First Monte Carlo analysis of fragmentation functions from single-inclusive e+e- annihilation (arXiv:1609.00899)

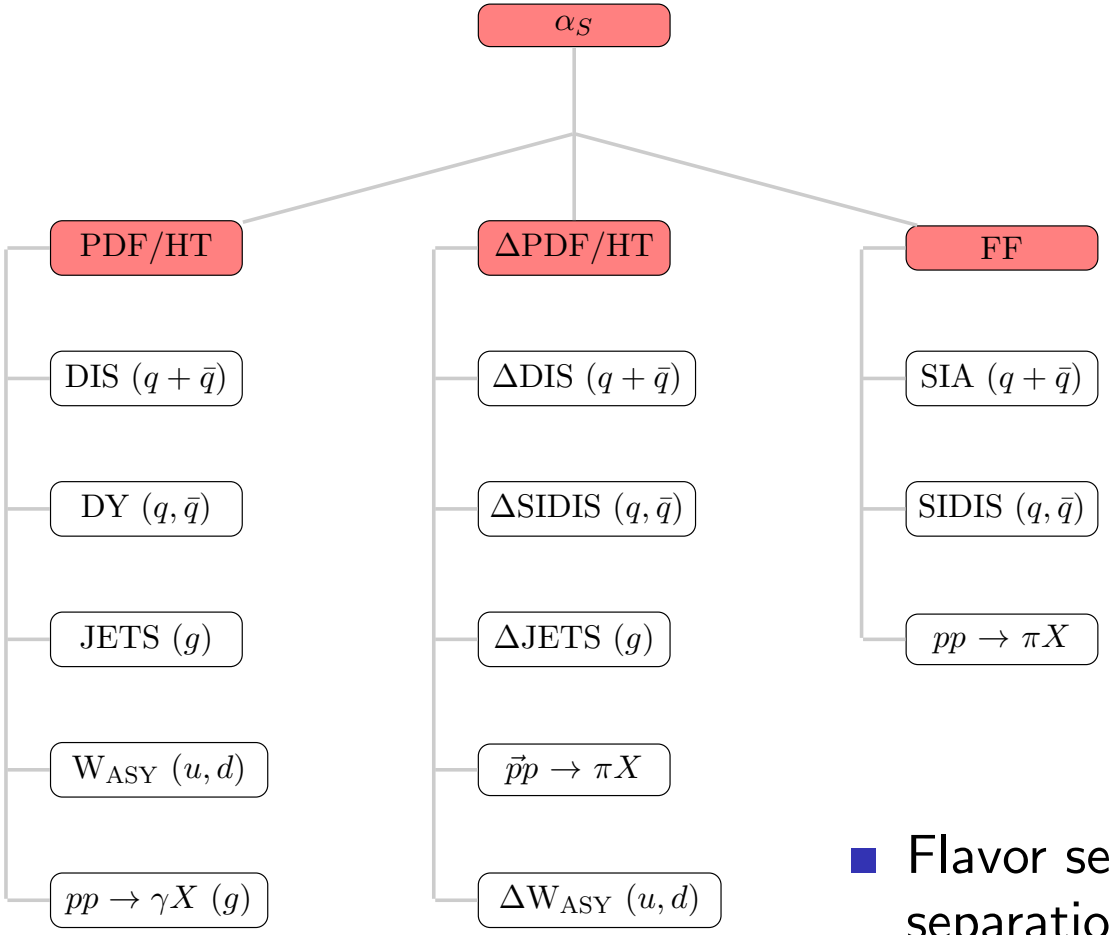
Nobuo Sato

In Collaboration with:

- J. J. Ethier (College of William and Mary)
- W. Melnitchouk (Jefferson Lab)
- A. Accardi (Hampton U. and Jefferson Lab)
- S.Kumano (KEK, J-PARC)
- M.Hirai (Nippon Institute of Technology)

Parton Radiation and Fragmentation from LHC to FCC-ee, 2016

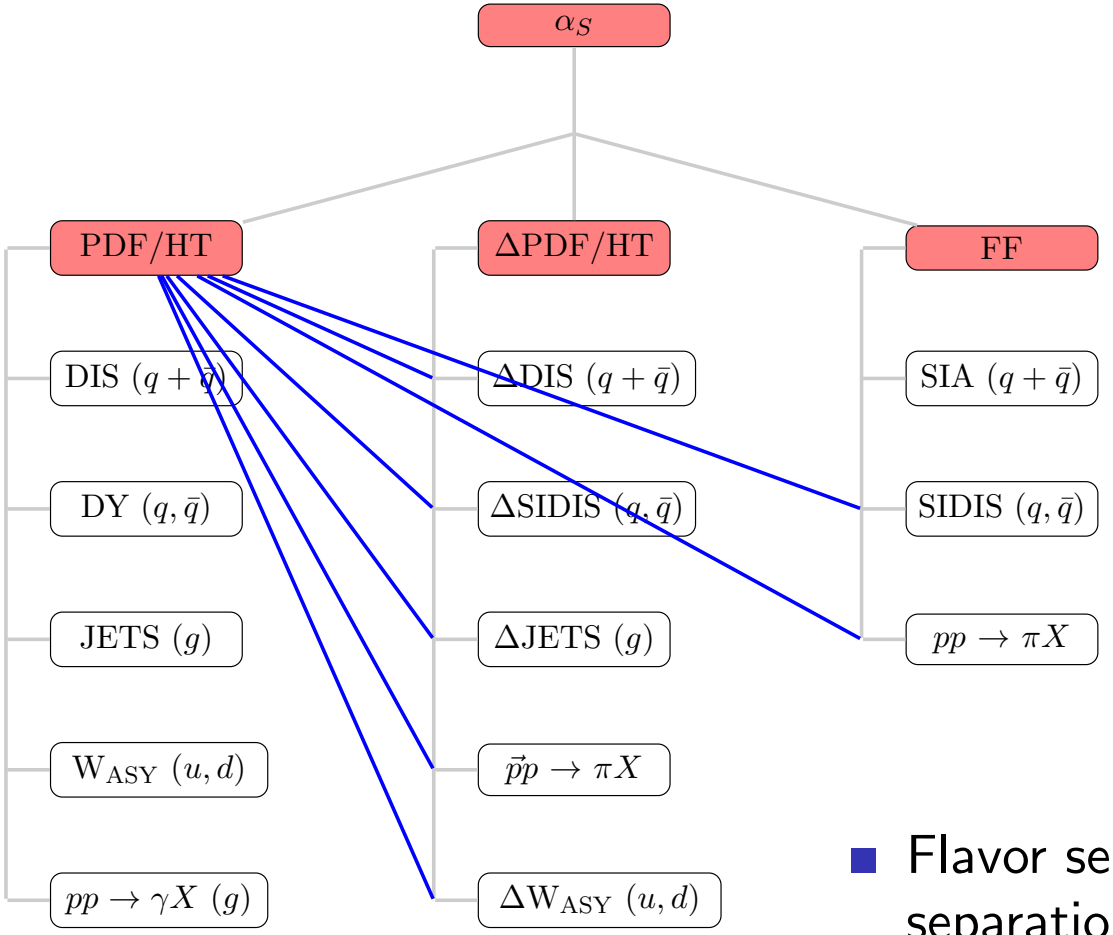
Motivations



- Testing factorization theorems in QCD (Collinear factorization, TMD factorization)
- Understand properties of nucleon structure in terms of PDF, ΔPDF
- Understand formation of hadrons from partons

- Flavor separation, quark antiquark separation can be achieved by considering observables sensitive to combinations of PDF, ΔPDF and FF.
- A full combined analysis is needed

