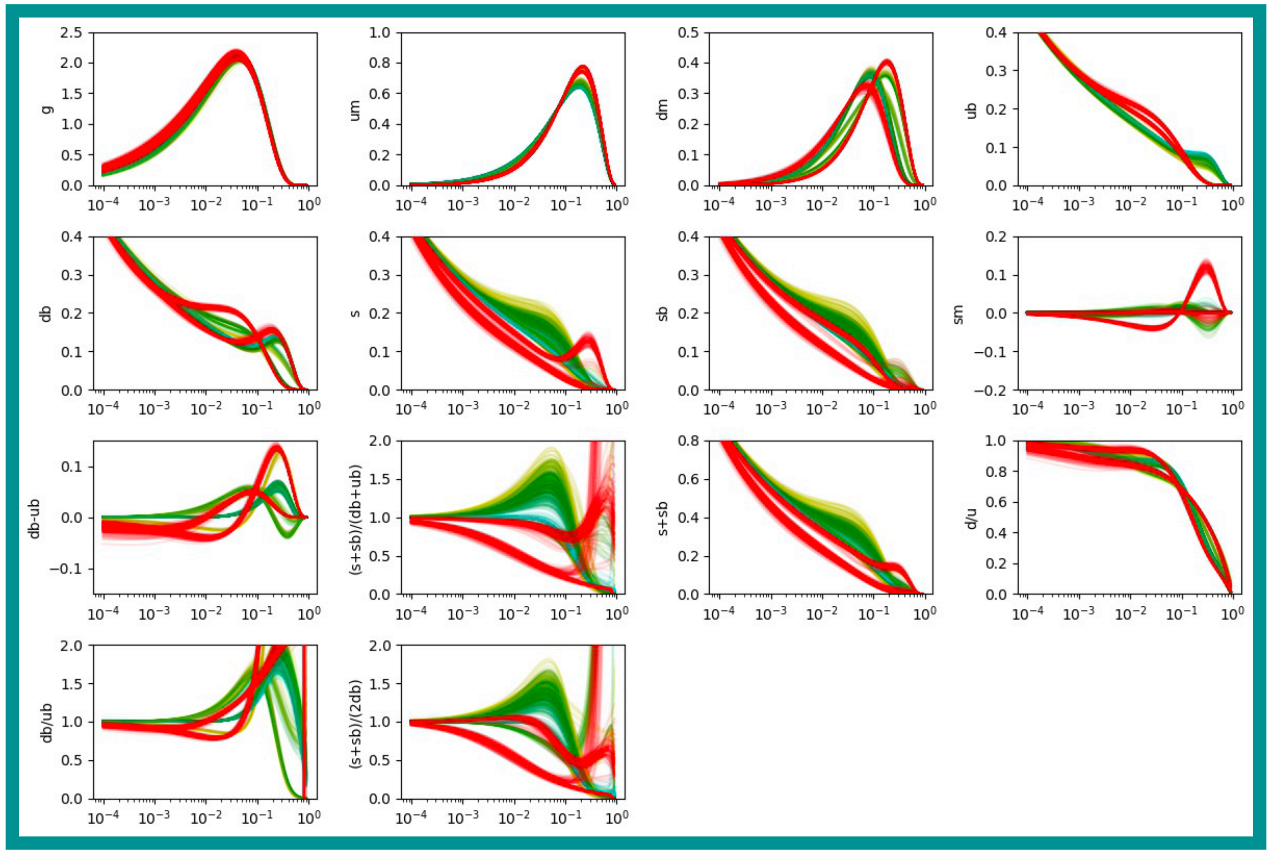
Constraints on
$s^+ \to K^+$

Constraints on Rs

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs

$x\,f(x)$



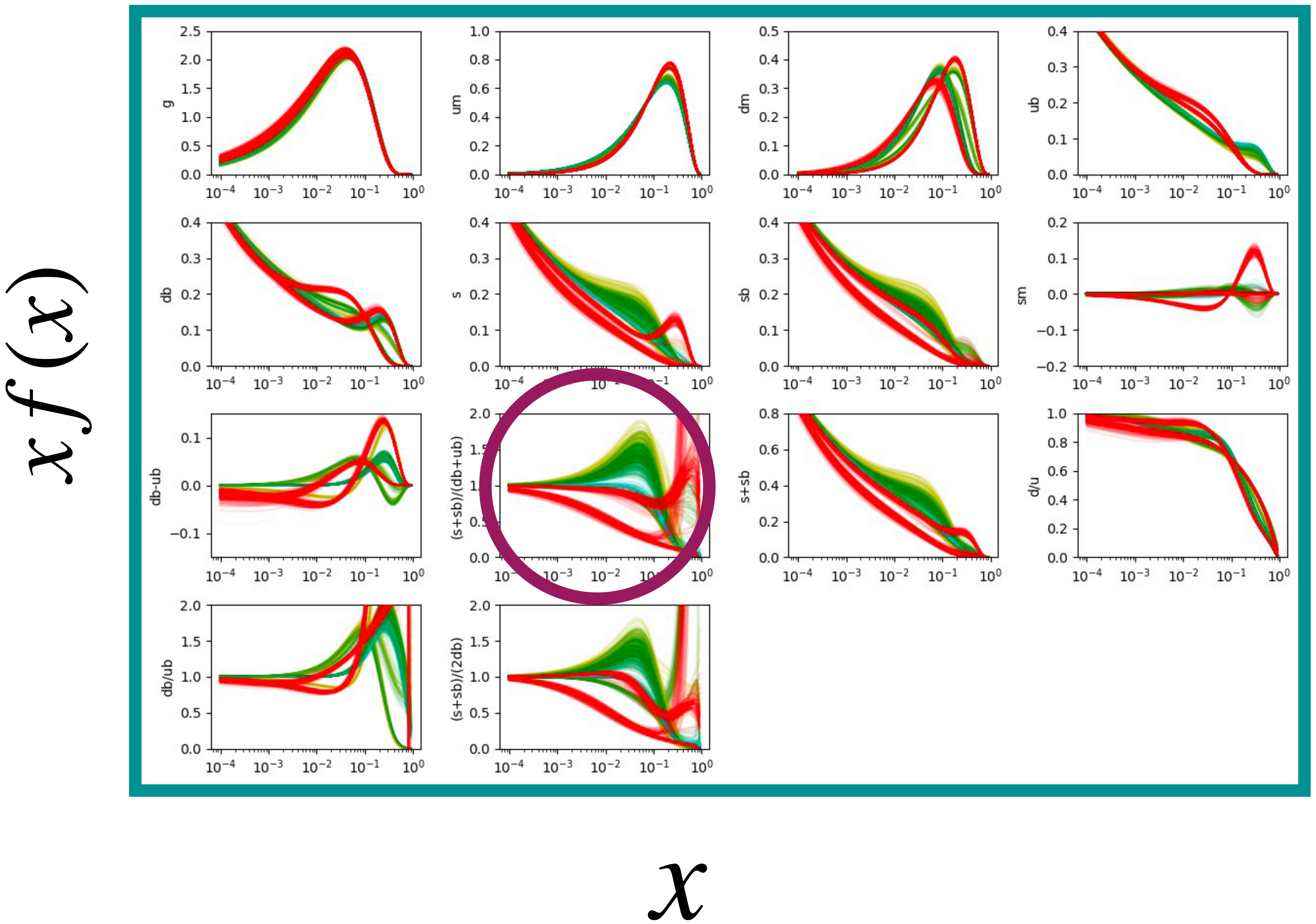
x

+ DIS data

Constraints on Rs

