Neural network QCD analysis of charged hadron fragmentation functions in the presence of SIDIS data

Based on PHYSICAL REVIEW D 105, 114018 (2022)

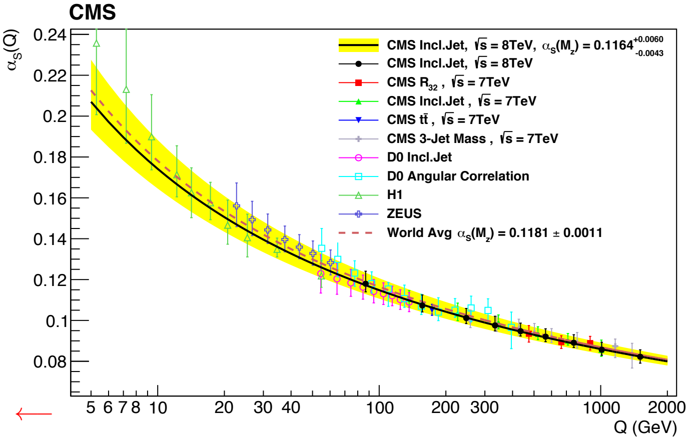
Hadi Hashamipour
h_hashamipour@ipm.ir

in collaboration with M. Soleymaninia and H. Khanpour

School of Particles and Accelerators
Institute for Research in Fundamental Sciences

Wednesday Weekly Seminar of the school of particles and accelerators
29th June 2022

Asymptotic freedom

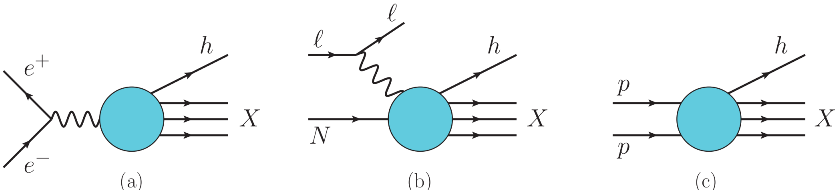


Confinement ⟵

Most strong-interaction processes can not be calculated directly with perturbative QCD, due to color confinement.

Factorization

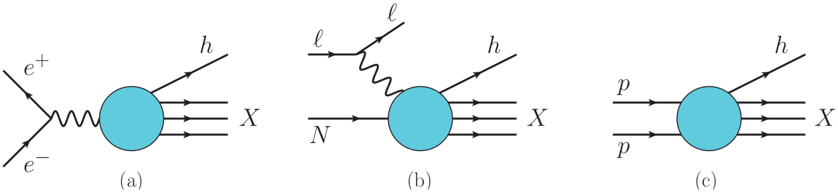
Hard and soft QCD processes contribute? they may be separable.



$$
\begin{aligned}
(a)\ \sigma^{e^+e^-\to hX} &= \hat{\sigma} \otimes FF, \\
(b)\ \sigma^{e^+e^-\to hX} &= \hat{\sigma} \otimes PDF \otimes FF, \\
(c)\ \sigma^{e^+e^-\to hX} &= \hat{\sigma} \otimes PDF \otimes PDF \otimes FF.
\end{aligned}
$$

Factorization

Hard and soft QCD processes contribute? they may be separable.



$$
\begin{aligned}
\text{(a) } \sigma^{\mathrm{SIA}} &= \hat{\sigma} \otimes FF, \\
\text{(b) } \sigma^{\mathrm{SIDIS}} &= \hat{\sigma} \otimes PDF \otimes FF, \\
\text{(c) } \sigma^{e^+e^- \to hX} &= \hat{\sigma} \otimes PDF \otimes PDF \otimes FF.
\end{aligned}
$$

Fragmentation functions in SIA

The differential cross-section for the single-inclusive e^+e^- annihilation giving a hadron,

$$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz} = \frac{1}{\sigma_{\text{tot}}}\left[F_T^h(z,Q^2) + F_L^h(z,Q^2)\right], \tag{1}$$

$$\begin{aligned} F_k^h(z,Q^2) &= \sigma_{\text{tot}}^{(0)}\left[D_S^h(z,Q^2)\otimes C_{k,q}^S(z,\alpha_s(Q)) + D_g^h(z,Q^2)\otimes C_{k,g}^S(z,\alpha_s(Q))\right] \\ &+ \sum_q \sigma_q^{(0)} D_{\text{NS,q}}^h(z,Q^2)\otimes C_{k,q}^{\text{NS}}(z,\alpha_s(Q)). \end{aligned} \tag{2}$$