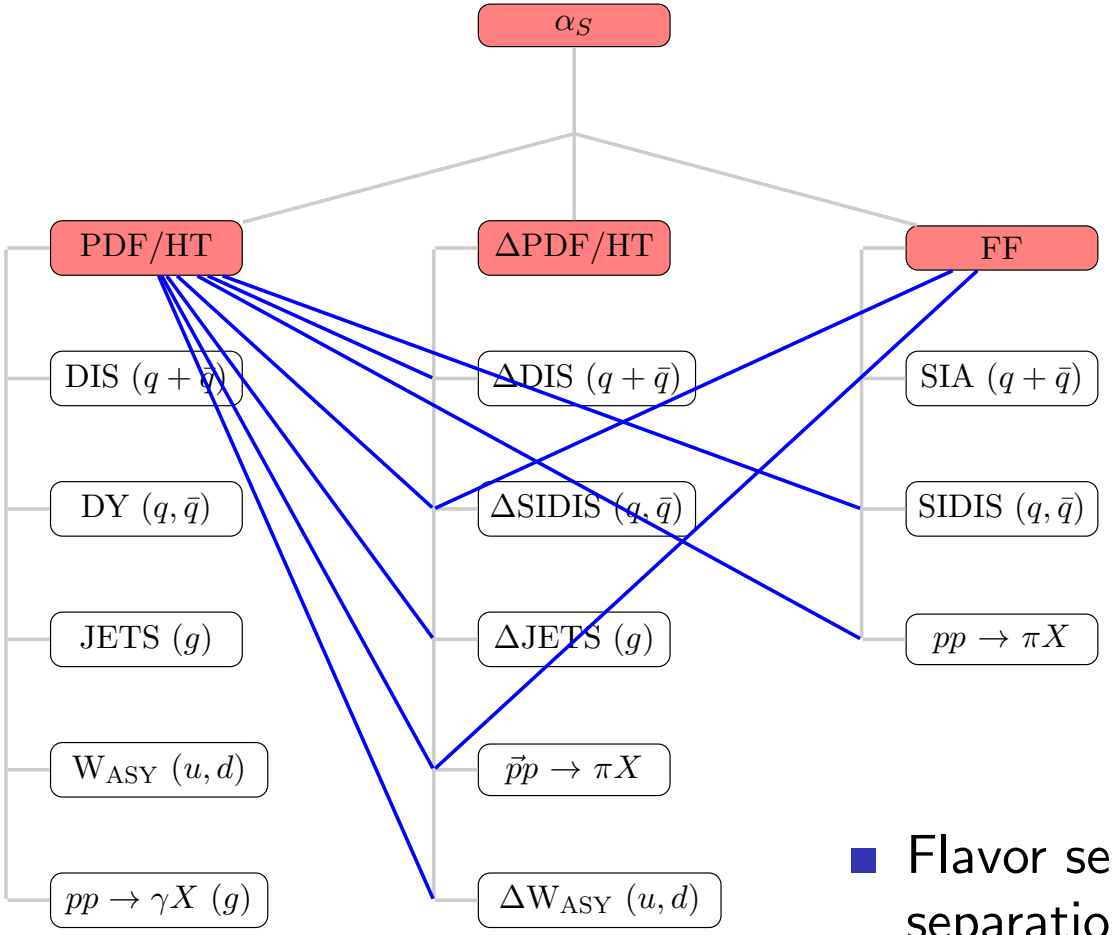
Motivations



- Testing factorization theorems in QCD (Collinear factorization, TMD factorization)
- Understand properties of nucleon structure in terms of PDF, ΔPDF
- Understand formation of hadrons from partons

- Flavor separation, quark antiquark separation can be achieved by considering observables sensitive to combinations of PDF, ΔPDF and FF.
- A full combined analysis is needed

Motivations



- Testing factorization theorems in QCD (Collinear factorization, TMD factorization)
- Understand properties of nucleon structure in terms of PDF, ΔPDF
- Understand formation of hadrons from partons

- Flavor separation, quark antiquark separation can be achieved by considering observables sensitive to combinations of PDF, ΔPDF and FF.
- A full combined analysis is needed

Theory of fitting

The goal is to estimate:

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

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

- $\mathcal{O}$=$\Delta q, D_q^h, ...$
- $\sqrt{V[\mathcal{O}]} = 1\sigma$
- $\boldsymbol{a} \rightarrow$ PDF/FF parameters
- $\mathcal{P}(\boldsymbol{a}|data) \propto \mathcal{L}(data|\boldsymbol{a})\pi(\boldsymbol{a})$
- $\mathcal{L}(data|\boldsymbol{a}) \propto \exp\left(-\frac{1}{2}\chi^2(\boldsymbol{a})\right)$
- $\chi^2(\boldsymbol{a}) = \sum_i \left(\frac{D_i - T_i(\boldsymbol{a})}{\delta D_i}\right)^2$
- $\pi(\boldsymbol{a}) \rightarrow$*priors*. i.e $\prod_i \theta(a_i - a_i^0)$

Theory of fitting

The goal is to estimate:

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

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

- $\mathcal{O}$=$\Delta q, D_q^h, ...$
- $\sqrt{V[\mathcal{O}]} = 1\sigma$
- $\boldsymbol{a} \rightarrow$ PDF/FF parameters
- $\mathcal{P}(\boldsymbol{a}|data) \propto \mathcal{L}(data|\boldsymbol{a})\pi(\boldsymbol{a})$
- $\mathcal{L}(data|\boldsymbol{a}) \propto \exp\left(-\frac{1}{2}\chi^2(\boldsymbol{a})\right)$
- $\chi^2(\boldsymbol{a}) = \sum_i \left(\frac{D_i - T_i(\boldsymbol{a})}{\delta D_i}\right)^2$
- $\pi(\boldsymbol{a}) \rightarrow$*priors*. i.e $\prod_i \theta(a_i - a_i^0)$
- **How to evaluate $E[\mathcal{O}], V[\mathcal{O}]$?**

Theory of fitting

The goal is to estimate:

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

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

Monte Carlo methods

- $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \{\boldsymbol{a}_k\}$
- $\mathrm{E}[\mathcal{O}] \approx \frac{1}{N} \sum_k \mathcal{O}(\boldsymbol{a}_k)$
- $\mathrm{V}[\mathcal{O}] \approx \frac{1}{N} \sum_k [\mathcal{O}(\boldsymbol{a}_k) - \mathrm{E}[\mathcal{O}]]^2$

Maximum Likelihood

- Maximize $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \boldsymbol{a}_0$
- $\mathrm{E}[\mathcal{O}] \approx \mathcal{O}\left(\boldsymbol{a}_0\right)$
- $\mathrm{V}[\mathcal{O}] \approx$hessian, $\Delta\chi^2$ envelope,...

Theory of fitting

The goal is to estimate:

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

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

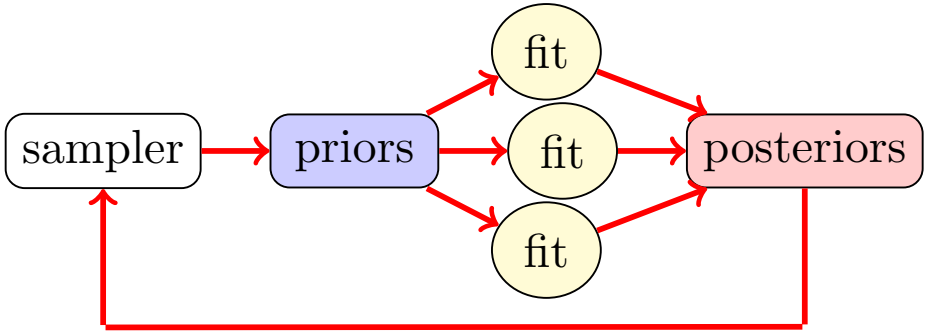
Monte Carlo methods

- $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \{\boldsymbol{a}_k\}$
- $\mathrm{E}[\mathcal{O}] \approx \frac{1}{N} \sum_k \mathcal{O}(\boldsymbol{a}_k)$
- $\mathrm{V}[\mathcal{O}] \approx \frac{1}{N} \sum_k [\mathcal{O}(\boldsymbol{a}_k) - \mathrm{E}[\mathcal{O}]]^2$

