Global fits for deep inelastic scattering and related processes

Nobuo Sato
ODU/JLab

Jefferson Lab Angular Momentum Collaboration

2019 Fall Meeting of the APS
Division of Nuclear Physics

Jefferson Lab
EXPLORING THE NATURE OF MATTER

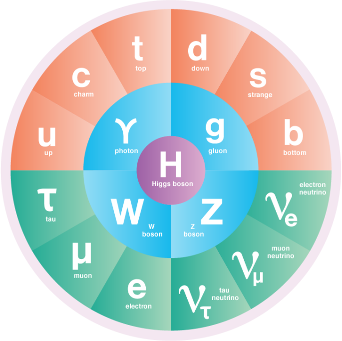
Outline

- Motivations
- Global QCD analysis in a nutshell
- Regression strategies
- JAM 19

Motivations

Motivations

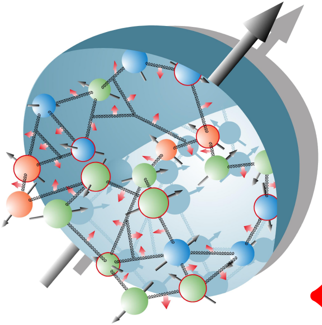
hadrons as emergent phenomena of QCD

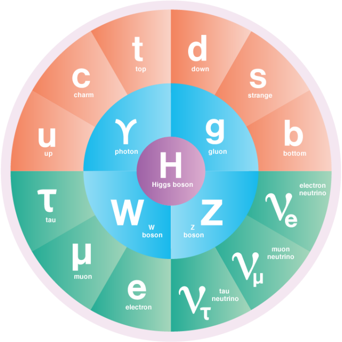


quarks and gluons

Motivations

hadrons as emergent phenomena of QCD

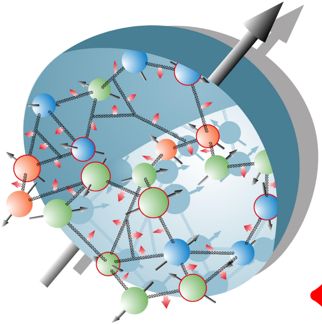


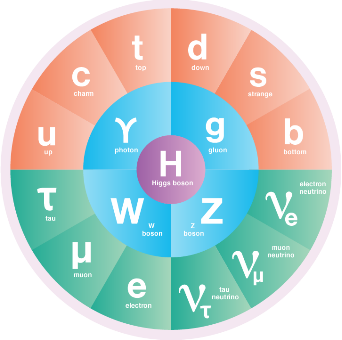


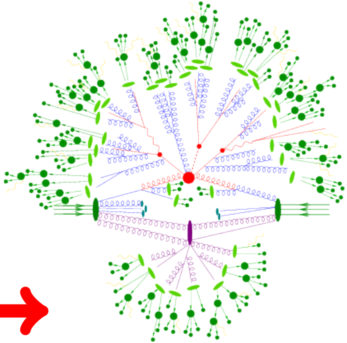
nucleon structure quarks and gluons

Motivations

hadrons as emergent phenomena of QCD

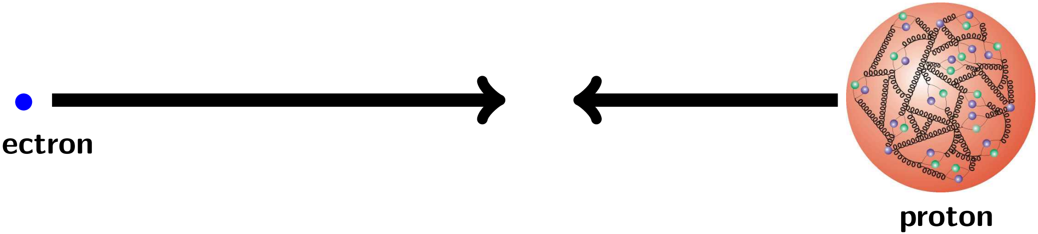




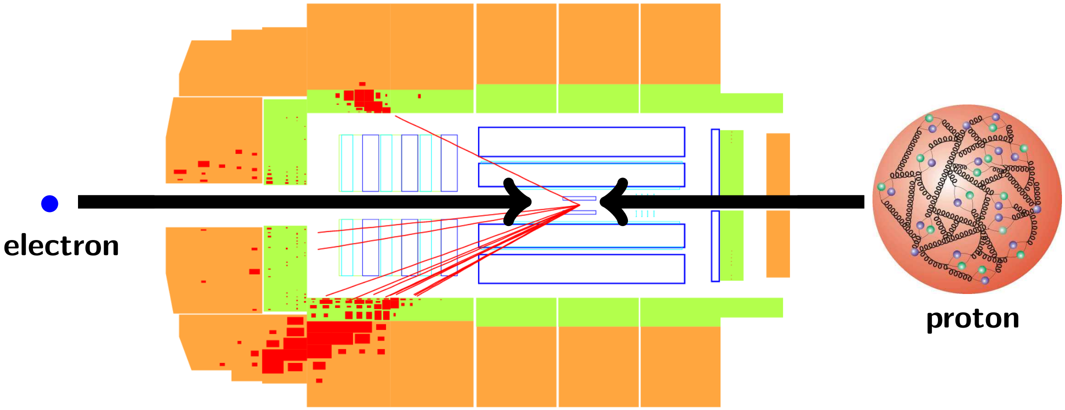


nucleon structure ← **quarks and gluons** → **hadronization**

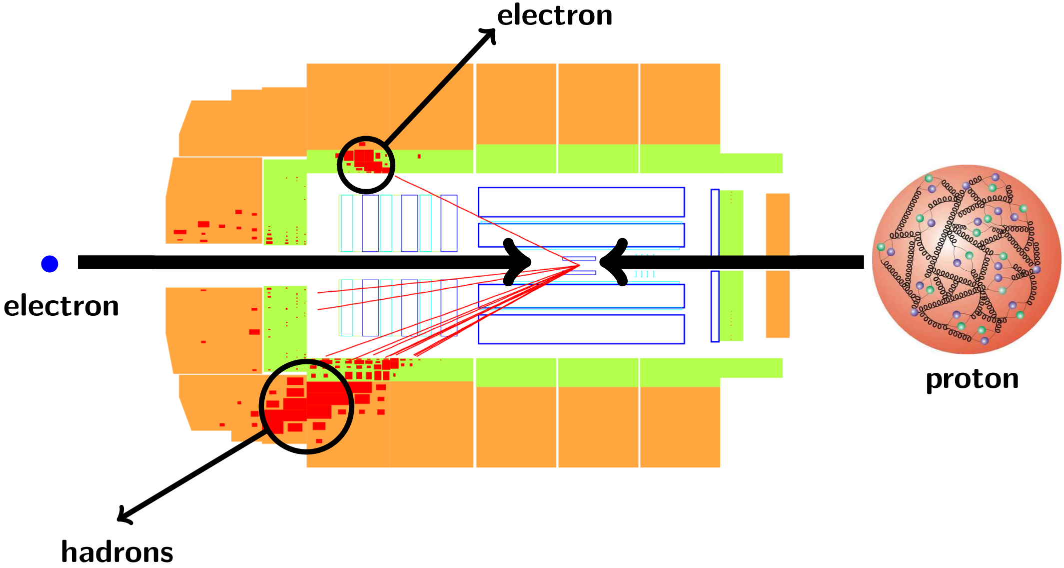
Motivations



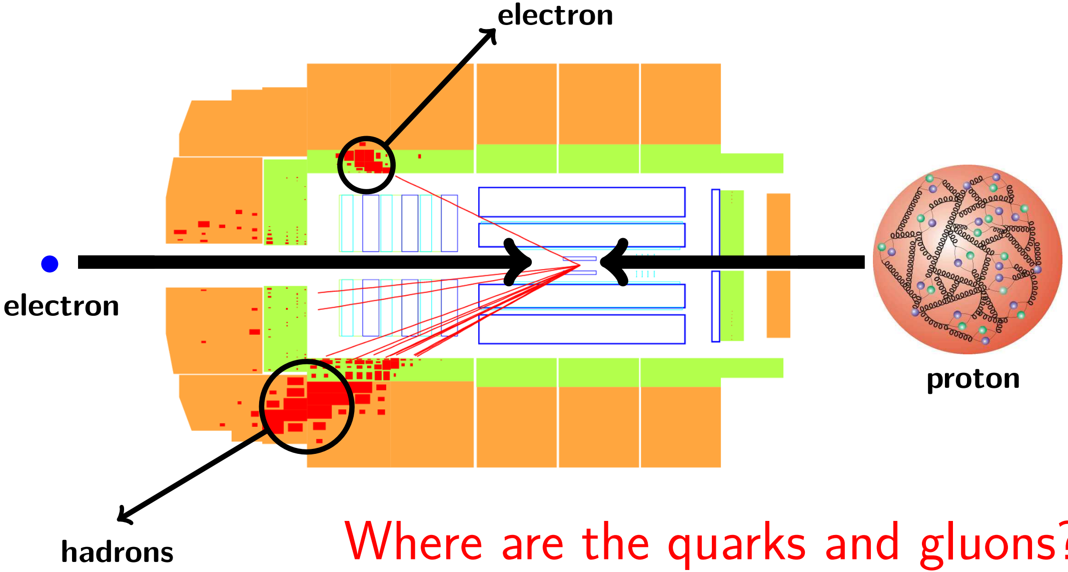
Motivations



Motivations



Motivations



Where are the quarks and gluons?

Motivations

- Quark and gluon d.o.f. **cannot be measured** directly

Motivations

- Quark and gluon d.o.f. **cannot be measured** directly

- Experimental measurements **can be interpreted** in terms of quark and gluon d.o.f.

Motivations

The interpretation relies on:

Motivations

The interpretation relies on:

- QCD factorization theorems (theory)

Motivations

The interpretation relies on:

- QCD factorization theorems (theory)
- Experimental cross section measurements

Motivations

The interpretation relies on:

- QCD factorization theorems (theory)
- Experimental cross section measurements
- Global QCD analysis (Bayesian regression)

Motivations

The strategy:

Motivations

The strategy:

1. Define **nucleon structure/hadronization** objects in QFT

Motivations

The strategy:

1. Define **nucleon structure/hadronization** objects in QFT
2. Identify cross sections that **factorize** in terms of such QFT objects

Motivations

The strategy:

1. Define **nucleon structure/hadronization** objects in QFT
2. Identify cross sections that **factorize** in terms of such QFT objects
3. Perform a **global QCD analysis**

Motivations

What do we mean by "**structure of nucleon**"? e.g.

Motivations

What do we mean by "**structure of nucleon**"? e.g.

$$f_{j/h}(\xi) = \int \frac{dw^-}{2\pi} e^{-i\xi P^+ w^-} \boxed{\left\langle P \left| \bar{\psi}_j(0, w^-, \mathbf{0}_\mathrm{T}) \frac{\gamma^+}{2} \psi_j(0) \right| P \right\rangle}$$

Motivations

What do we mean by "**structure of nucleon**"? e.g.

$$f_{j/h}(\xi) = \int \frac{dw^-}{2\pi} e^{-i\xi P^+ w^-} \boxed{\left\langle P \left| \bar{\psi}_j(0, w^-, \mathbf{0}_\mathrm{T}) \frac{\gamma^+}{2} \psi_j(0) \right| P \right\rangle}$$