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs

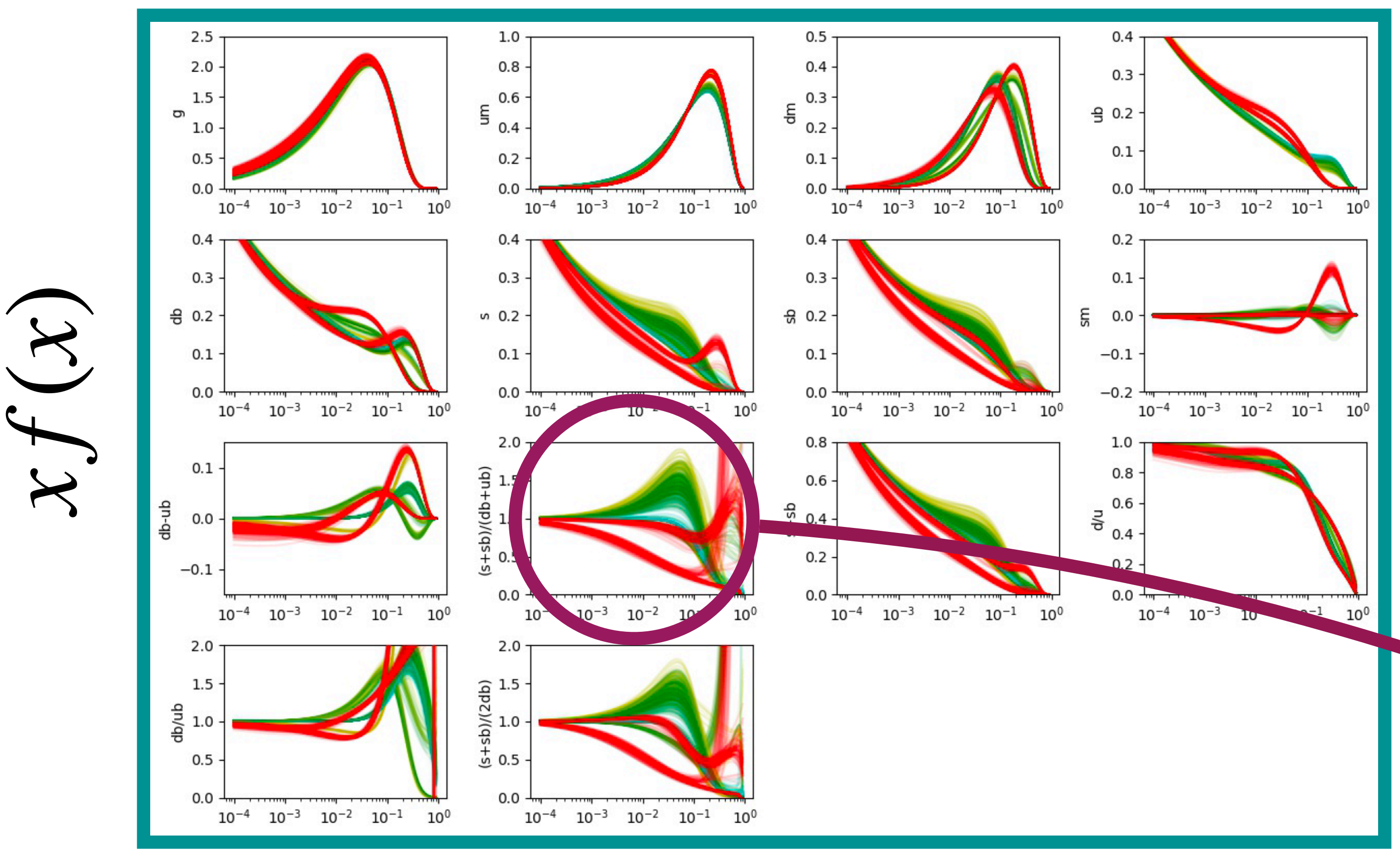


+ DIS data

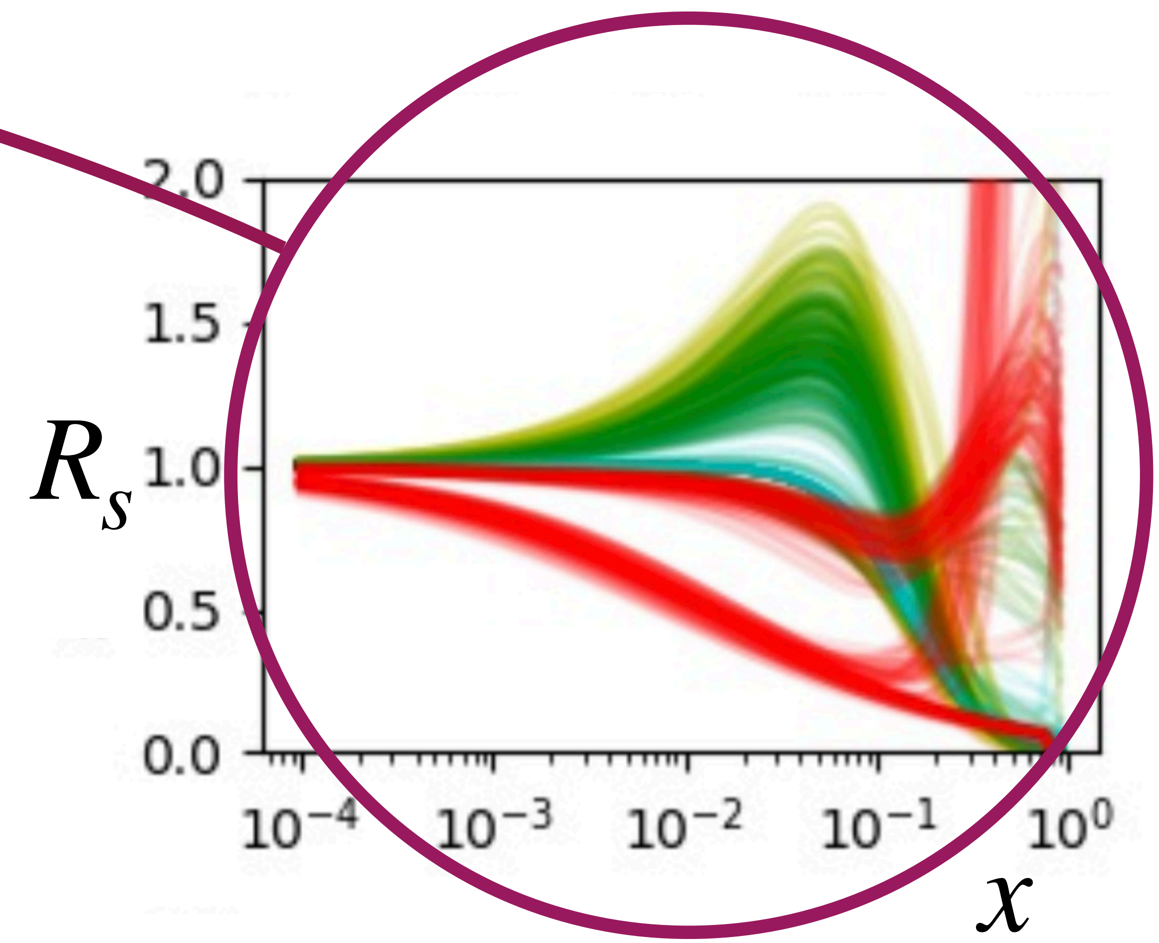
Constraints on Rs

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs



+ DIS data