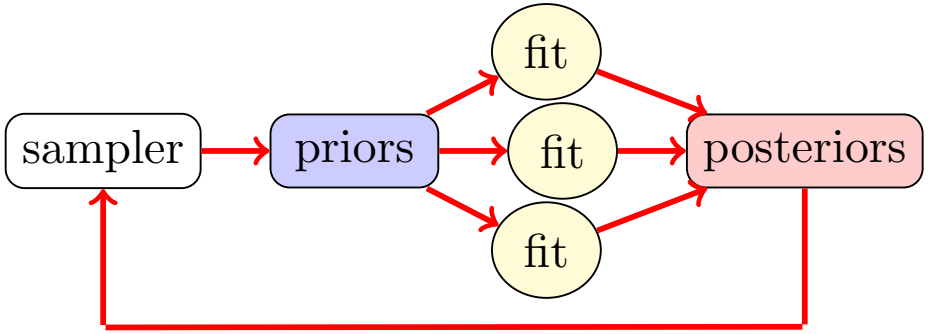
Maximum Likelihood

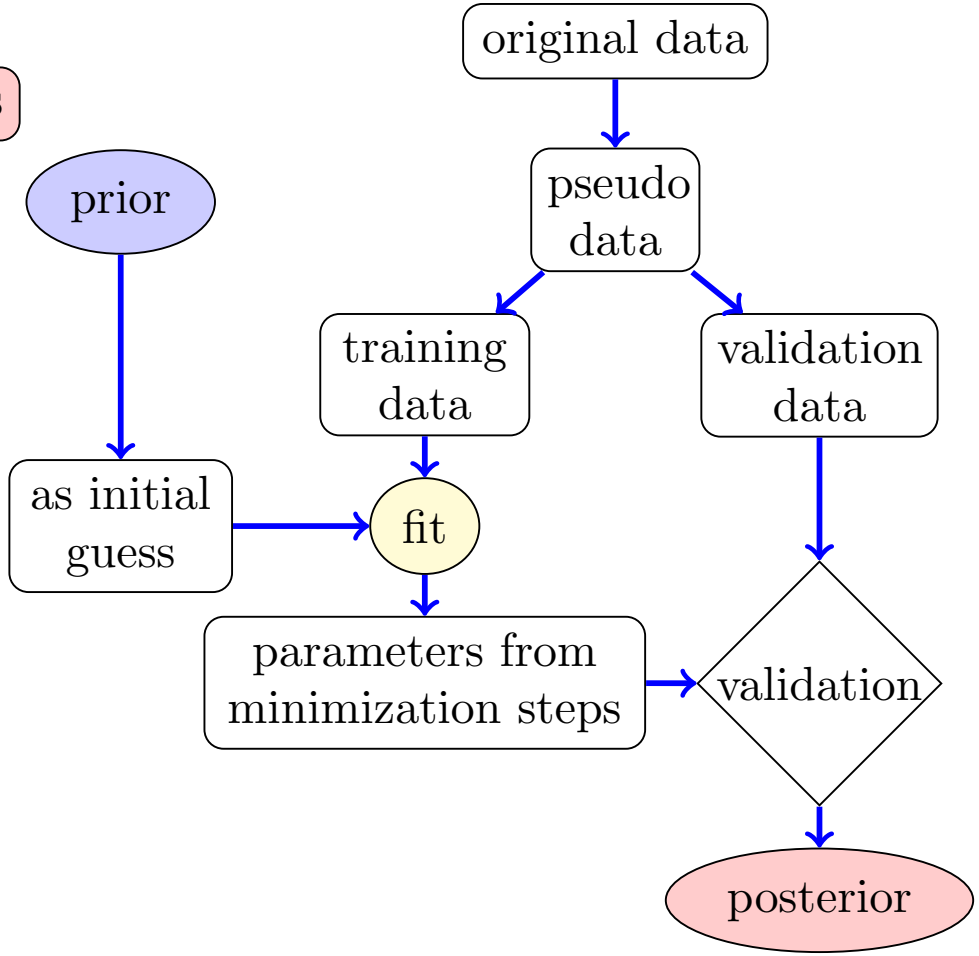
- Maximize $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \boldsymbol{a}_0$
- $\mathrm{E}[\mathcal{O}] \approx \mathcal{O}\left(\boldsymbol{a}_0\right)$
- $\mathrm{V}[\mathcal{O}] \approx$hessian, $\Delta\chi^2$ envelope,...

Iterative Monte Carlo analysis (IMC)

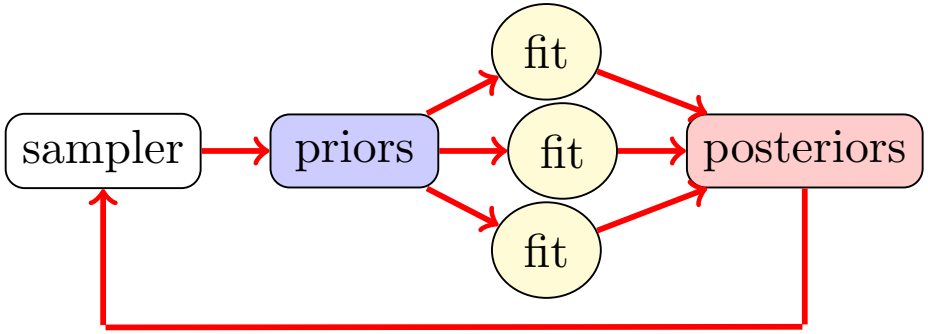


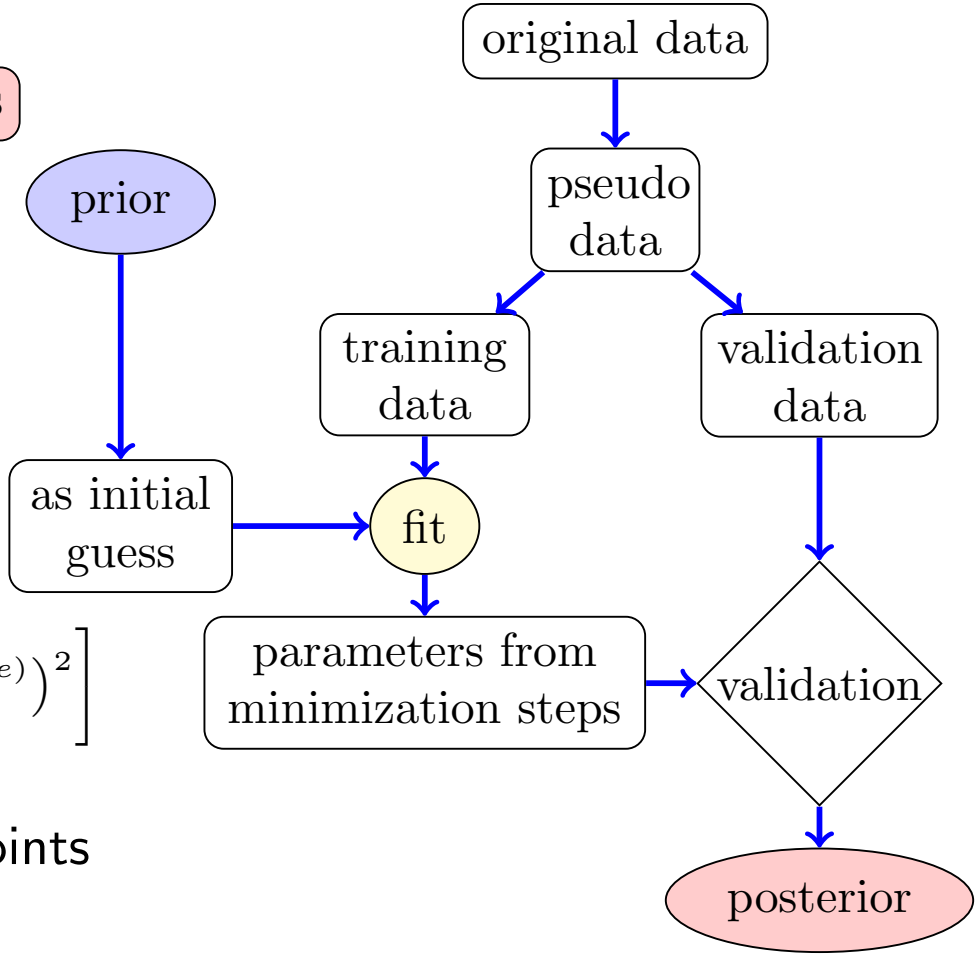
Iterative Monte Carlo analysis (IMC)





Iterative Monte Carlo analysis (IMC)