In this equation, symbol $\otimes$ also denotes the standard convolution integral defined as

$$D(z)\otimes C(z) \equiv \int_z^1 dy D(y)C(z/y). \tag{3}$$

Fragmentation function in SIDIS

The basic cross-section for the charged hadron production in deep inelastic scattering of a lepton from a nucleon can be written as,

$$\frac{d\sigma^h}{dx\,dy\,dz_h} = \frac{2\,\pi\alpha^2}{Q^2}\left[\frac{(1+(1-y)^2)}{y}2\,F_1^h(x,z_h,Q^2) + \frac{2(1-y)}{y}F_L^h(x,z_h,Q^2)\right]. \quad (4)$$

The structure functions F_1^h and F_L^h in Eq. 4 are the relevant inclusive DIS structure functions in which at NLO accuracy are given by,

$$F_1^h(x,z_h,Q^2) = \frac{1}{2}\sum_{q,\bar{q}} e_q^2 \Big\{ q(x,Q^2)D_q^h(z_h,Q^2) + \frac{\alpha_s(Q^2)}{2\pi}\Big[q\otimes C_{qq}^1\otimes D_q^h + q\otimes C_{gq}^1\otimes D_g^h + g\otimes C_{qg}^1\otimes D_q^h\Big](x,z_h,Q^2)\Big\},$$

and a similar relation for F_L.

The convolution symbol $\otimes$ in equations above is defined as,

$$q(x) \otimes C(x, z_h) \otimes D^h(z_h) \equiv \int_x^1 \frac{dx'}{x'} \int_{z_h}^1 \frac{dz'_h}{z'_h} q(\frac{x}{x'}) c(x', z'_h) D(\frac{z_h}{z'_h}) \,.$$

Evolution

The fragmentation functions are not perturbatively calculable, but their scale dependence (scaling violation) is given by the DGLAP equation:

$$\mu \frac{\partial}{\partial \mu} D_i^h(x, \mu) = \sum_j P_{ij}(z, \alpha_s(\mu)) \otimes D_j^h(x/z, \mu) \tag{5}$$

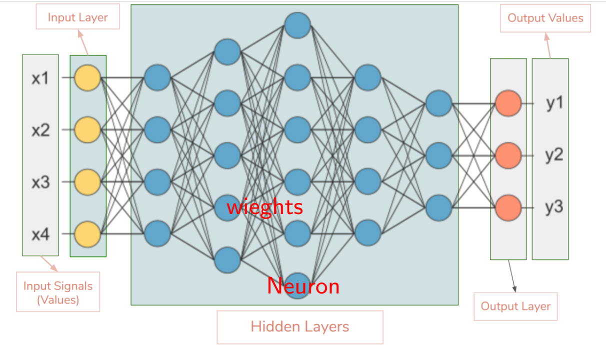
What is Machine learning?

Artificial intelligence (AI) is intelligence demonstrated by machines, as opposed to the natural intelligence displayed by animals including humans.

Machine learning (ML) is a field of inquiry devoted to understanding and building methods that 'learn', It is often seen as a part of artificial intelligence.

An artificial neural network(NN) is a simulation of a biological brain. In other words, they're a way that a machine can process data, that's inspired by human and animal brains.

Artificial Neural Network



$$x_i^{(l)} = \sum_{j=1}^{N_l-1} \omega_{ij}^{(l)} y_j^{(l-1)} + \theta_i^{(l)} \quad (6)$$

the output of each layer is then fed into an activation function to simulate non-linear behavior.

Why NNs?

There are a number of activation functions to use we have used `ReLU`.

Why?

- NNs are flexible, one can fit any function using just simple feed forward NNs.
- They are redundant (by construction). They have many adjustable parameters.
- The above features together result in NN being unbiased.
- NNs have the ability to learn and model non-linear and complex relationships.