- Not currently computable from first principles

Motivations

What do we mean by "**structure of nucleon**"? e.g.

$$f_{j/h}(\xi) = \int \frac{dw^-}{2\pi} e^{-i\xi P^+ w^-} \boxed{\left\langle P \left| \bar{\psi}_j(0, w^-, \mathbf{0}_\mathrm{T}) \frac{\gamma^+}{2} \psi_j(0) \right| P \right\rangle}$$

- Not currently computable from first principles
- Needs to be inferred from data

Motivations

What do we mean by "**hadronization**"? e.g.

Motivations

What do we mean by "**hadronization**"? e.g.

$$d_{h/j}(\zeta) = \frac{\mathrm{Tr}_{\text{color,Dirac}}}{4N_{c,j}} \sum_X \zeta \int \frac{dw^+}{2\pi} e^{i(p_h^-/\zeta)w^+}$$

$$\times \boxed{\gamma^- \langle 0 \left| \bar{\psi}_j(0, w^+, \mathbf{0}_\mathrm{T}) \right. \left| p_h, X \right\rangle \left\langle p_h, X \right| \psi_j(0) \left| 0 \right\rangle}$$

Motivations

What do we mean by "**hadronization**"? e.g.

$$d_{h/j}(\zeta) = \frac{\mathrm{Tr}_{\text{color,Dirac}}}{4N_{c,j}} \sum_X \zeta \int \frac{dw^+}{2\pi} e^{i(p_h^-/\zeta)w^+}$$

$$\times \boxed{\gamma^- \langle 0 \left| \bar{\psi}_j(0, w^+, \mathbf{0}_\mathrm{T}) \right| p_h, X \rangle \langle p_h, X | \psi_j(0) | 0 \rangle}$$

- Not currently computable from first principles

Motivations

What do we mean by "**hadronization**"? e.g.

$$d_{h/j}(\zeta) = \frac{\mathrm{Tr}_{\text{color,Dirac}}}{4N_{c,j}} \sum_X \zeta \int \frac{dw^+}{2\pi} e^{i(p_h^-/\zeta)w^+}$$

$$\times \boxed{\gamma^- \langle 0 \left| \bar{\psi}_j(0, w^+, \mathbf{0}_\mathrm{T}) \right| p_h, X \rangle \langle p_h, X | \psi_j(0) | 0 \rangle}$$

- Not currently computable from first principles
- Needs to be inferred from data

Motivations

Acronyms for 1D distributions

Motivations

Acronyms for 1D distributions

- $f_{j/h}(\xi)$: “Parton Distribution Functions”

  PDFs

Motivations

Acronyms for 1D distributions

- $f_{j/h}(\xi)$: “Parton Distribution Functions”
  PDFs

- $d_{h/j}(\zeta)$: “Fragmentation Functions”
  FFs

Motivations

What do we mean by “**factorization**”? e.g DIS

Motivations

What do we mean by "**factorization**"? e.g DIS

$$F_2(x, Q) = x \sum_j e_j^2 \int_x^1 \frac{d\xi}{\xi} \quad C_2(\xi, \mu) \quad\quad f_j\left(\frac{x}{\xi}, \mu\right)$$

Motivations

What do we mean by "**factorization**"? e.g DIS

$$F_2(x,Q) = x \sum_j e_j^2 \int_x^1 \frac{d\xi}{\xi} \boxed{C_2(\xi,\mu)} \quad f_j\left(\frac{x}{\xi},\mu\right)$$

- $\boxed{C_2}$ is calculable in perturbative QCD

Motivations

What do we mean by "**factorization**"? e.g DIS

$$F_2(x,Q) = x\sum_j e_j^2 \int_x^1 \frac{d\xi}{\xi} \boxed{C_2(\xi,\mu)} \boxed{f_j\left(\frac{x}{\xi},\mu\right)}$$

- $\boxed{C_2}$ is calculable in perturbative QCD
- $\boxed{f_j}$ cannot be solved in closed form