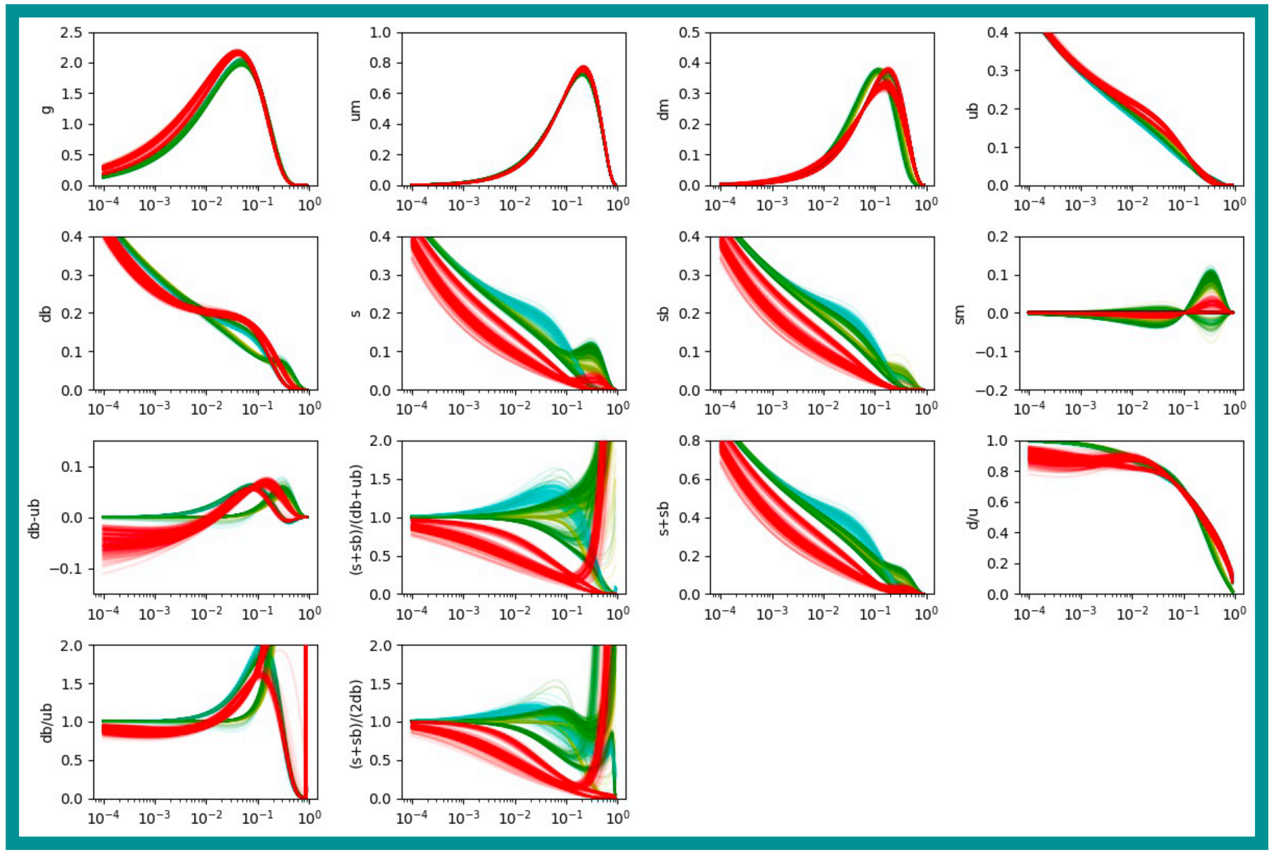


Constraints on Rs

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs

$x\,f(x)$



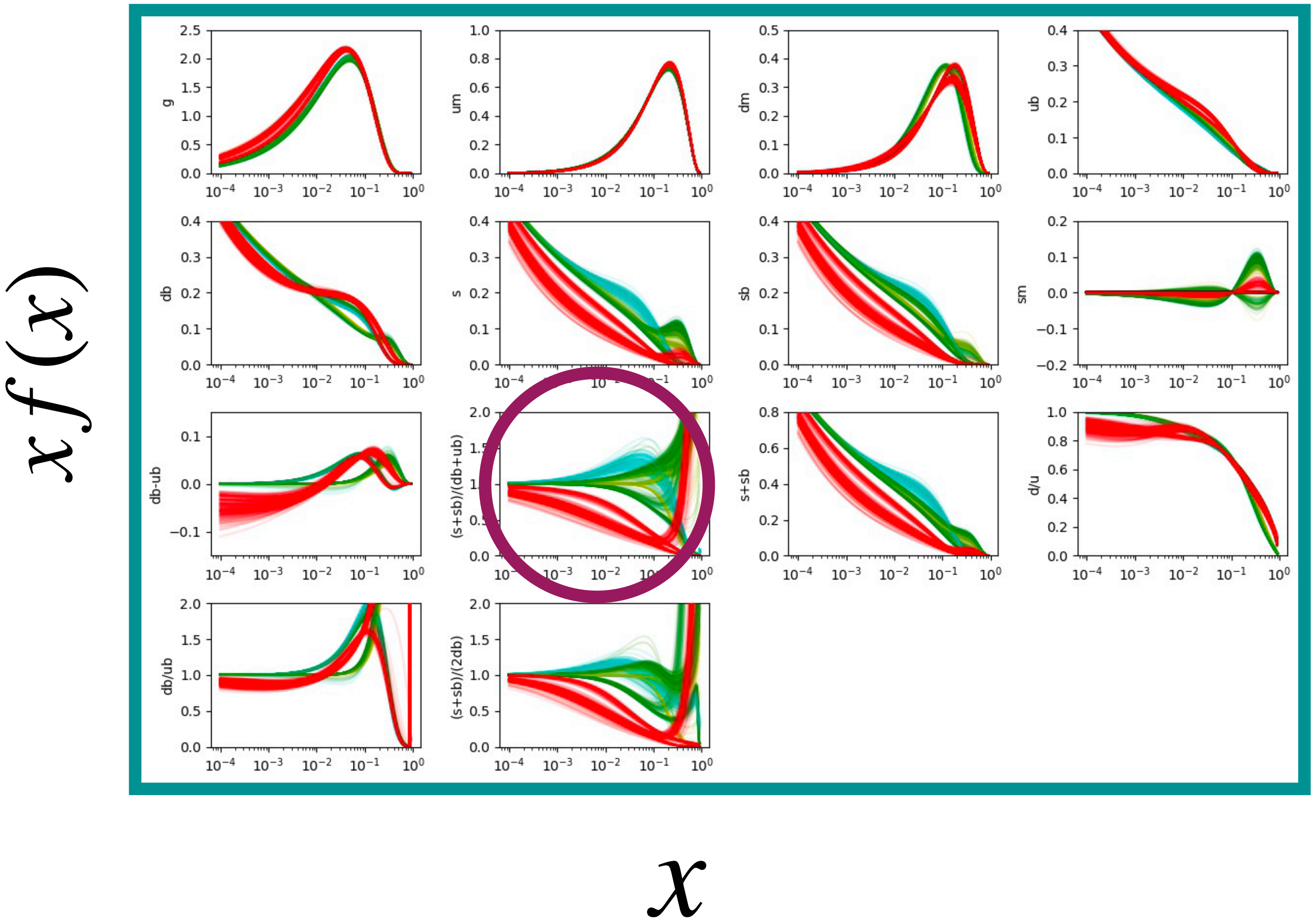
x

+ DIS data
+ DY data

Constraints on Rs

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs