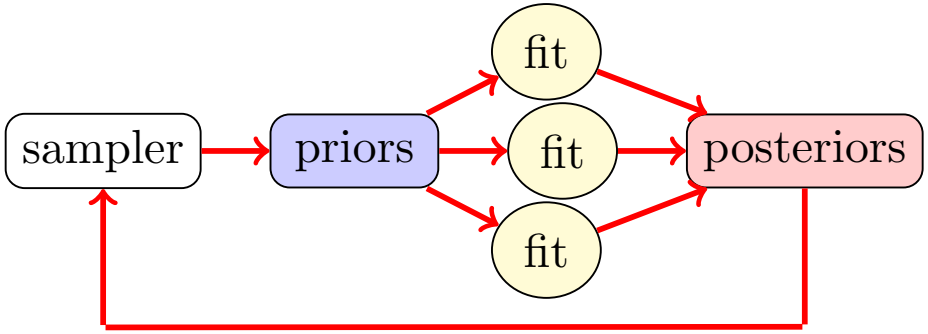
- Chi-squared function

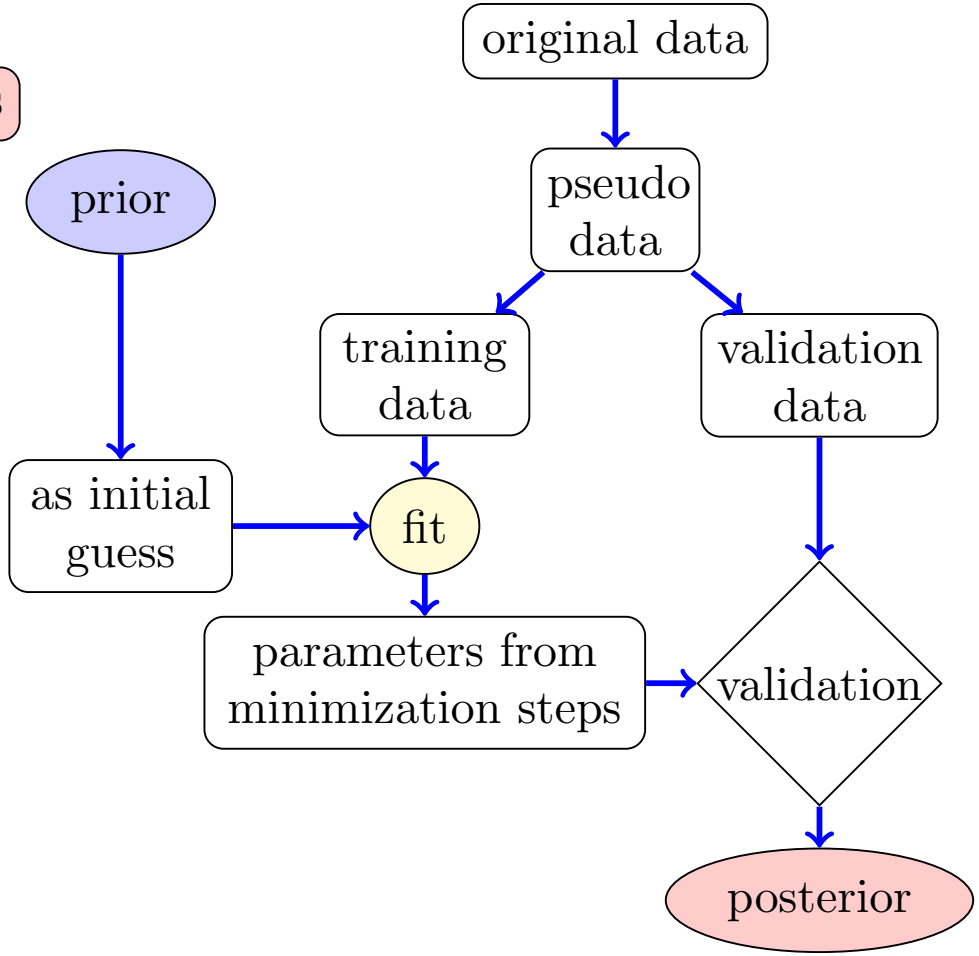
$$N_i^{(e)} = 1 - \sum_k \frac{r_k^{(e)} \beta_{k,i}^{(e)}}{\mathcal{D}_i^{(e)}}$$

$$\chi^2(\boldsymbol{a}) = \sum_e \left[\sum_i \left(\frac{\mathcal{D}_i^{(e)} N_i^{(e)} - T_i^{(e)}}{\alpha_i^{(e)} N_i^{(e)}} \right)^2 + \sum_k \left(r_k^{(e)} \right)^2 \right]$$

- $\mathcal{D}_i^{(e)}$ and $T_i^{(e)}$: the data and theory points
- $\alpha_i^{(e)} = \sqrt{\sigma_{\text{uncorr}}^2 + \sigma_{\text{corr}}^2}$
- $\beta_{k,i}^{(e)}$: correlated systematic uncertainties
- $r_k^{(e)}$: Hyperparameters

Iterative Monte Carlo analysis (IMC)





- Use traditional anzat
  $xf(x) = Nx^a(1-x)^b(1+c\sqrt{x}+dx)$
- Keep all the parameters free.
  No assumptions on the exponents
- Avoid over-fitting by Cross-Validation
- **Iterative procedure**
  $\rightarrow$ Adaptive MC integration (like in Vegas)
- Robust estimation of uncertainties

Details of SIA analysis

- SIA theory calculated at NLO in pQCD
- We use all existing SIA data
- $z > 0.05$ for data at $Q \simeq M_z$
- $z > 0.1$ for data for low energy data
- We only fit $D^h_{q^+} = D^h_q + D^h_{\bar{q}}$ and D^h_g
- Template function $\rightarrow \mathrm{T}(z;\boldsymbol{a}) = M \frac{z^\alpha (1-z)^\beta}{\int_0^1 dz\, z^{1+\alpha}(1-z)^\beta}$

- For pions

$$D^{\pi^+}_{u^+} = D^{\pi^+}_{d^+} = \mathrm{T}(z;\boldsymbol{a}^\pi_{ud}) + \mathrm{T}(z;\boldsymbol{a}'^\pi_{ud})$$
$$D^{\pi^+}_{s^+,\,c^+,\,b^+,\,g} = \mathrm{T}(z;\boldsymbol{a}^\pi_{s,\,c,\,b,\,g})$$

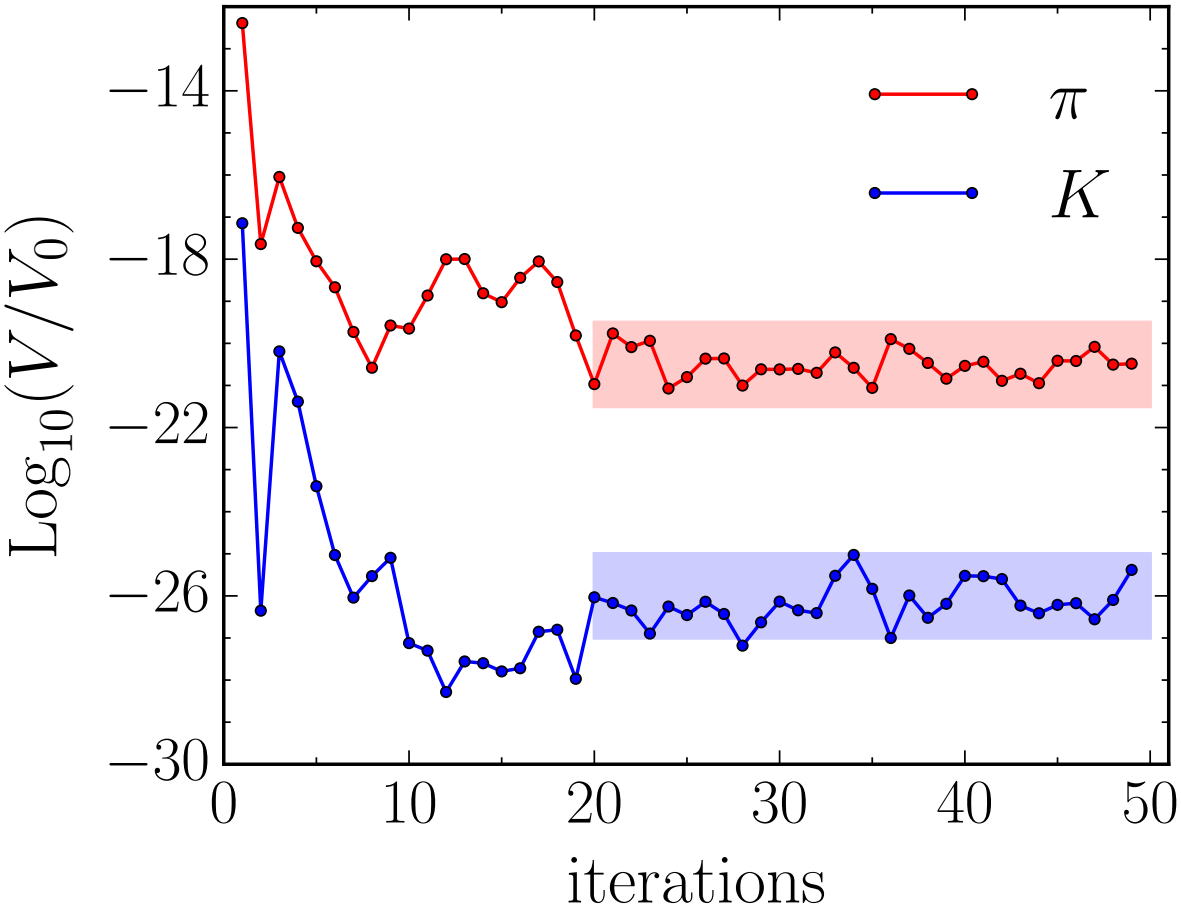
- For kaons

$$D^{K^+}_{s^+} = \mathrm{T}(z;\boldsymbol{a}^K_s) + \mathrm{T}(z;\boldsymbol{a}'^K_s)$$
$$D^{K^+}_{u^+} = \mathrm{T}(z;\boldsymbol{a}^K_u) + \mathrm{T}(z;\boldsymbol{a}'^K_u)$$
$$D^{K^+}_{d^+,\,c^+,\,b^+,\,g} = \mathrm{T}(z;\boldsymbol{a}^K_{d,\,c,\,b,\,g})$$