$\rightarrow$ **inverse problem**

Motivations

Another example: SIDIS

$$F_1^h(x,z,Q) = x\sum_j e_j^2 \int_x^1 \frac{d\xi}{\xi}\int_z^1 \frac{d\zeta}{\zeta} \;\; C_1(\xi,\zeta,\mu) \quad f_j\left(\frac{x}{\xi},\mu\right) \quad d_j\left(\frac{z}{\zeta},\mu\right)$$

Motivations

Another example: SIDIS

$$F_1^h(x,z,Q) = x\sum_j e_j^2 \int_x^1 \frac{d\xi}{\xi}\int_z^1 \frac{d\zeta}{\zeta}\, \boxed{C_1(\xi,\zeta,\mu)} \quad f_j\left(\frac{x}{\xi},\mu\right) \quad d_j\left(\frac{z}{\zeta},\mu\right)$$

Motivations

Another example: SIDIS

$$F_1^h(x,z,Q) = x\sum_j e_j^2 \int_x^1 \frac{d\xi}{\xi}\int_z^1 \frac{d\zeta}{\zeta}\ \boxed{C_1(\xi,\zeta,\mu)}\ \boxed{f_j\left(\frac{x}{\xi},\mu\right)}\ \boxed{d_j\left(\frac{z}{\zeta},\mu\right)}$$

- $\boxed{C_1}$ is calculable in perturbative QCD

Motivations

Another example: SIDIS

$$F_1^h(x,z,Q) = x\sum_j e_j^2 \int_x^1 \frac{d\xi}{\xi} \int_z^1 \frac{d\zeta}{\zeta} \boxed{C_1(\xi,\zeta,\mu)} \boxed{f_j\left(\frac{x}{\xi},\mu\right)} \boxed{d_j\left(\frac{z}{\zeta},\mu\right)}$$

- $\boxed{C_1}$ is calculable in perturbative QCD
- $\boxed{f_j}$ and $\boxed{d_j}$ cannot be solved in closed form

  $\rightarrow$ **inverse problem**

Motivations

Universality $\rightarrow$ the predictive power of QCD

Motivations

Universality $\rightarrow$ the predictive power of QCD

$$\sigma^{\mathrm{EXP}}_{l+P\rightarrow l+X} = \boxed{C_{l+k\rightarrow l+X}} \otimes \boxed{f}$$

Motivations

Universality $\to$ the predictive power of QCD

$$\sigma^{\mathrm{EXP}}_{l+P\to l+X} = \boxed{C_{l+k\to l+X}} \otimes \boxed{f}$$

$$\sigma^{\mathrm{EXP}}_{l+P\to l+H+X} = \boxed{C_{l+k\to l+k+X}} \otimes \boxed{f} \otimes \boxed{d}$$

Motivations

Universality $\rightarrow$ the predictive power of QCD

$$\sigma^{\mathrm{EXP}}_{l+P\rightarrow l+X} = \boxed{C_{l+k\rightarrow l+X}} \otimes \boxed{f}$$

$$\sigma^{\mathrm{EXP}}_{l+P\rightarrow l+H+X} = \boxed{C_{l+k\rightarrow l+k+X}} \otimes \boxed{f} \otimes \boxed{d}$$

$$\sigma^{\mathrm{EXP}}_{P+P\rightarrow l+\bar{l}+X} = \boxed{C_{k+k\rightarrow l+\bar{l}+X}} \otimes \boxed{f} \otimes \boxed{f}$$

Motivations

Universality $\rightarrow$ the predictive power of QCD

$$\sigma^{\mathrm{EXP}}_{l+P\to l+X} = \boxed{C_{l+k\to l+X}} \otimes \boxed{f}$$

$$\sigma^{\mathrm{EXP}}_{l+P\to l+H+X} = \boxed{C_{l+k\to l+k+X}} \otimes \boxed{f} \otimes \boxed{d}$$

$$\sigma^{\mathrm{EXP}}_{P+P\to l+\bar{l}+X} = \boxed{C_{k+k\to l+\bar{l}+X}} \otimes \boxed{f} \otimes \boxed{f}$$

$$\sigma^{\mathrm{EXP}}_{l+\bar{l}\to H+X} = \boxed{C_{l+\bar{l}\to k+X}} \otimes \boxed{d}$$

Global QCD analysis in a nutshell

Global QCD analysis in a nutshell

1. Parametrize $\boxed{f}$'s and $\boxed{d}$'s

Global QCD analysis in a nutshell

1. Parametrize $\boxed{f}$'s and $\boxed{d}$'s

$$\boxed{f_j(\xi)} = N_j \xi^{a_j}(1-\xi)^{b_j} P(\xi; \boldsymbol{w}_j)$$
$$\boxed{d_j(\zeta)} = \tilde{N}_j \zeta^{\tilde{a}_j}(1-\zeta)^{\tilde{b}_j} P(\zeta; \tilde{\boldsymbol{w}}_j)$$

Global QCD analysis in a nutshell

1. Parametrize $\boxed{f}$'s and $\boxed{d}$'s

$$\boxed{f_j(\xi)} = N_j \xi^{a_j}(1-\xi)^{b_j} P(\xi; \boldsymbol{w}_j)$$

$$\boxed{d_j(\zeta)} = \tilde{N}_j \zeta^{\tilde{a}_j}(1-\zeta)^{\tilde{b}_j} P(\zeta; \tilde{\boldsymbol{w}}_j)$$

$$\boldsymbol{p} = (\ldots, N_j, a_j, b_j, \boldsymbol{w}_j \ldots, \tilde{N}_j, \tilde{a}_j, \tilde{b}_j, \tilde{\boldsymbol{w}}_j, \ldots)$$

Global QCD analysis in a nutshell

2. Sample the Bayesian posterior distribution

Global QCD analysis in a nutshell

2. Sample the Bayesian posterior distribution

$$\rho\,(\boldsymbol{p}|\mathrm{data}) \propto \mathcal{L}(\boldsymbol{p},\mathrm{data})\pi(\boldsymbol{p})$$

Global QCD analysis in a nutshell

2. Sample the Bayesian posterior distribution

$$\rho\left(\boldsymbol{p}\middle|\mathrm{data}\right) \propto \mathcal{L}(\boldsymbol{p}, \mathrm{data})\pi(\boldsymbol{p})$$

$$\mathrm{E}[\mathcal{O}] = \frac{1}{N}\sum_k \mathcal{O}(\boldsymbol{p}_k) \qquad \mathrm{V}[\mathcal{O}] = \frac{1}{N}\sum_k \left[\mathcal{O}(\boldsymbol{p}_k) - \mathrm{E}[\mathcal{O}]\right]^2$$

Global QCD analysis in a nutshell

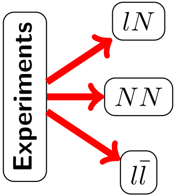
2. Sample the Bayesian posterior distribution

$$\rho\left(\boldsymbol{p}\middle|\mathrm{data}\right) \propto \mathcal{L}(\boldsymbol{p}, \mathrm{data})\pi(\boldsymbol{p})$$