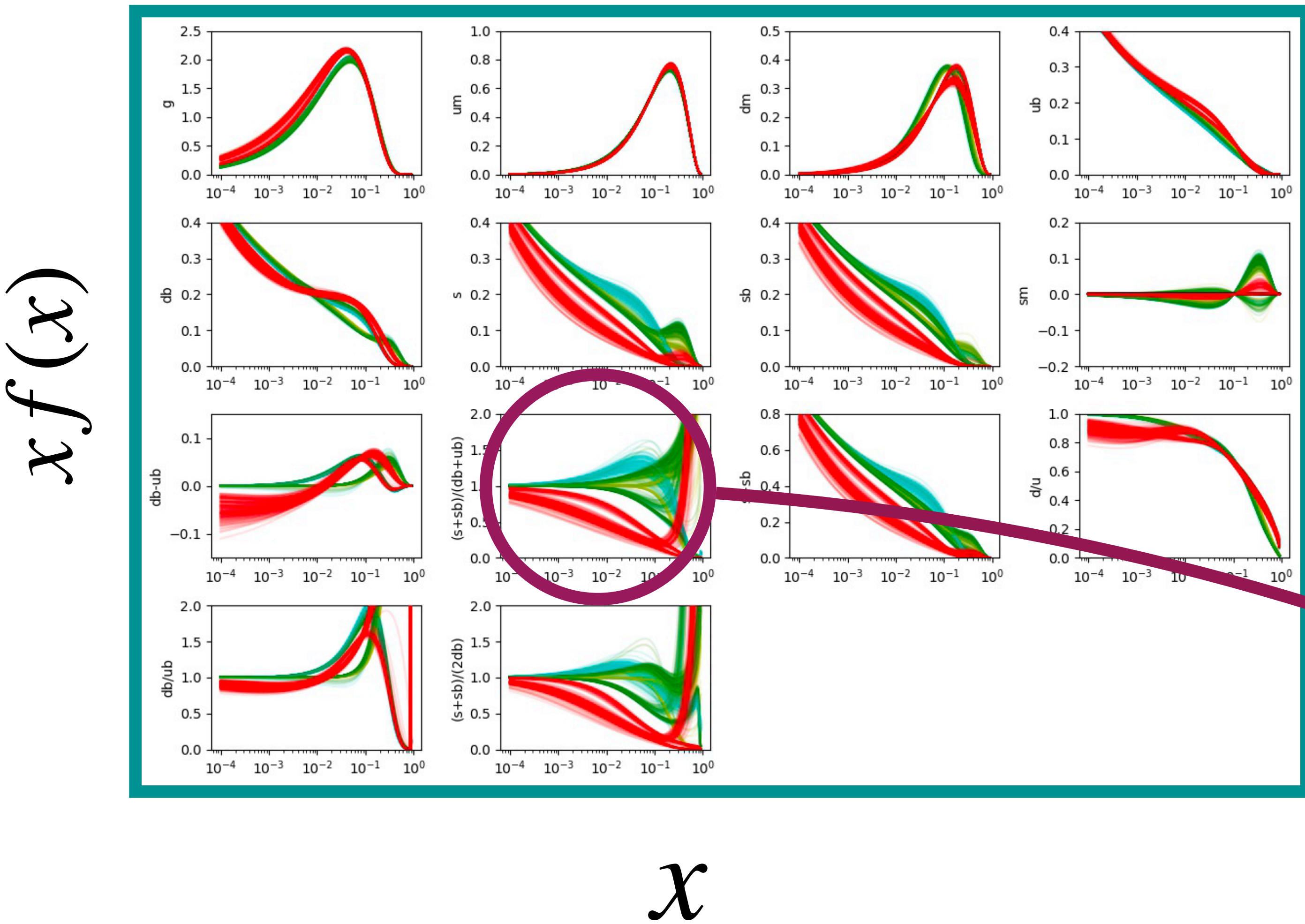


+ DIS data
+ DY data

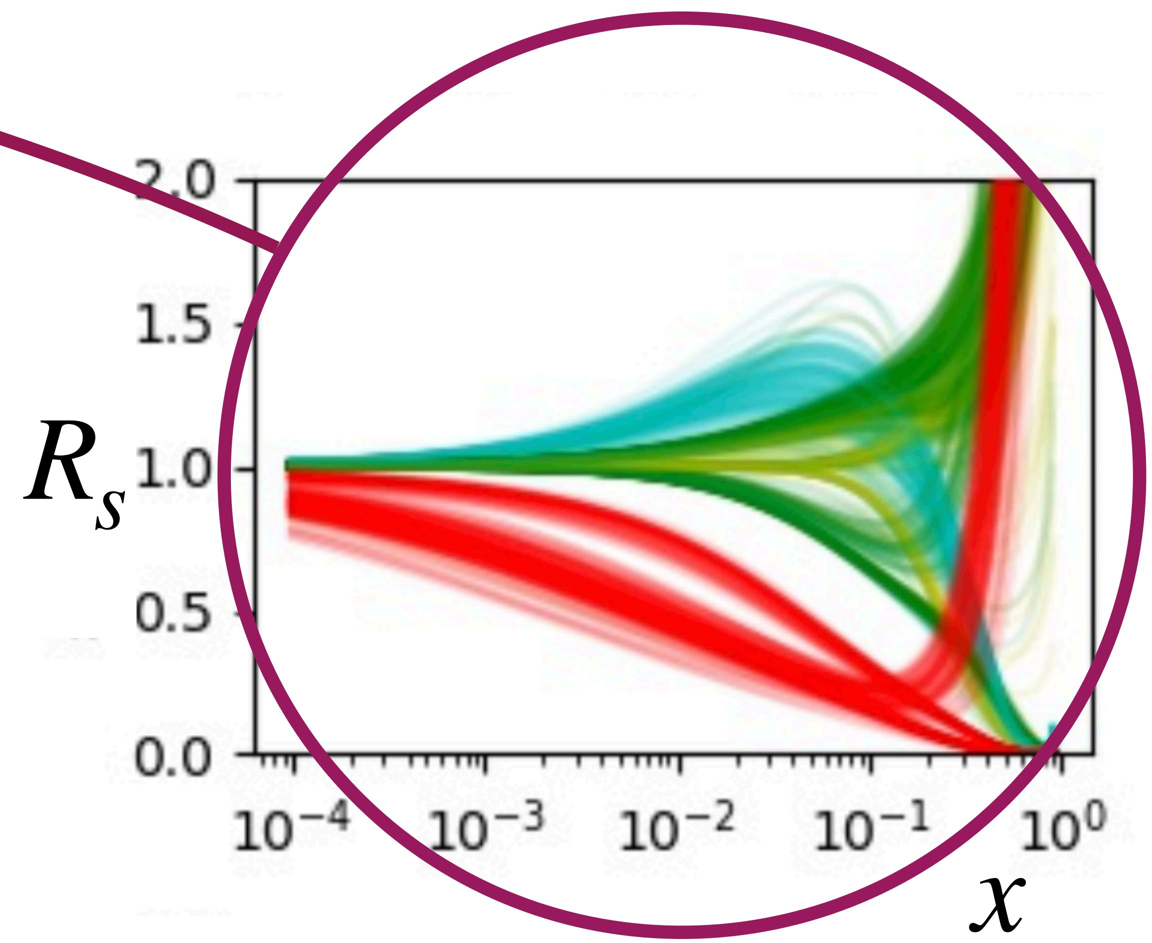
Constraints on Rs

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs



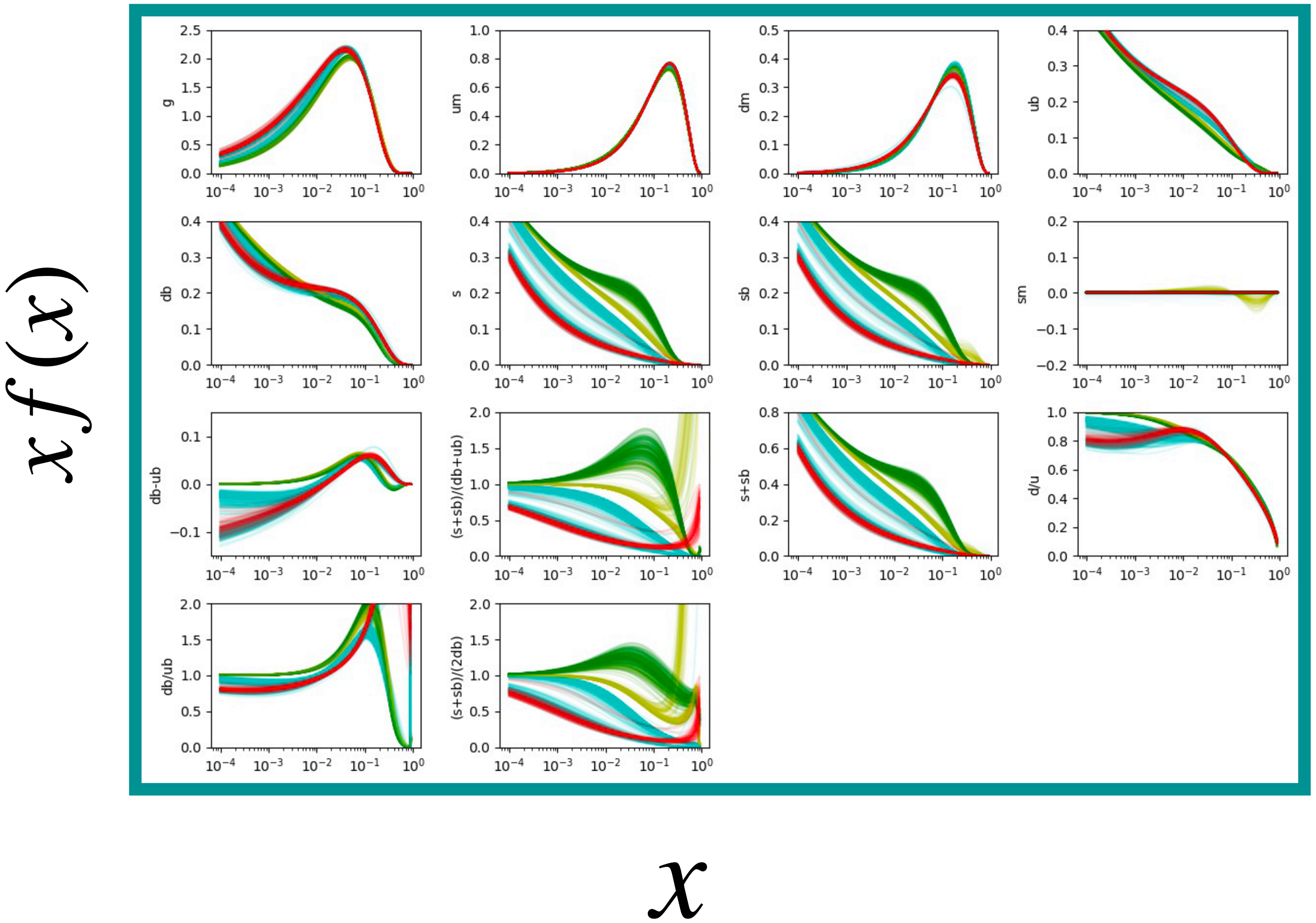