Results

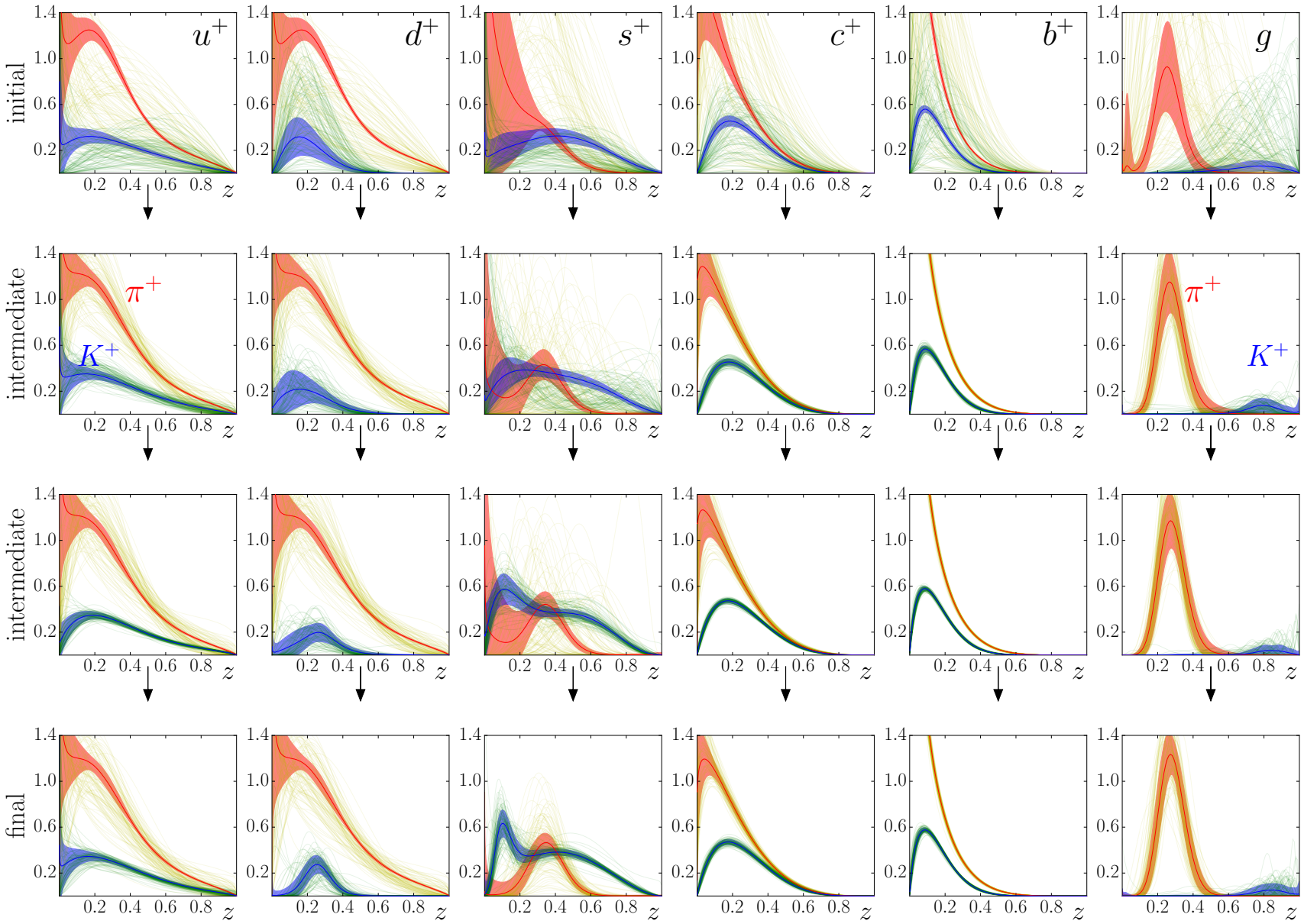
experiment	observable	Q (GeV)	pions $N_{\rm dat}$	pions norm.	pions χ^2	kaons $N_{\rm dat}$	kaons norm.	kaons χ^2
ARGUS	inclusive	9.98	35	1.024 (1.058)	51.1 (55.8)	15	1.007	8.5
Belle	inclusive	10.52	78	0.900 (0.919)	37.6 (21.7)	78	0.988	10.9
BaBaR	inclusive	10.54	39	0.993 (0.948)	31.6 (70.7)	30	0.992	4.9
TASSO	inclusive	12-44	29	$(*)$	37.0 (38.8)	18	$(*)$	14.3
TPC	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	inclusive	29.00	2	1	4.2 (6.2)	3	1	0.3
TOPAZ	inclusive	58.00	4	1	4.8 (6.3)	3	1	0.9
OPAL	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
ALEPH	inclusive	91.20	22	0.987 (0.989)	15.6 (20.4)	18	1.008	6.1
DELPHI	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	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
					$\chi^2/N_{\rm dat}$ =1.31 (1.46)			$\chi^2/N_{\rm dat}$ =1.01