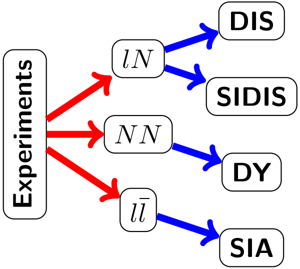
$$\mathrm{E}[\mathcal{O}] = \frac{1}{N}\sum_k \mathcal{O}(\boldsymbol{p}_k) \qquad \mathrm{V}[\mathcal{O}] = \frac{1}{N}\sum_k \left[\mathcal{O}(\boldsymbol{p}_k) - \mathrm{E}[\mathcal{O}]\right]^2$$

$$\mathcal{O} = f, d, \sigma, \ldots$$

Global QCD analysis in a nutshell

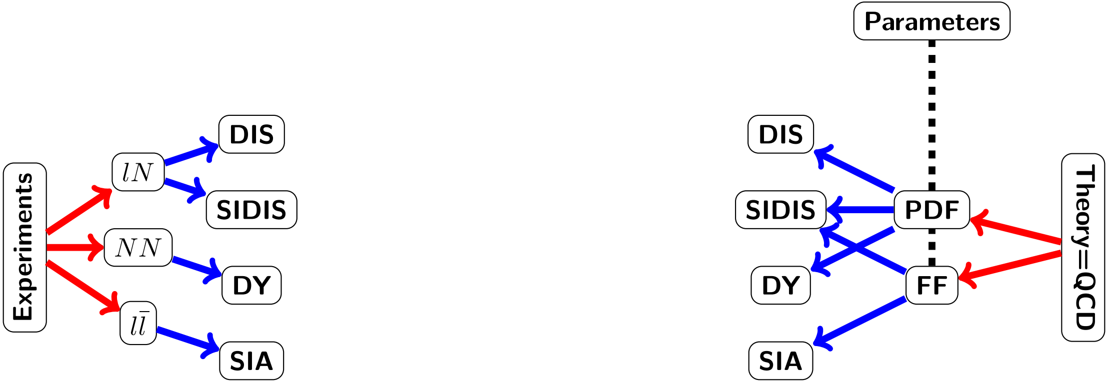


Global QCD analysis in a nutshell

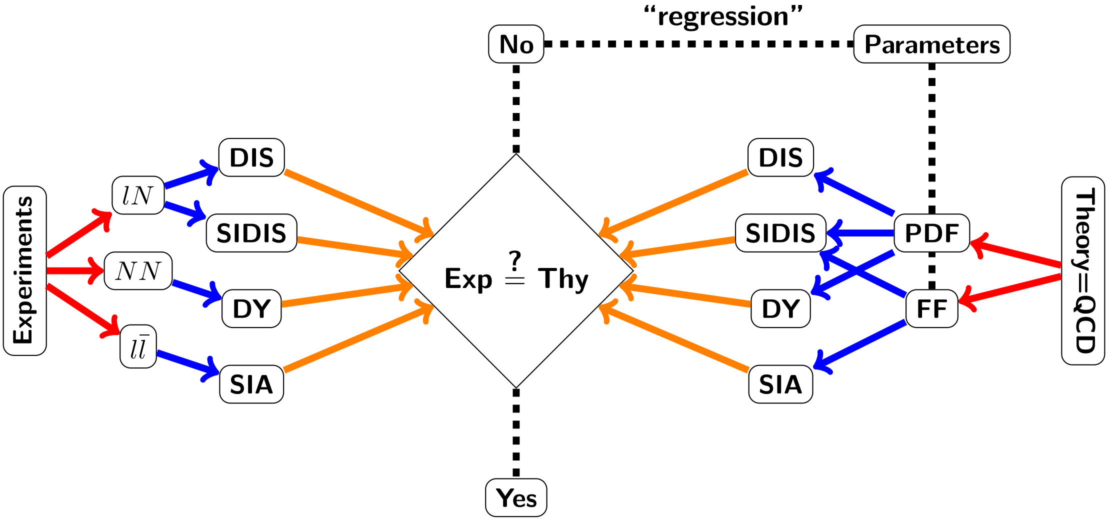


Global QCD analysis in a nutshell

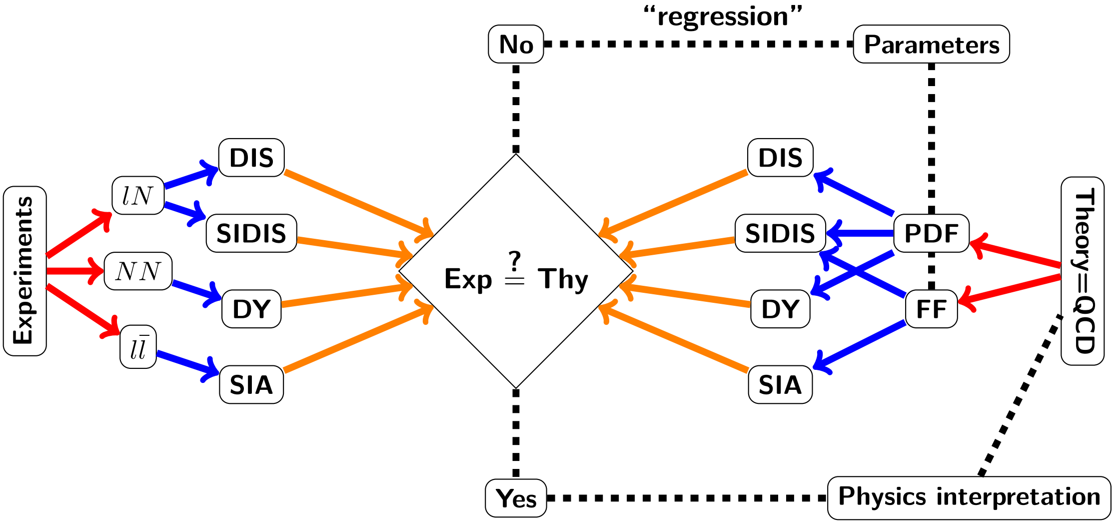
Global QCD analysis in a nutshell



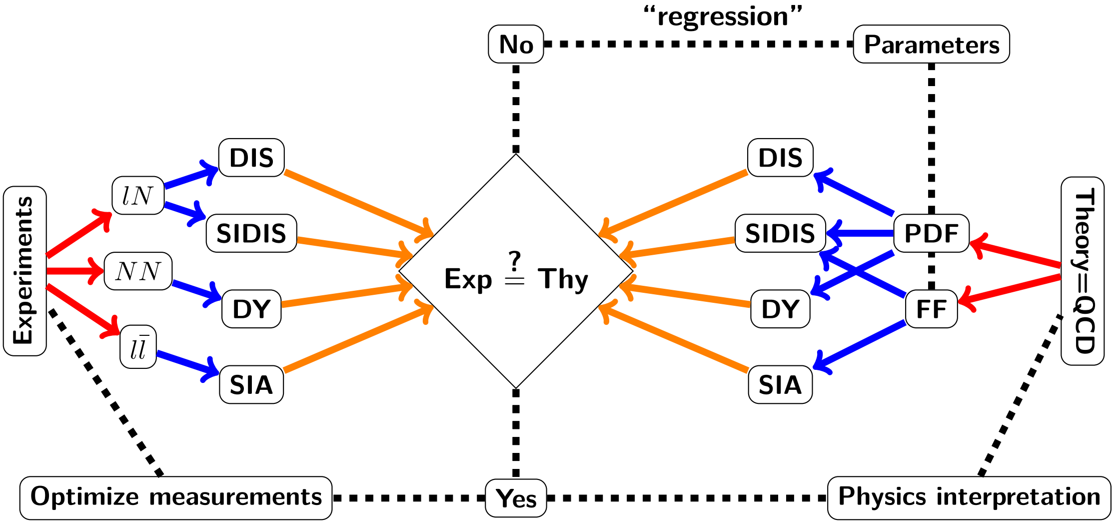
Global QCD analysis in a nutshell



Global QCD analysis in a nutshell



Global QCD analysis in a nutshell



Regression strategies

Regression strategies

- Maximum likelihood (CJ, CT, MMHT,...)