Structure of our NNs

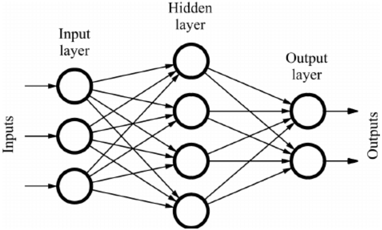


Figure: The feed forward model is the simplest form of neural network as information is only processed in one direction.

- We chose {1-20-7} as our NN architecture.
- Number of adjustable parameters in this configuration is 1*20+20+20*7+7=187, compared to 30-50 in conventional analyses.
- The FF at $Q_0 = 5$ GeV is given as,

$$zD_i^{h^+}(z, Q_0) = \left(N_i(z,\theta) - N_i(1,\theta)\right)^2 . \tag{7}$$

Data
a glimpse

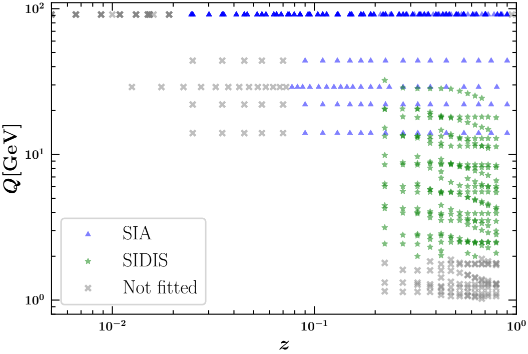


Figure: Kinematic coverage of the data sets analyzed in SHK22.h study.

Data points shown by a cross were excluded from the fit using well established kinematic cuts.

Data set

Dataset facts

- a total of 18 data sets including tagged and untagged SIA and SIDIS for $h^{\pm}$ and h^{+} and h^{-} separately.
- These data sets include the TASSO experiment at DESY; the TPC and SLD experiment at SLAC, ALEPH, DELPHI, and OPAL experiments at CERN.
- number of data points in the fit,
  SIA: 370 data, SIDIS: 314 data, TOTAL: 684
- These data include all of SIDIS and all SIA data sets except for the `TASSO 35` data set, that was left out because of tension with `COMPASS` data sets.
- We checked an analysis including the `TASSO 35` and the results were practically unchanged, except for larger error bands and worse fit quality(larger $\chi^2/N_{\rm dat}$).
- Other TASSO data sets also show a little bit of tension with `COMPASS`, but they were kept in the analysis.

Fitting procedure

Flavor decomposition

- It has been shown that adding the SIDIS datasets to the data sample could provide a direct constraint on the individual q and $\bar{q}$ FFs for light quarks.

$$D_u^{h^+},\ D_{\bar{u}}^{h^+},\ D_{d+s}^{h^+},\ D_{\bar{d}+\bar{s}}^{h^+},\ D_{c^+}^{h^+},\ D_{b^+}^{h^+},\ D_g^{h^+}\ . \quad (8)$$

χ^2 function

We seek to minimize a χ^2 function of form:

$$\chi^{2(k)} \equiv \left(\mathbf{T}(\boldsymbol{\theta}^{(k)}) - \boldsymbol{x}^{(k)}\right)^{\mathrm{T}} \cdot \mathbf{C}^{-1} \cdot \left(\mathbf{T}(\boldsymbol{\theta}^{(k)}) - \boldsymbol{x}^{(k)}\right)\ . \quad (9)$$

Here k is the replica number, we get back to this in unc. propagation.

Minimization

Minimization techniques

- Explicit differentiation
- Genetic algorithm
- Stochastic Gradient Descent methods (SGD)
- Trust-region methods
  implemented in MontBlanc

Uncertainty propagation

Monte Carlo method

- The Monte Carlo approach assumes: the experimental data follow a multivariate Gaussian distribution,

$$\mathcal{G}(\boldsymbol{x}^k) \propto \exp\left(\left(\boldsymbol{x}^k - \boldsymbol{\mu}\right)^T \cdot \boldsymbol{C}^{-1}\left(\boldsymbol{x}^k - \boldsymbol{\mu}\right)\right). \quad (10)$$

where $\boldsymbol{x}^k = \left\{x_1^k, x_2^k, \cdots, x_{N_{\mathrm{dat}}}^k\right\}$ are so called "replicas".