Results

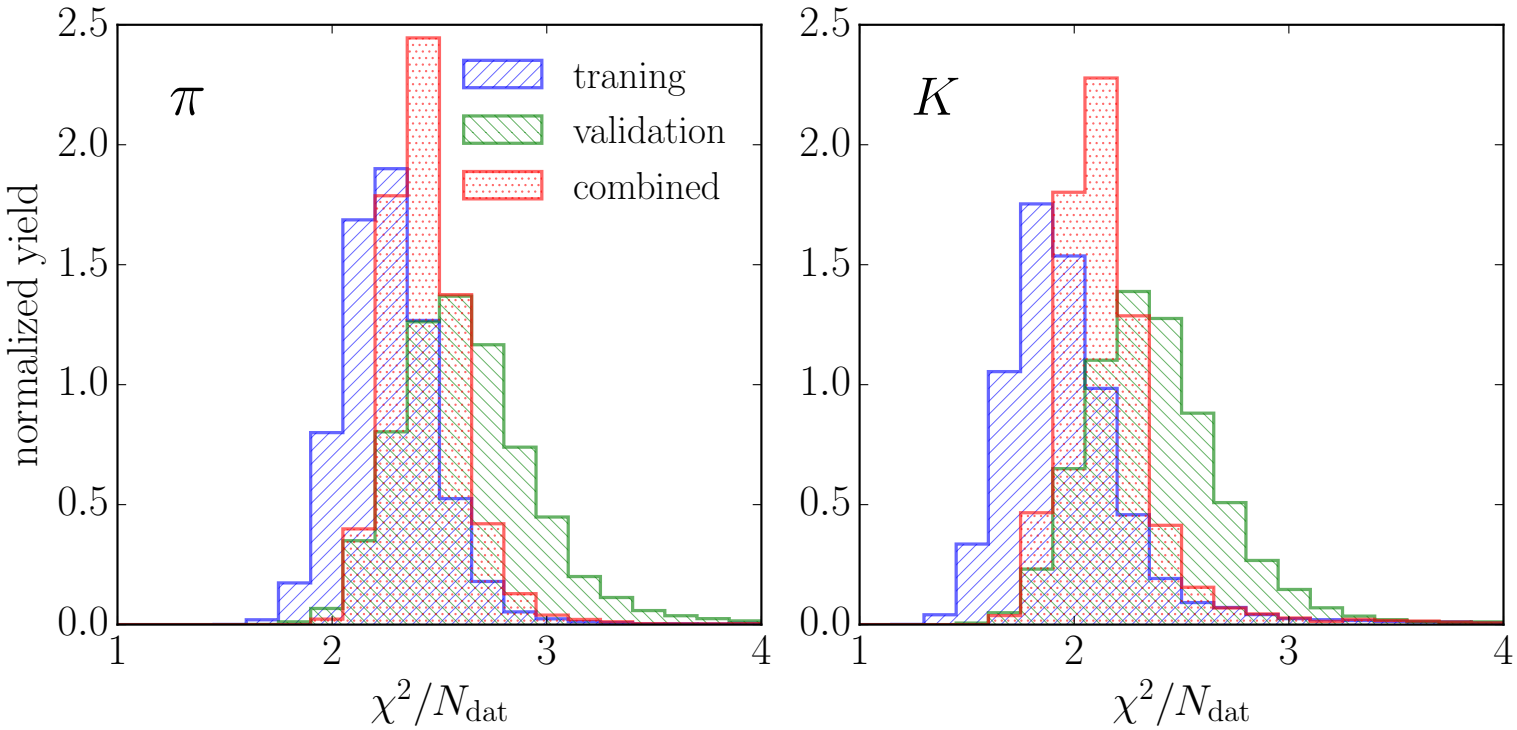


- Approximate convergence of the priors is achieved after 20 iterations

IMC in action

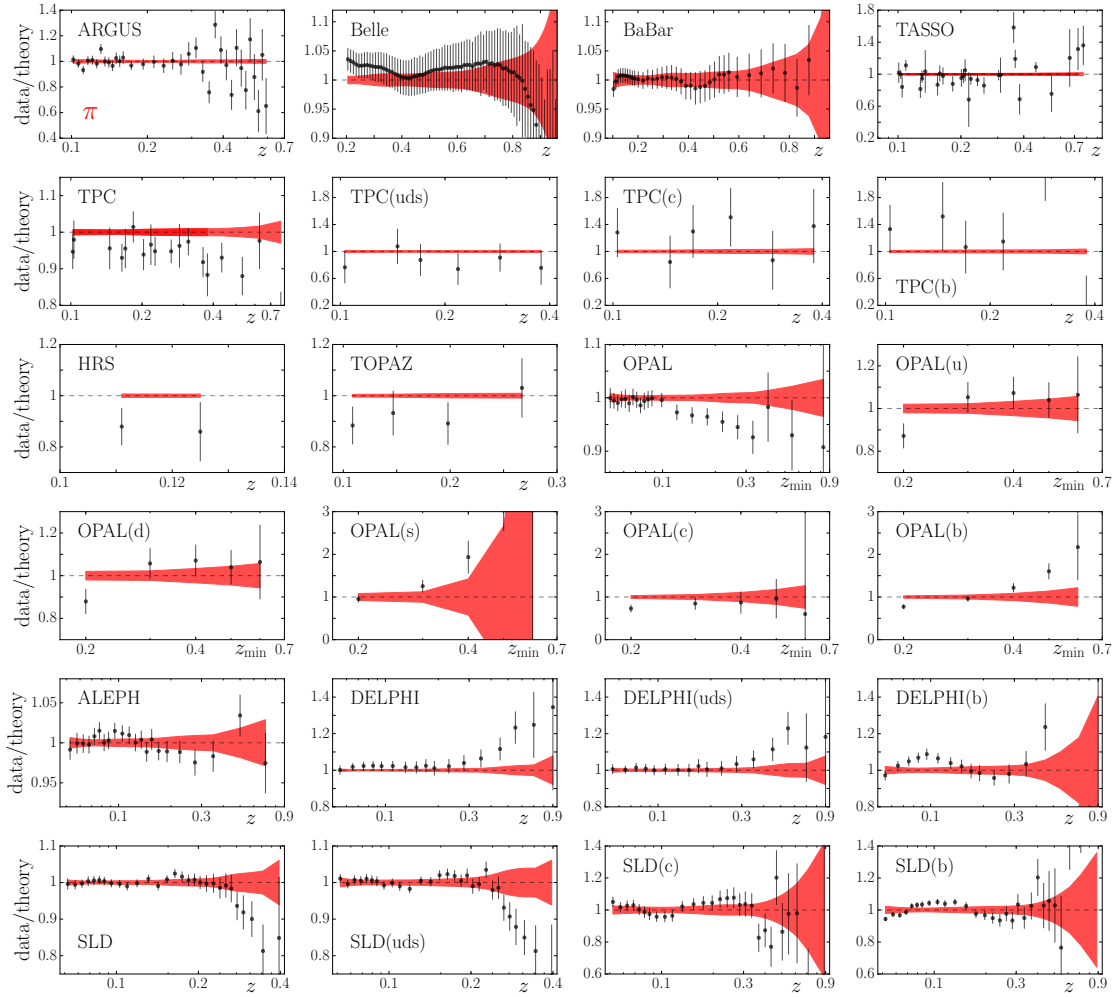


Results



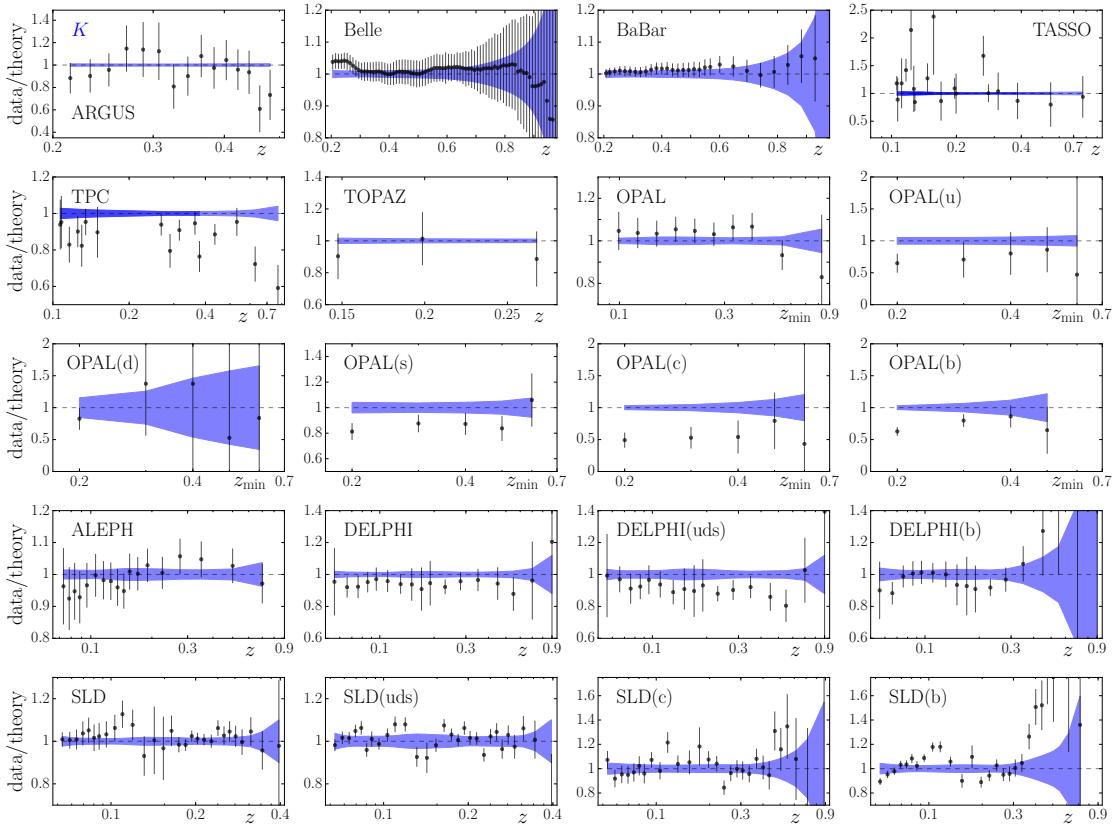
- The shifts between training and validations shows tensions among the data sets
- The Chi-squared is centered around 2 as expected from data resampling