Regression strategies

- Maximum likelihood (CJ, CT, MMHT,...)

$$\mathrm{E}[\mathcal{O}] = \frac{1}{N}\sum_k \mathcal{O}(\boldsymbol{p}_k) \sim \mathcal{O}(\boldsymbol{p}_0)$$

Regression strategies

- Maximum likelihood (CJ, CT, MMHT,...)

$$\mathrm{E}[\mathcal{O}] = \frac{1}{N}\sum_k \mathcal{O}(\boldsymbol{p}_k) \sim \mathcal{O}(\boldsymbol{p}_0)$$

$$\mathrm{V}[\mathcal{O}] = \frac{1}{N}\sum_k [\mathcal{O}(\boldsymbol{p}_k) - \mathrm{E}[\mathcal{O}]]^2$$

$$= \text{hessian, lagrange}$$

Regression strategies

- Data resampling (JAM, NNPDF)

Regression strategies

- Data resampling (JAM, NNPDF)
  - Generate N resampled data $\boxed{\sigma_{i,k} = \sigma_i + R_{i,k}\delta\sigma_i}$

Regression strategies

- Data resampling (JAM, NNPDF)

  + Generate N resampled data $\boxed{\sigma_{i,k} = \sigma_i + R_{i,k}\delta\sigma_i}$

  + $\{\boldsymbol{p}_k : 1...N\}$ from N fits to resampled data

Regression strategies

- Data resampling (JAM, NNPDF)
  - Generate N resampled data $\boxed{\sigma_{i,k} = \sigma_i + R_{i,k}\delta\sigma_i}$
  - $\{\boldsymbol{p}_k : 1...N\}$ from N fits to resampled data
  - Use flat priors as guess for the N fits

Regression strategies

- Other approaches

Regression strategies

- Other approaches
  - Hybrid Markov Chain (Gbedo, Mangin-Brinet)

Regression strategies

- Other approaches
  - + Hybrid Markov Chain (Gbedo, Mangin-Brinet)
  - + Nested sampling (JAM)
    $\rightarrow$ challenging for higher dimensions $O(100)$

JAM19: “A less strange proton”

arXiv:1905.03788

NS, Andres, Ethier, Melnitchouk

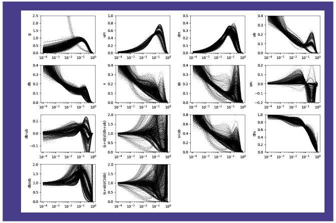
Session KH: Nucleon Structure I
9:50AM , Wednesday, October 16, 2019
Room: Salon B

The JAM 19 challenge

- Simultaneous extraction of $\boxed{f}$s and $\boxed{d}$s
- Dimension of parameter space is $\mathcal{O}(100)$
- NLL evaluation ~ 1 min per point in parameter space

JAM19 multi-step strategy

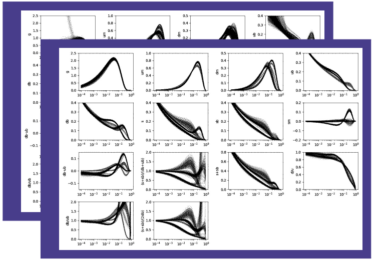
PDFs



+DIS (No HERA)

JAM19 multi-step strategy

PDFs

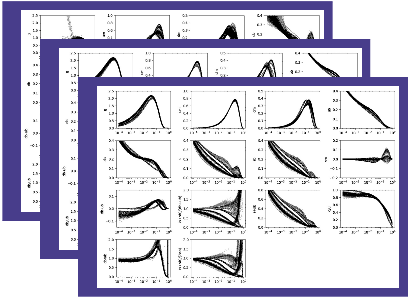


+DIS (No HERA)

+DIS HERA

JAM19 multi-step strategy

PDFs



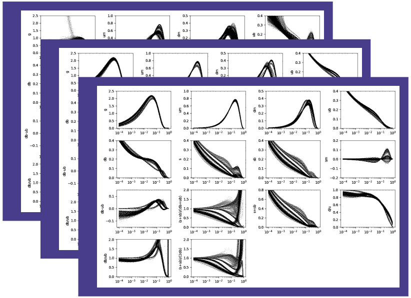
+DIS (No HERA)

+DIS HERA

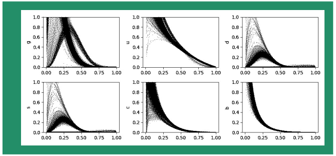
+DY

JAM19 multi-step strategy

PDFs



pion FFs



+DIS (No HERA)

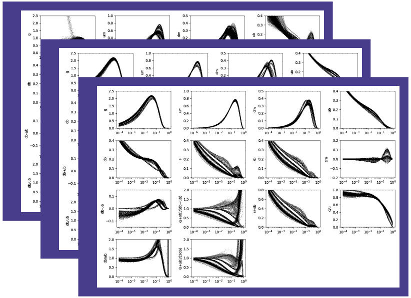
+DIS HERA

+DY

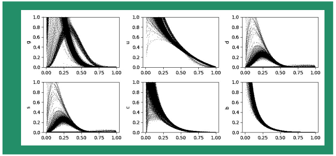
+SIA pions

JAM19 multi-step strategy

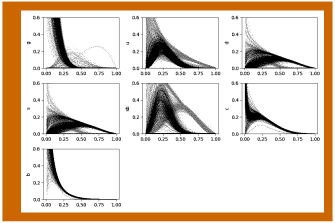
PDFs



pion FFs



kaon FFs



+DIS (No HERA)

+DIS HERA

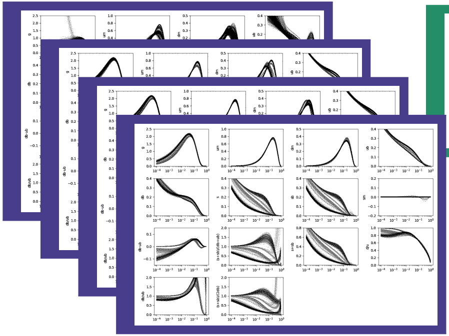
+DY

+SIA pions

+SIA kaons

JAM19 multi-step strategy

PDFs

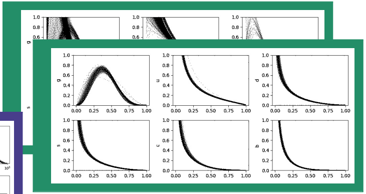


+DIS (No HERA)