The elements of the covariance matrix are defined as in Ref. [**?**],

$$C_{ij} = \delta_{ij}\sigma_{i,\mathrm{unc}}^2 + \sum_\beta \sigma_{i,\mathrm{corr}}^{(\beta)}\sigma_{j,\mathrm{corr}}^{(\beta)}, \quad (11)$$

Monte Carlo method

Statistical framework

1. The covariance matrix $\boldsymbol{C}$ is decomposed to a lower triangular $\boldsymbol{L}$ using Cholesky method ($\boldsymbol{C} = \boldsymbol{L} \cdot \boldsymbol{L}^T$).
2. Apply this matrix to an N_{dat}-dimensional Gaussian random vector, $\boldsymbol{r}^k$, a vector with the covariance properties of the experimental data is obtained.
3. Accordingly the pseudo dataset is constructed as follows,

$$\boldsymbol{x}^k = \boldsymbol{\mu} + \boldsymbol{L} \cdot \boldsymbol{r}^k \tag{12}$$

Monte Carlo method

Statistical framework

5. for sufficiently large N_{rep}, the set of replicas satisfies the following conditions,

$$\frac{1}{N_{\text{rep}}}\sum_{k}^{N_{\text{rep}}} x_i^k \simeq \mu_i\,, \quad \frac{1}{N_{\text{rep}}}\sum_{k}^{N_{\text{rep}}} x_i^k x_j^k \simeq \mu_i\mu_j + C_{ij}\,. \tag{13}$$

6. Perform the QCD analysis, N_{rep} times, trained NNs are gained,

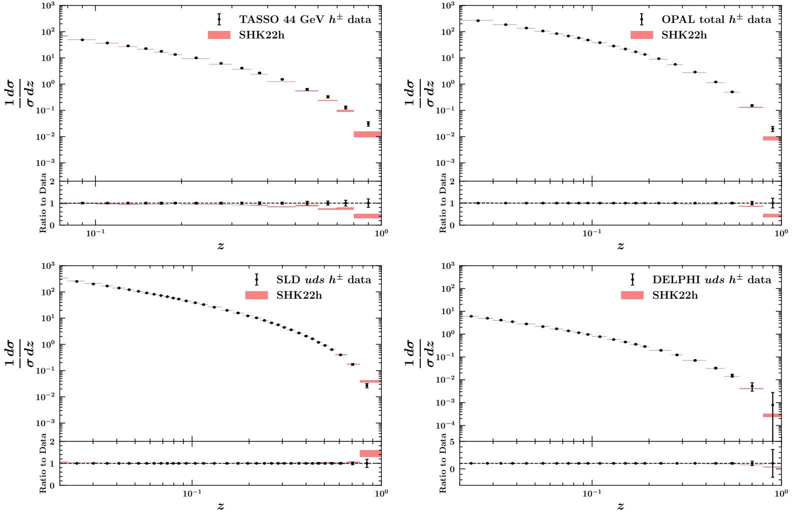
$$\begin{aligned}\langle\mathcal{D}(x)\rangle &= \frac{1}{N_{\text{rep}}}\sum_{k=1}^{N_{\text{rep}}} \mathcal{D}^k(x)\,,\\ \sigma_{\mathcal{D}} &= \sqrt{\langle\mathcal{D}^2\rangle - \langle\mathcal{D}\rangle^2}\,,\\ c_{\mathcal{D}}^{ij} &= \frac{\langle\mathcal{D}_i\mathcal{D}_j\rangle - \langle\mathcal{D}_i\rangle\langle\mathcal{D}_j\rangle}{\sigma_i\sigma_j}\,.\end{aligned} \tag{14}$$