+ DIS data
+ DY data



Constraints on R_s

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs



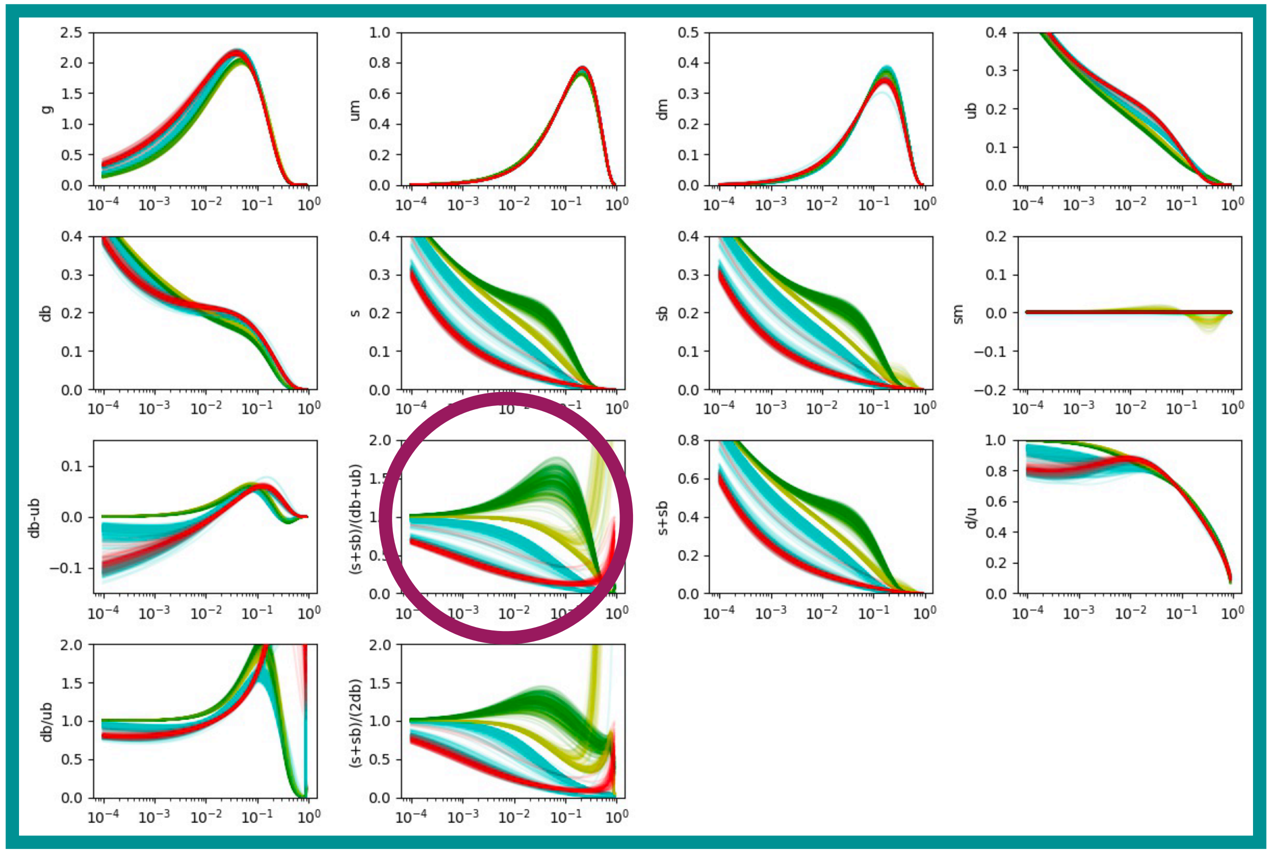
+ DIS data
+ DY data
+ SIA + SIDIS data

Constraints on R_s