+DIS HERA

+DY

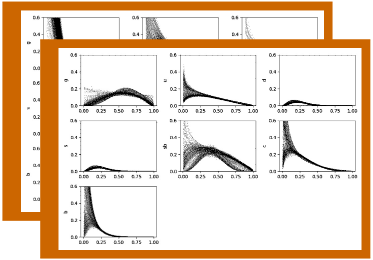
pion FFs



+SIA pions

+SIDIS pions

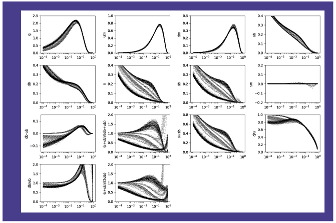
kaon FFs



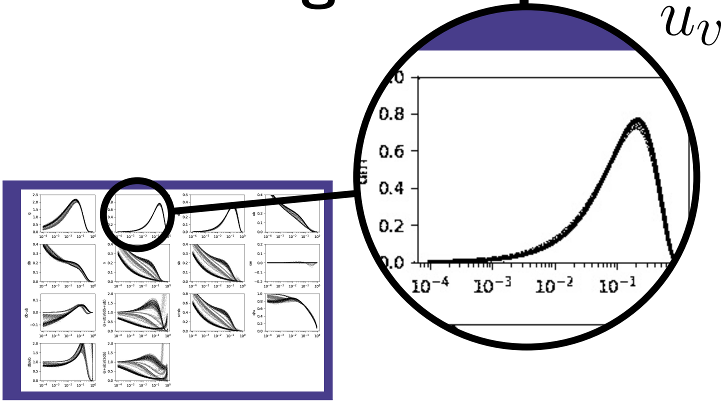
+SIA kaons

+SIDIS kaons

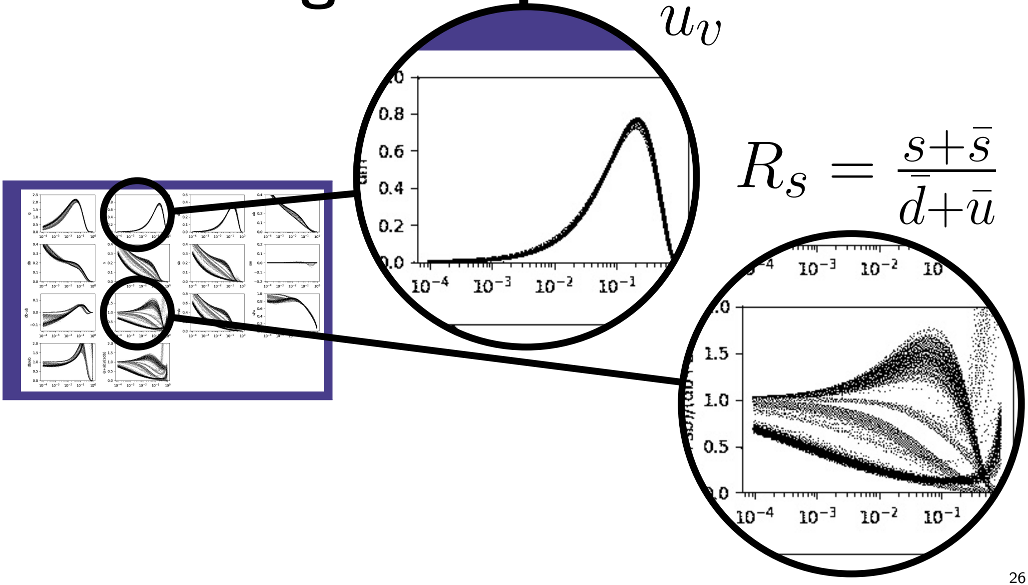
Discriminating multiple solutions



Discriminating multiple solutions

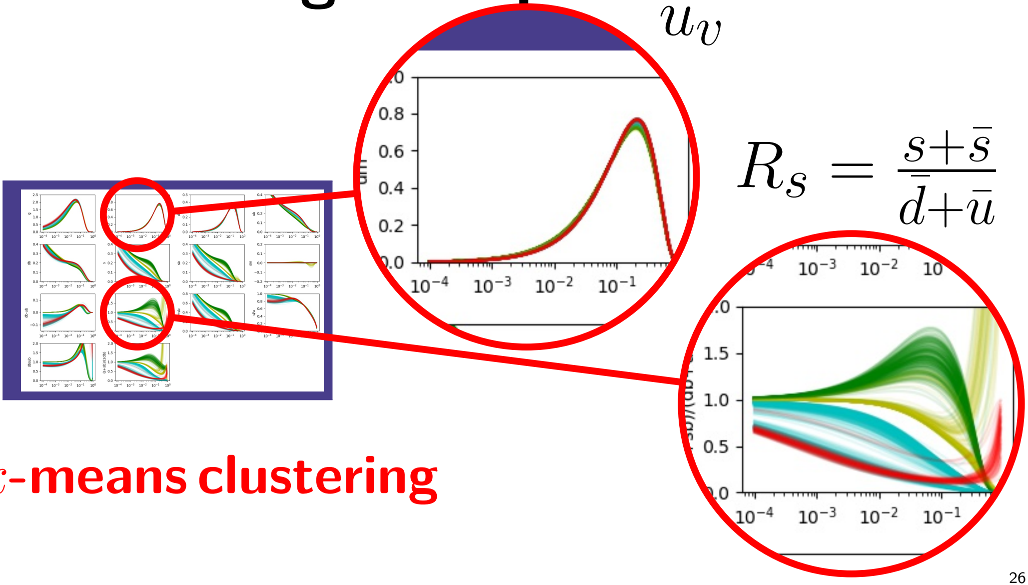


Discriminating multiple solutions



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

Discriminating multiple solutions



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

k-means clustering

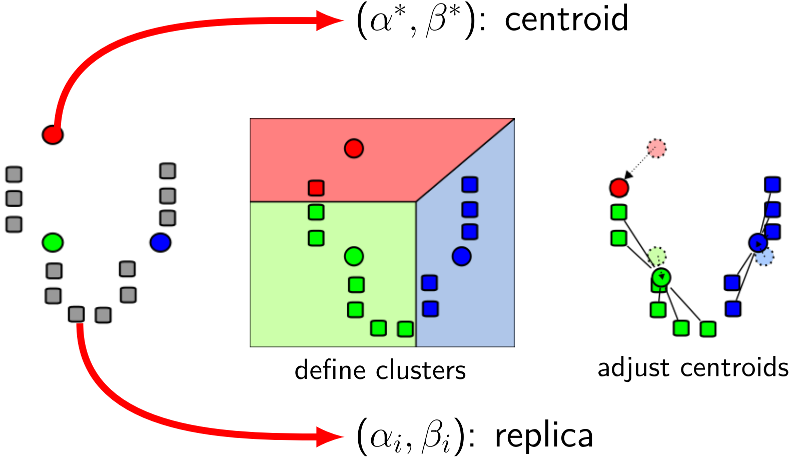
k-means clustering: 2D example

k-means clustering: 2D example

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

k-means clustering: 2D example

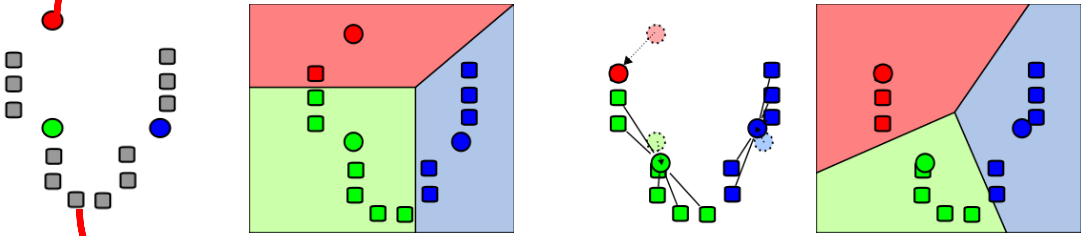
e.g $\boxed{f(x) = x^{\alpha}(1-x)^{\beta}}$