Fit quality

Experiment	Observable	$\chi^2/N_{\rm dat}$	$N_{\rm dat}$
TASSO 14 GeV $h^\pm$	$\frac{1}{\sigma}\frac{d\sigma}{dz}$	1.791	14
TASSO 22 GeV $h^\pm$		1.254	14
TASSO 44 GeV $h^\pm$		2.912	14
TPC $h^\pm$		0.659	21
ALEPH $h^\pm$		0.825	32
DELPHI total $h^\pm$		0.610	21
DELPHI *uds* $h^\pm$		0.380	21
DELPHI bottom $h^\pm$		1.028	21
OPAL total $h^\pm$		1.821	19
OPAL *uds* $h^\pm$		0.794	19
OPAL charm $h^\pm$		0.599	19
OPAL bottom $h^\pm$		0.299	19
SLD total $h^\pm$		1.047	34
SLD *uds* $h^\pm$		0.946	34
SLD charm $h^\pm$		1.034	34
SLD bottom $h^\pm$		1.102	34
COMPASS h^-	$\frac{dM^h}{dz}$	0.907	157
COMPASS h^+		1.338	157
Global dataset		**1.079**	**684**

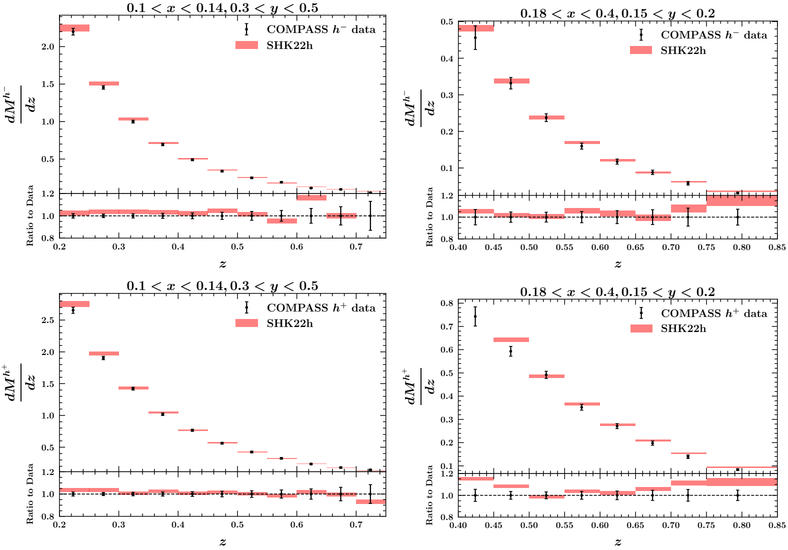
The χ^2 per data point after the cut.