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs

$x\, f(x)$



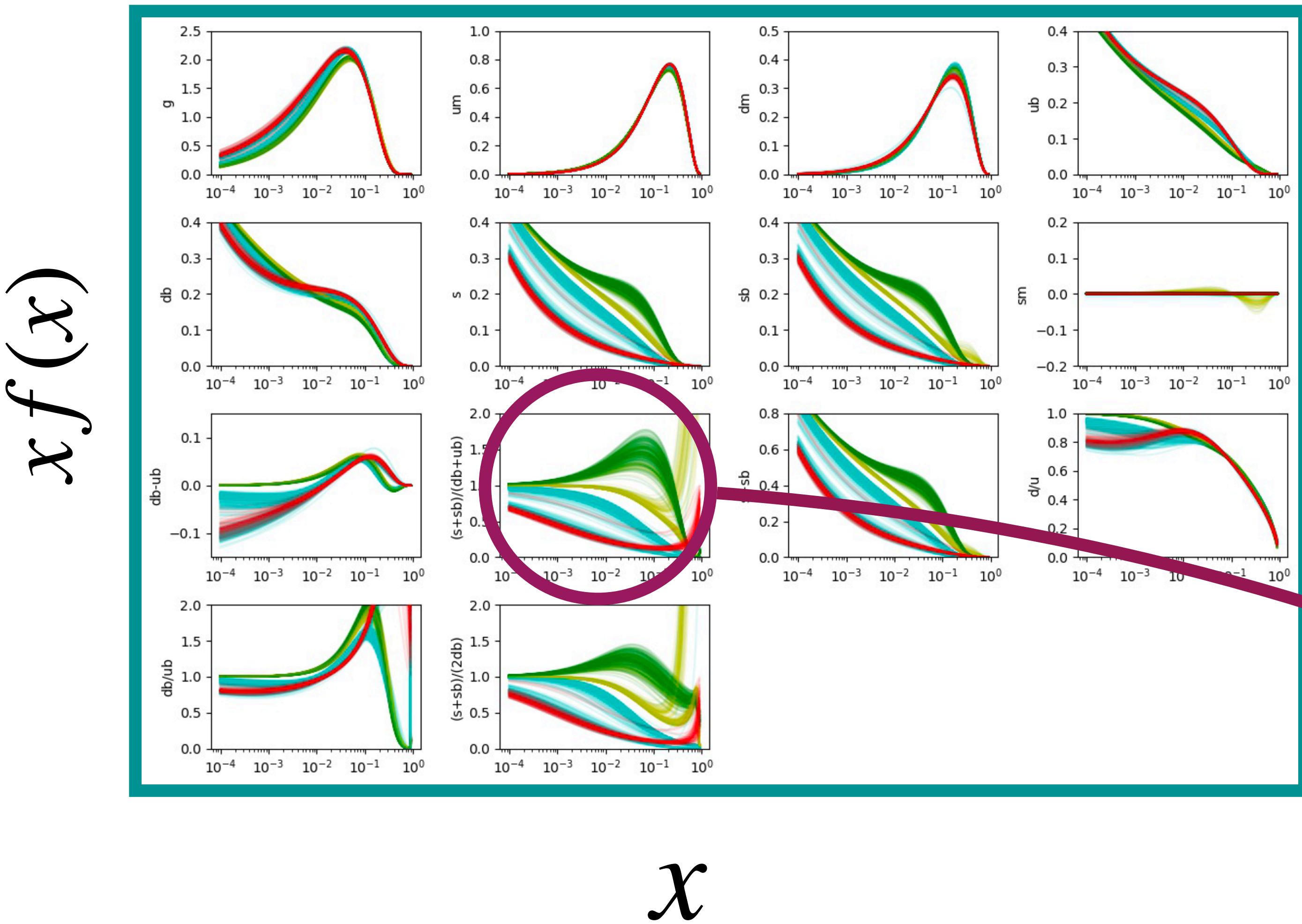
x

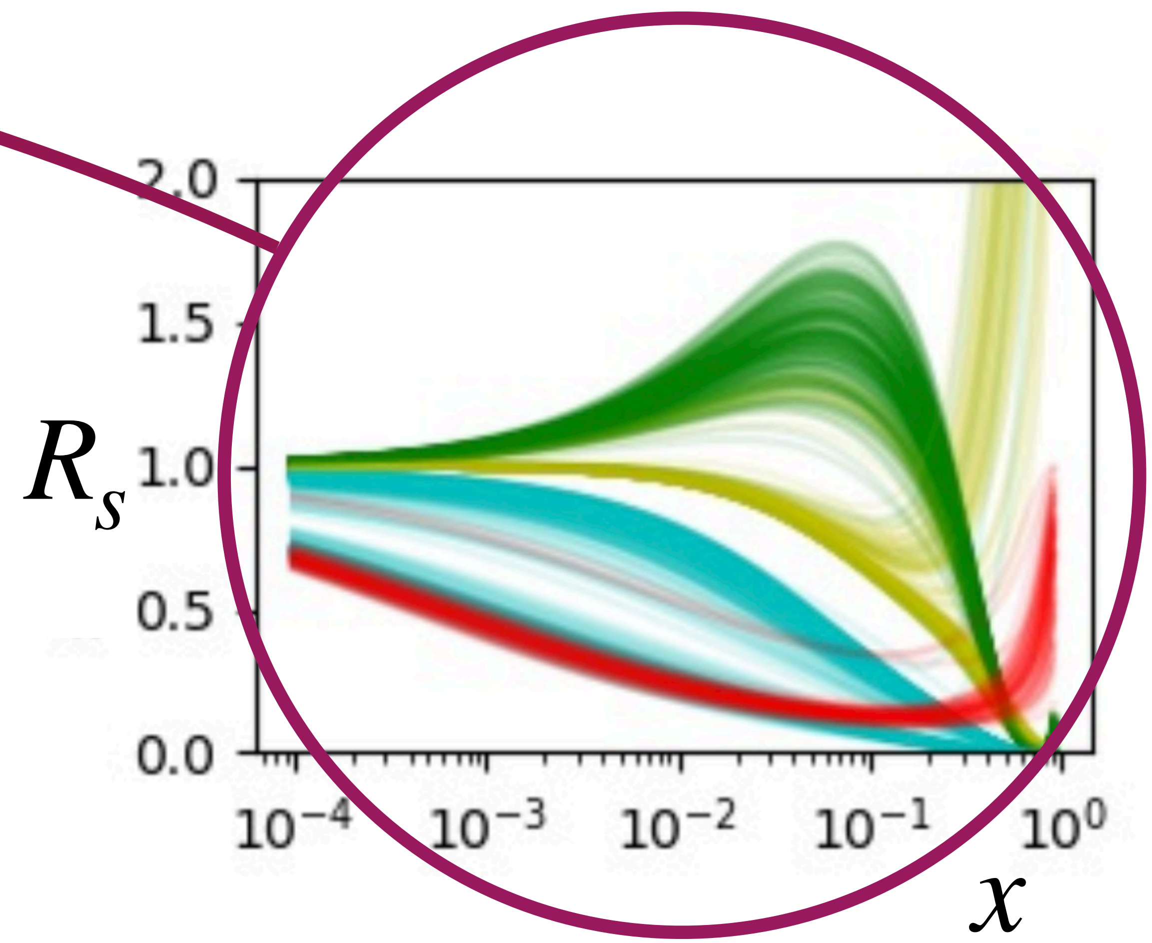
+ DIS data
+ DY data
+ SIA + SIDIS data

Constraints on R_s

$$R_s = \frac{s + \bar{s}}{\bar{u} + \bar{d}}$$

PDFs





+ DIS data
+ DY data
+ SIA + SIDIS data

Chi2

Reaction	N_{dat}	χ^2	χ^2/N_{dat}
SIDIS	992	1243.12	1.25
SIA	444	562.80	1.27
DIS	2680	3437.96	1.28
DY	250	416.29	1.67

Reaction	N_{dat}	χ^2	χ^2/N_{dat}
SIDIS ($\pi^{\pm}$)	498	585.48	1.18
SIDIS($K^{\pm}$)	494	657.64	1.33
SIA($\pi^{\pm}$)	231	247.27	1.07
SIA ($K^{\pm}$)	213	315.53	1.48

Experiment	target	hadron	N_{dat}	χ^2/N_{dat}
COMPASS	d	π^+	249	1.26
COMPASS	d	π^-	249	1.09
COMPASS	d	K^+	247	1.24
COMPASS	d	K^-	247	1.43