k-means clustering: 2D example

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

(α^*, β^*): centroid



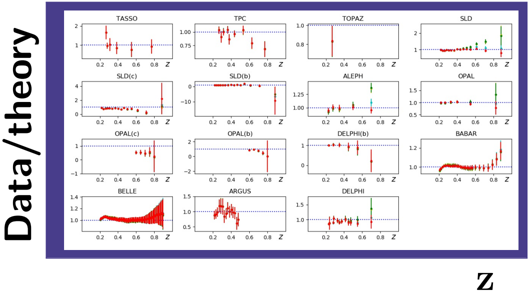
define clusters

adjust centroids

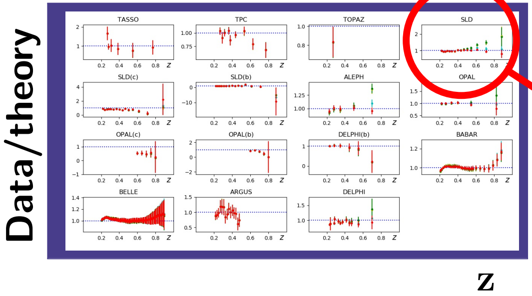
get new clusters

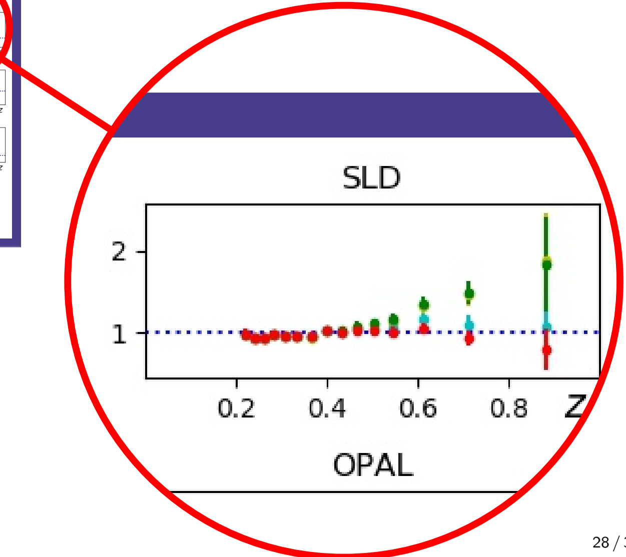
(α_i, β_i): replica

data over theory

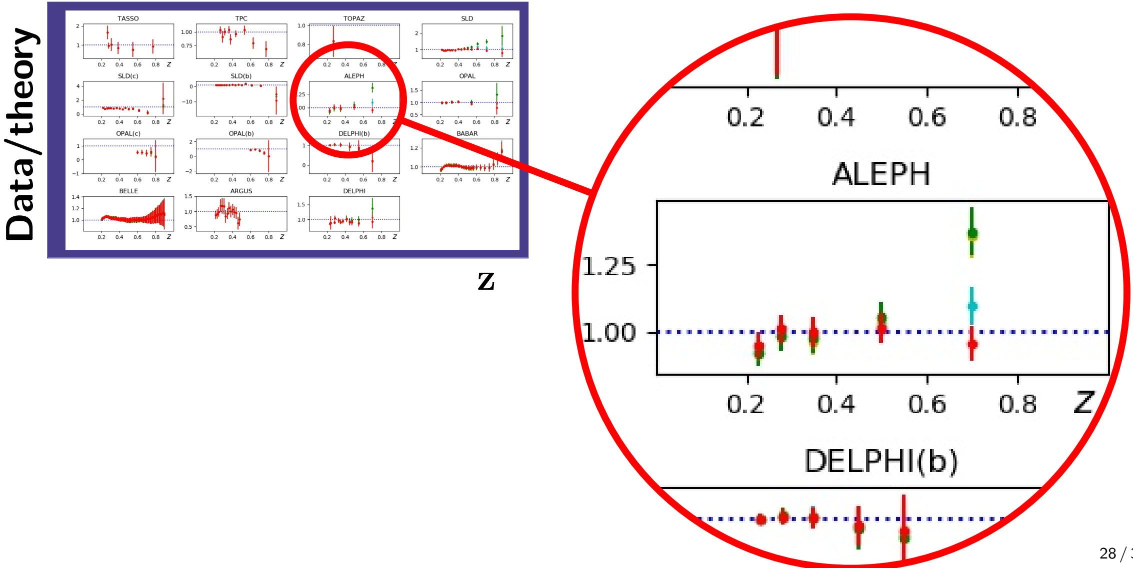


data over theory

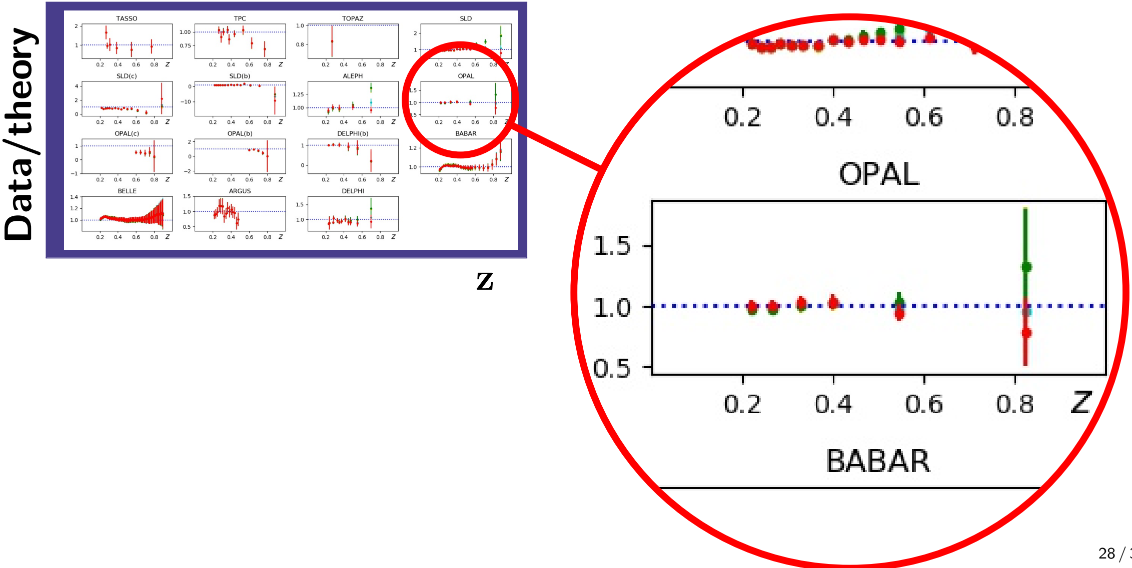




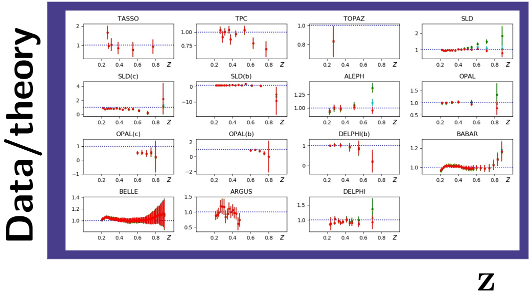
data over theory

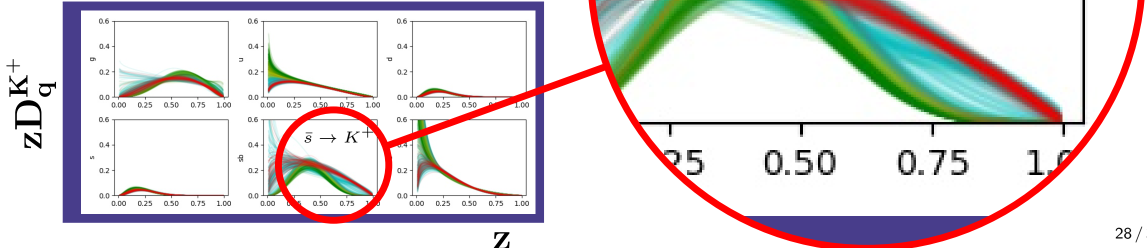


data over theory

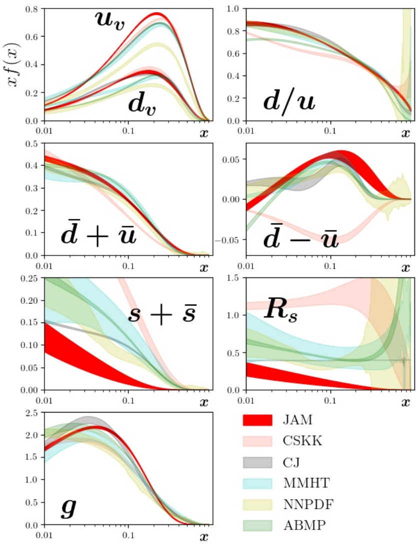


data over theory



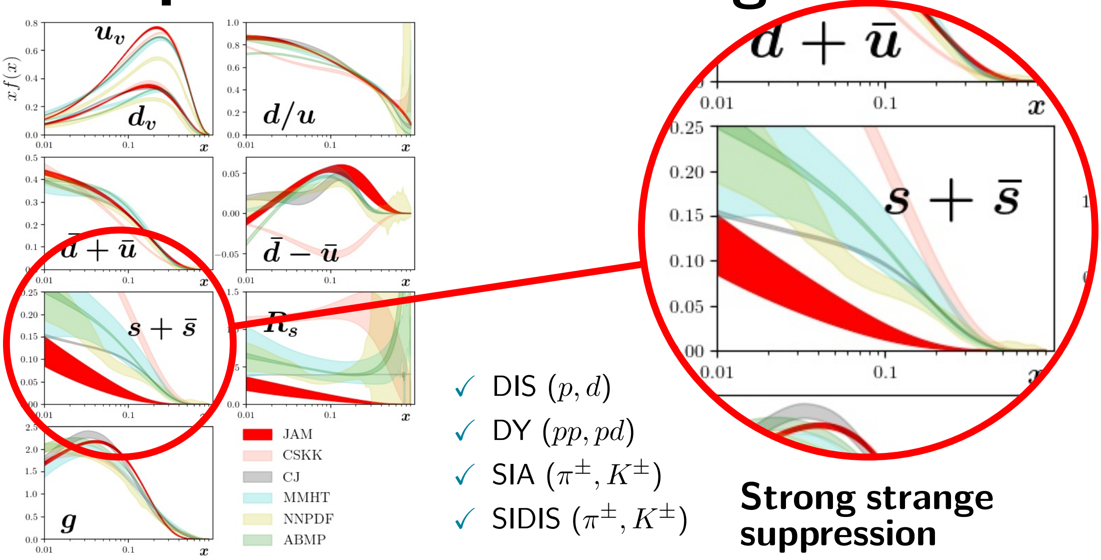


Comparison with other groups



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

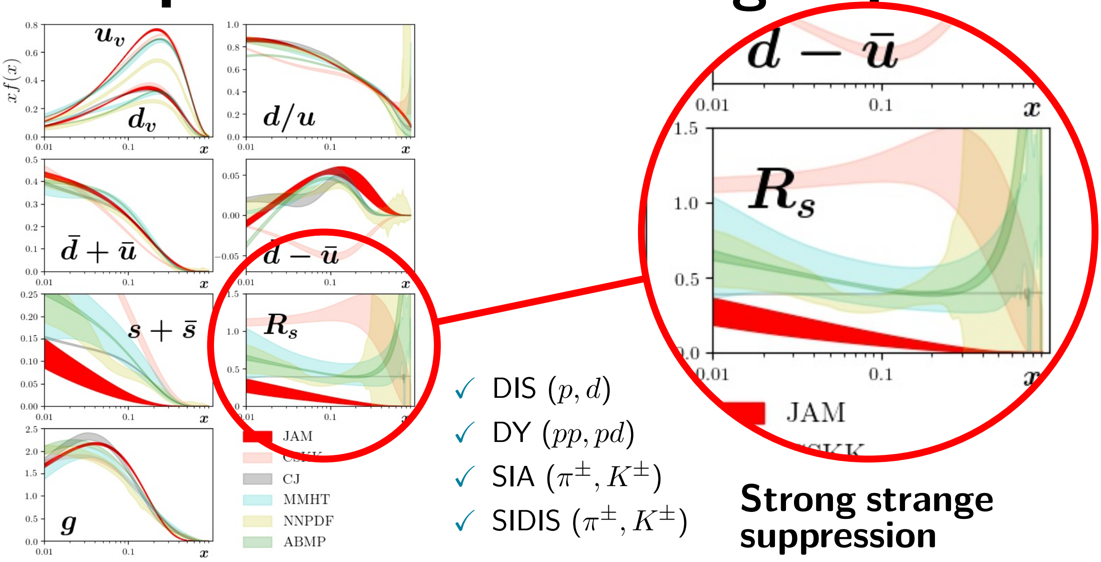
Comparison with other groups



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

Strong strange suppression

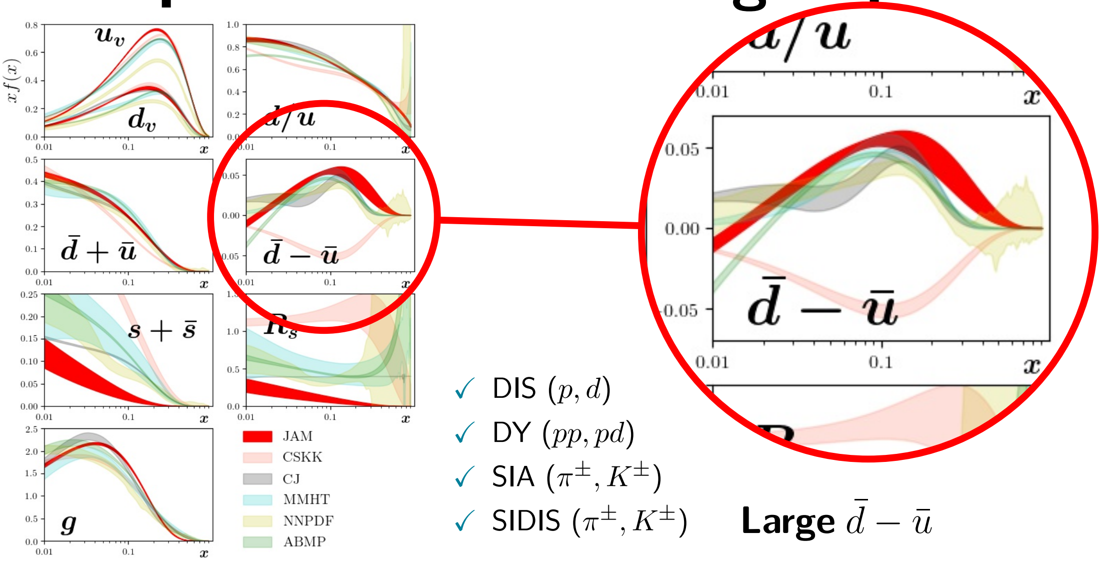
Comparison with other groups



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

Strong strange suppression

Comparison with other groups



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

Large $\bar{d}-\bar{u}$

Summary and outlook

- Understanding hadrons as emergent phenomena of QCD

Summary and outlook

- Understanding hadrons as emergent phenomena of QCD

  + Factorization theorems

Summary and outlook

- Understanding hadrons as emergent phenomena of QCD

  + Factorization theorems
  + Experimental cross sections

Summary and outlook