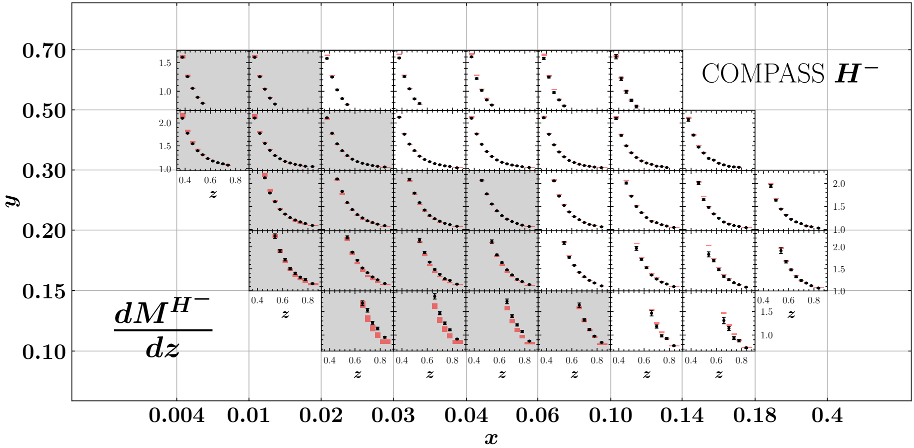
Fit quality - SIA experiments



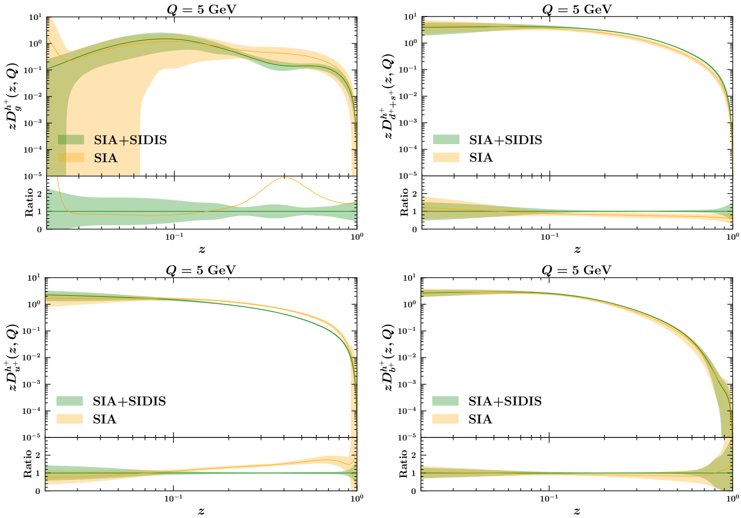
From top left very high to right bottom very good χ^2.

Fit quality - COMPASS - SIDIS

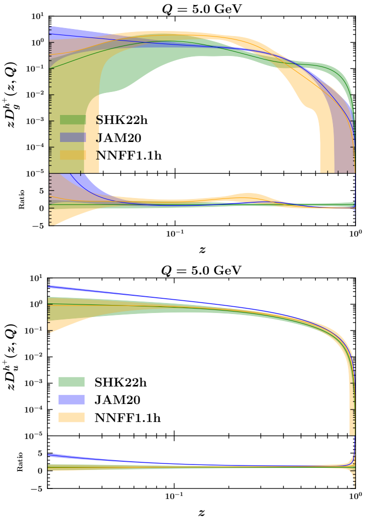



COMPASS H^-
$\frac{dM^{H^-}}{dz}$
y
0.70
0.50
0.30
0.20
0.15
0.10
x
0.004
0.01
0.02
0.03
0.04
0.06
0.10
0.14
0.18
0.4
z
0.4
0.6
0.8
1.0
1.5
2.0

Effect of SIDIS data

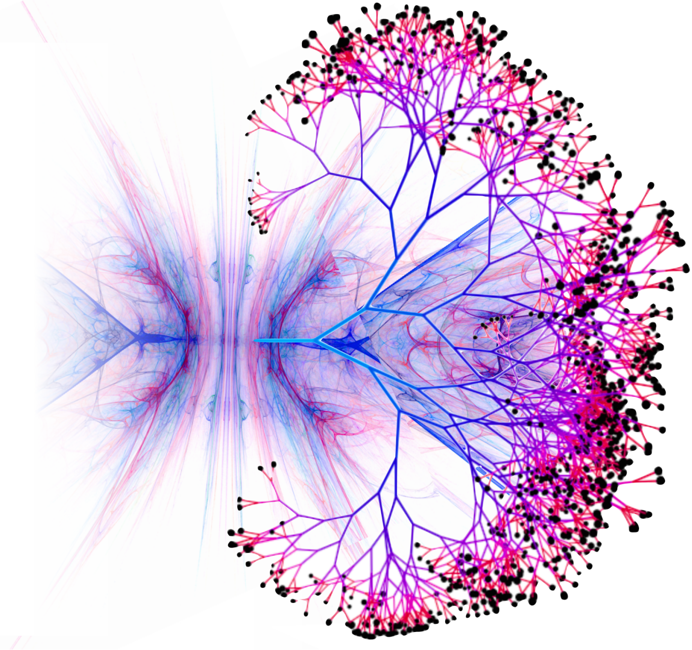


Comparison with other groups



Summary

- In summary, we have presented a new global QCD analysis of light-charged hadron FFs, `SHK22.h`, with several new features and some methodological improvements.
- We have used the machine learning framework to extract the `SHK22.h` FFs sets, along with the Monte Carlo uncertainty analysis.
- In addition to the comprehensive set of high-energy lepton-lepton annihilation (SIA), we have added the lepton-hadron scattering (SIDIS) data sets.
- We have shown that SIDIS data sets have significant effect on the FFs, and more specifically on the gluon FFs and the reduction of its uncertainty.
- The detailed comparisons to the existing FFs sets (NNFF1.1h and JAM20) demonstrate a reasonable agreement.
- Overall, the resulting NLO theory predictions for the SIA and SIDIS cross sections show very good agreement with the corresponding analyzed experimental datasets, as confirmed by the reported total χ^2 per data point.



Thank You!