π analysis $(\chi^2/N_{\rm npts} = 1.31)$



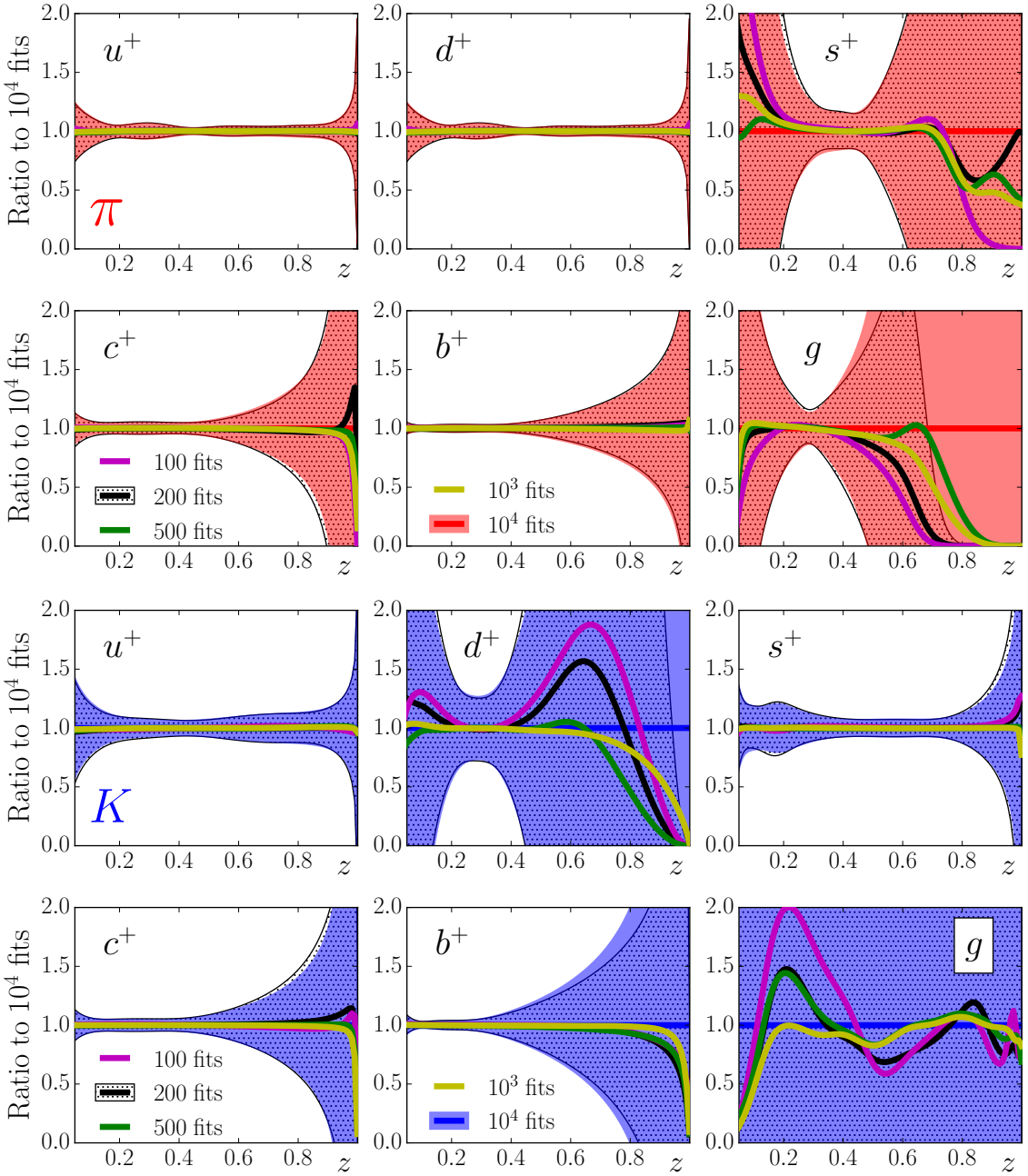
- $z_{\rm cut} > 0.1$ for low energies
- $z_{\rm cut} > 0.05$ for high energies
- We use BaBar prompt data set
- Belle data set needs 10% normalization
- Good agreement at low–z for inclusive data sets
- Data inconsistencies at large–z for $Q^2 = M_z^2$

K analysis $(\chi^2/N_{\mathrm{npts}} = 1.01)$



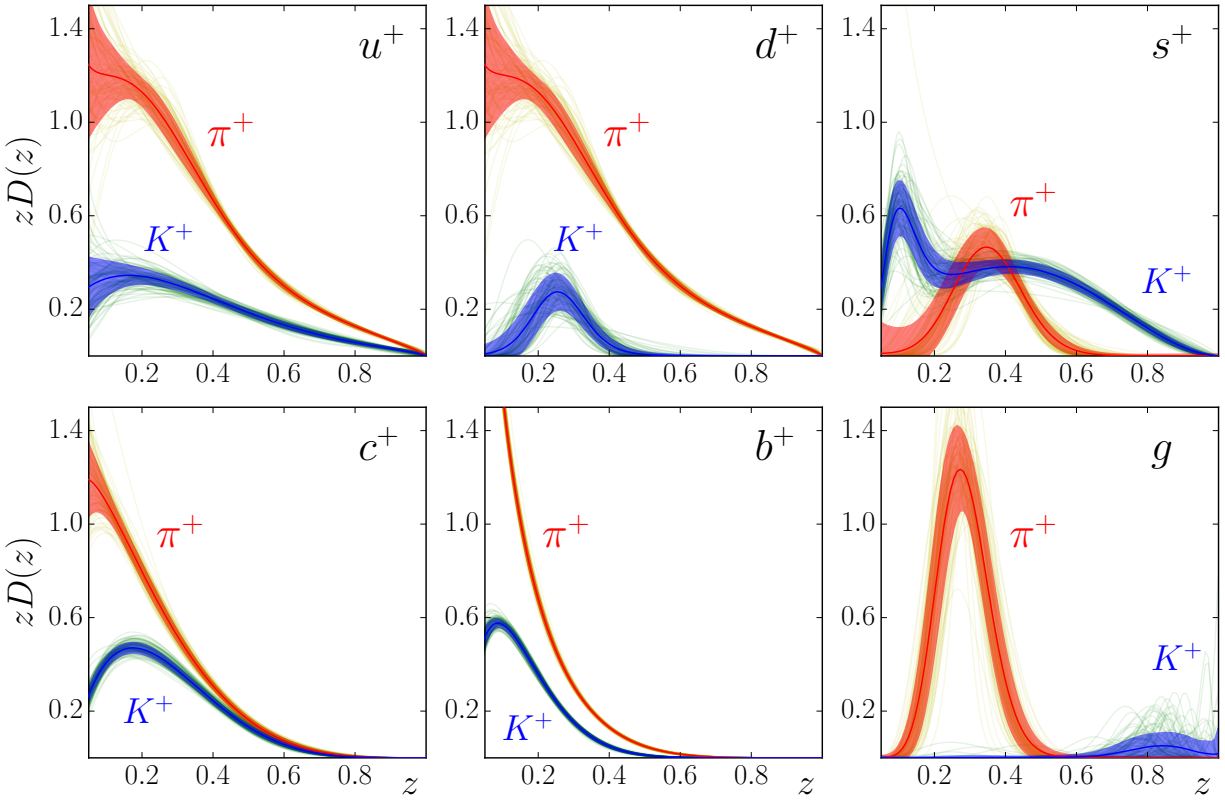
- $z_{\mathrm{cut}} > 0.2$ for low energies to avoid hadron mass corrections
- $z_{\mathrm{cut}} > 0.05$ for high energies
- Smaller χ^2 that π due to larger errors
- Consistent shapes across all z
- Inconsistencies mostly due to normalizations