- Understanding hadrons as emergent phenomena of QCD

  + Factorization theorems
  + Experimental cross sections
  + Global analysis of nucleon structures and hadronization

Summary and outlook

- Challenges of the inverse problem

Summary and outlook

- Challenges of the inverse problem
  - + Efficient sampling of the posterior distribution

Summary and outlook

- Challenges of the inverse problem
  - + Efficient sampling of the posterior distribution
  - + Identification of the best solution

Summary and outlook

- Challenges of the inverse problem
  - Efficient sampling of the posterior distribution
  - Identification of the best solution
  - Treatment of non compatible data sets
    (not discussed in this talk)

Summary and outlook

- Next generation of global analysis tools using Machine Learning

Summary and outlook

- Next generation of global analysis tools using Machine Learning
  - **M. Kuchera** Session FE: Mini-Symposium: Towards a US Electron Ion Collider: Physics, Accelerator, and Detectors II
    11:00 AM, Tuesday, October 15, 2019
    Room: Salon 5

Summary and outlook

- Next generation of global analysis tools using Machine Learning
  - **M. Kuchera** Session FE: Mini-Symposium: Towards a US Electron Ion Collider: Physics, Accelerator, and Detectors II
    11:00 AM, Tuesday, October 15, 2019
    Room: Salon 5
  - **M. Houk & E. Tsitinidi** Session HA: Conference Experience for Undergraduates Poster Session
    4:00 PM, Tuesday, October 15, 2019
    Room: Salon 1