Results



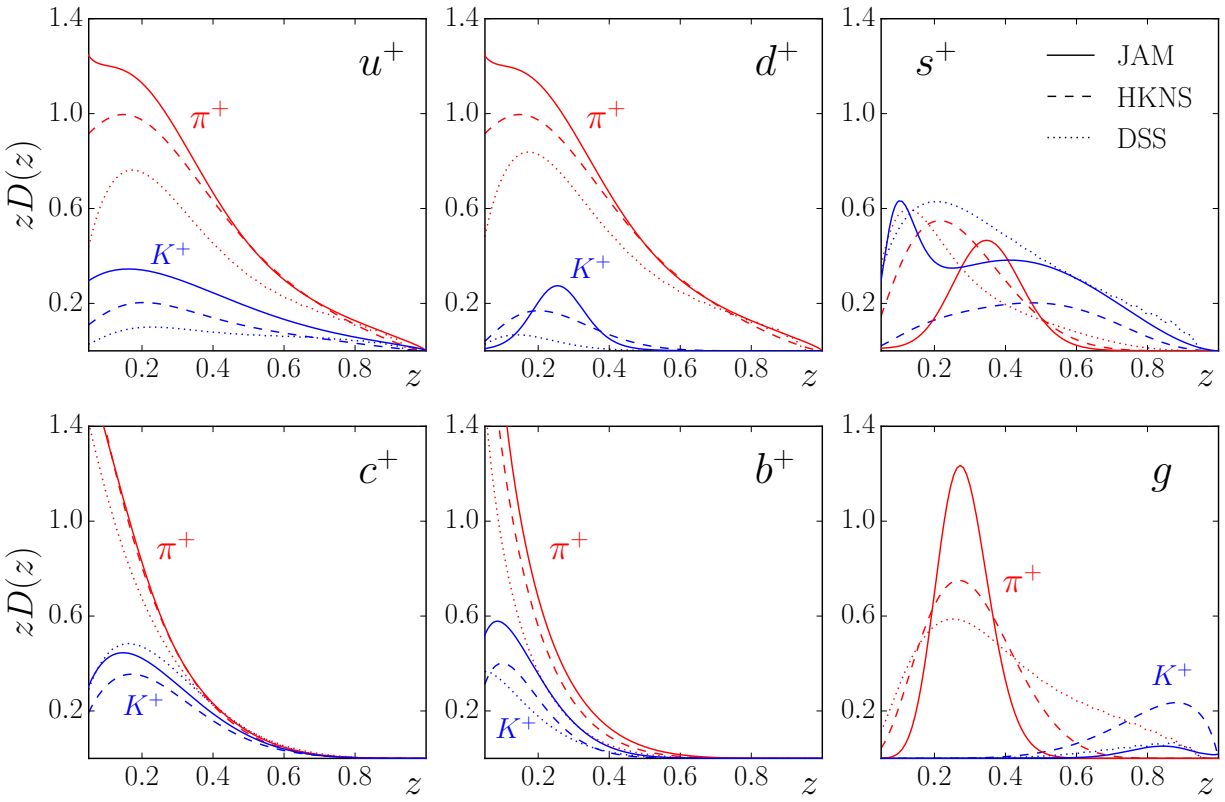
- Larger constraints on favored FFs than unfavored FFs
- More sensitivity on D_g^π than D_g^K
- D_g^K is unknown
- HQ tagged data provides similar constraints for π and K
- Convergence attained with ~ 200 posteriors

Fragmentation functions



- Similar behavior of unfavored $D^{\pi}_{s^+}$ and $D^{K}_{d^+}$
- In contrast the D^{π}_{g} and D^{K}_{g} behave differently
- The charm and bottom FFs become compatible at large–z
- Favored $D^{\pi}_{u^+}$ and $D^{K}_{s^+}$ have similar shape at large–z

Fragmentation functions



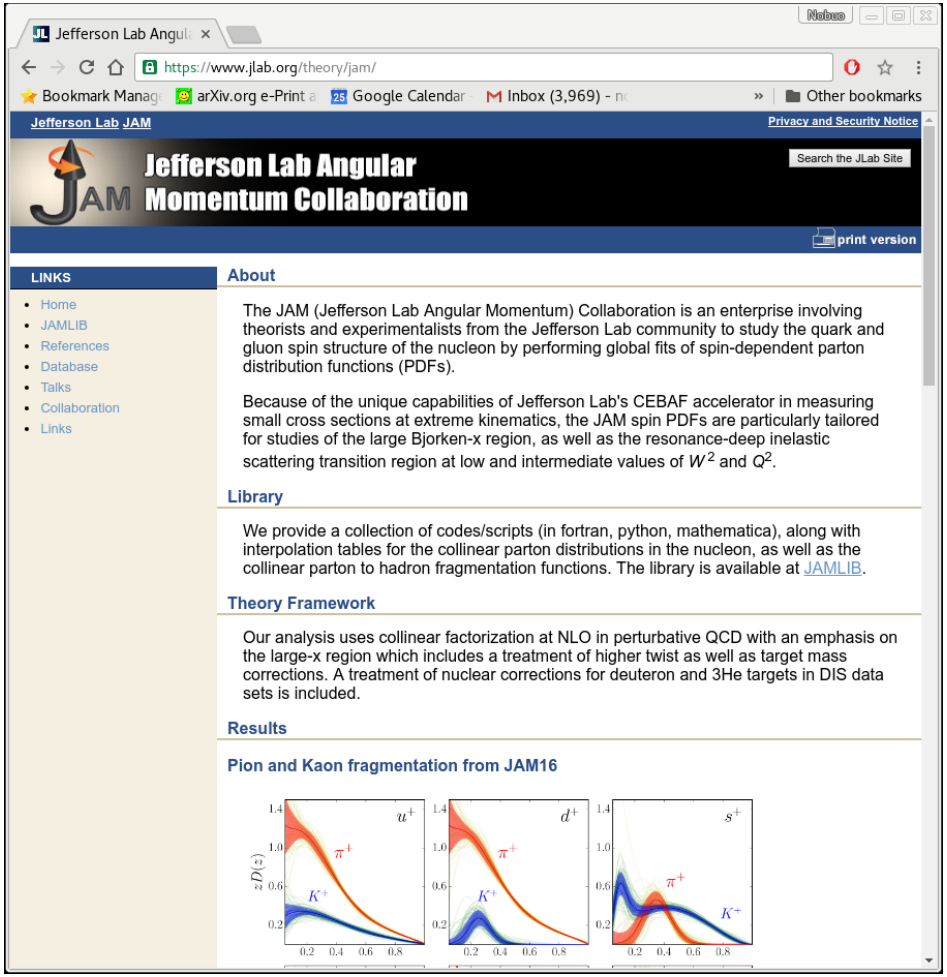
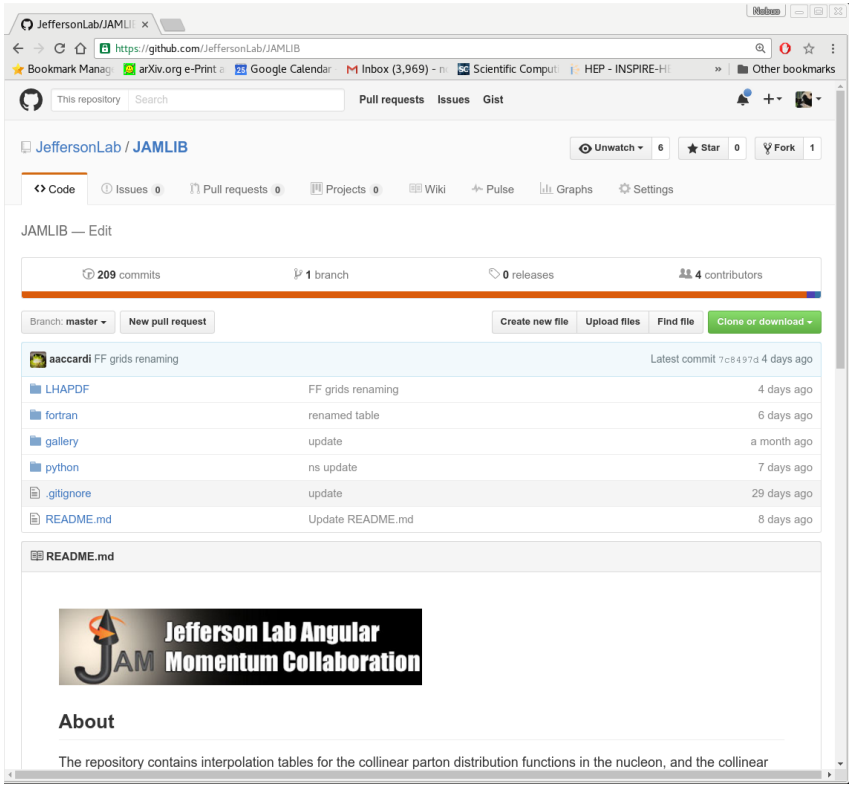
- Overall similar behavior at large-z
- JAM $D_{s^+}^K$ is more compatible with DSS. This might clarify the Δs puzzle.

Summary and outlook

SIA data

- New study of SIA data including recent Belle and BaBar data
- New extraction of fragmentation functions
- The setup for combined pol DIS, SIA and pol SIDIS is ready!

JAMLIB

- JAM SPDFs and FFs are available at github
- Python, Fortran and LHAPDF interfaces provided

JAM for public