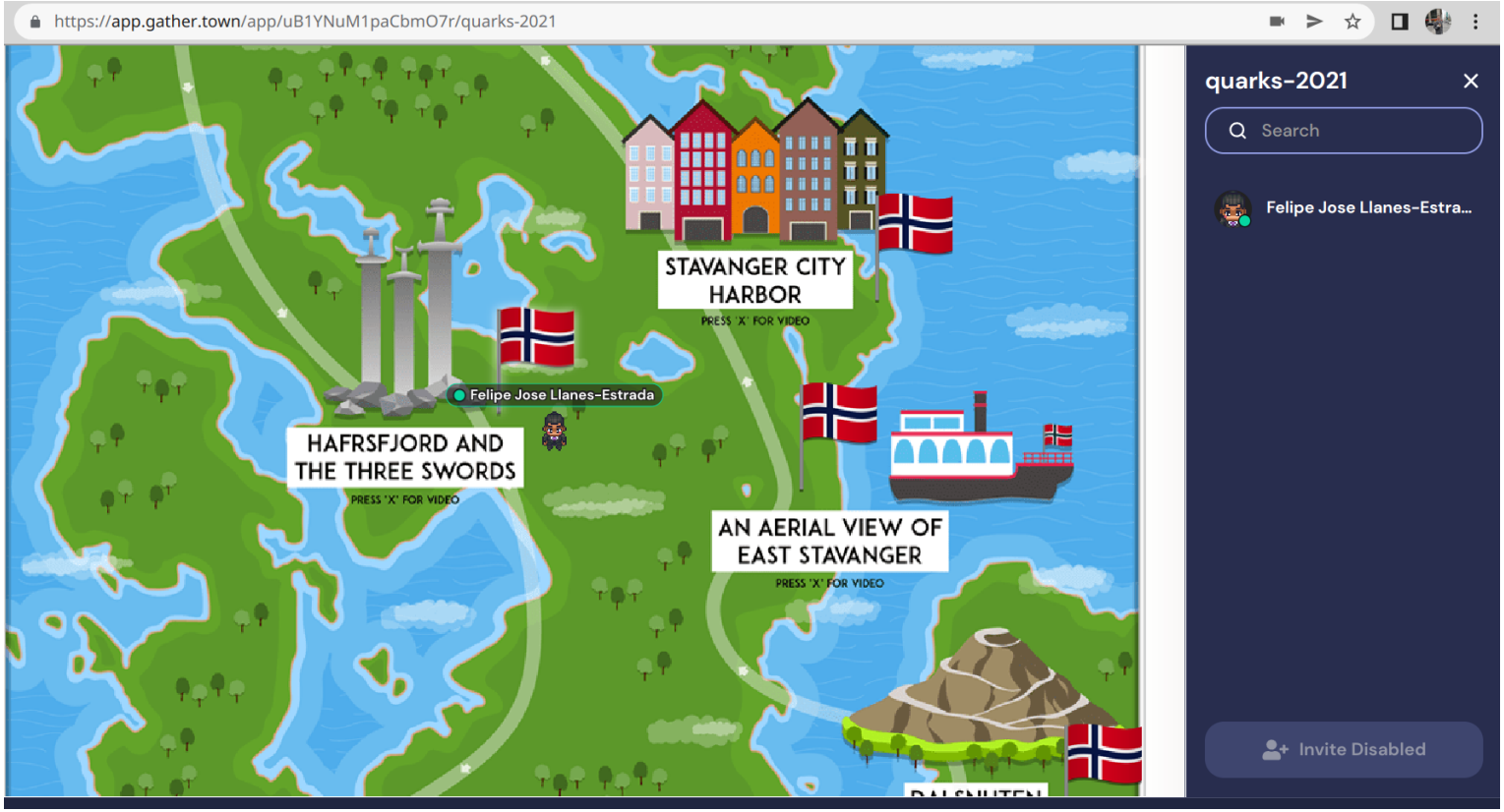
Information entropy and fragmentation functions

Felipe J. Llanes-Estrada
Guillermo Benito Calviño & Javier García Olivares

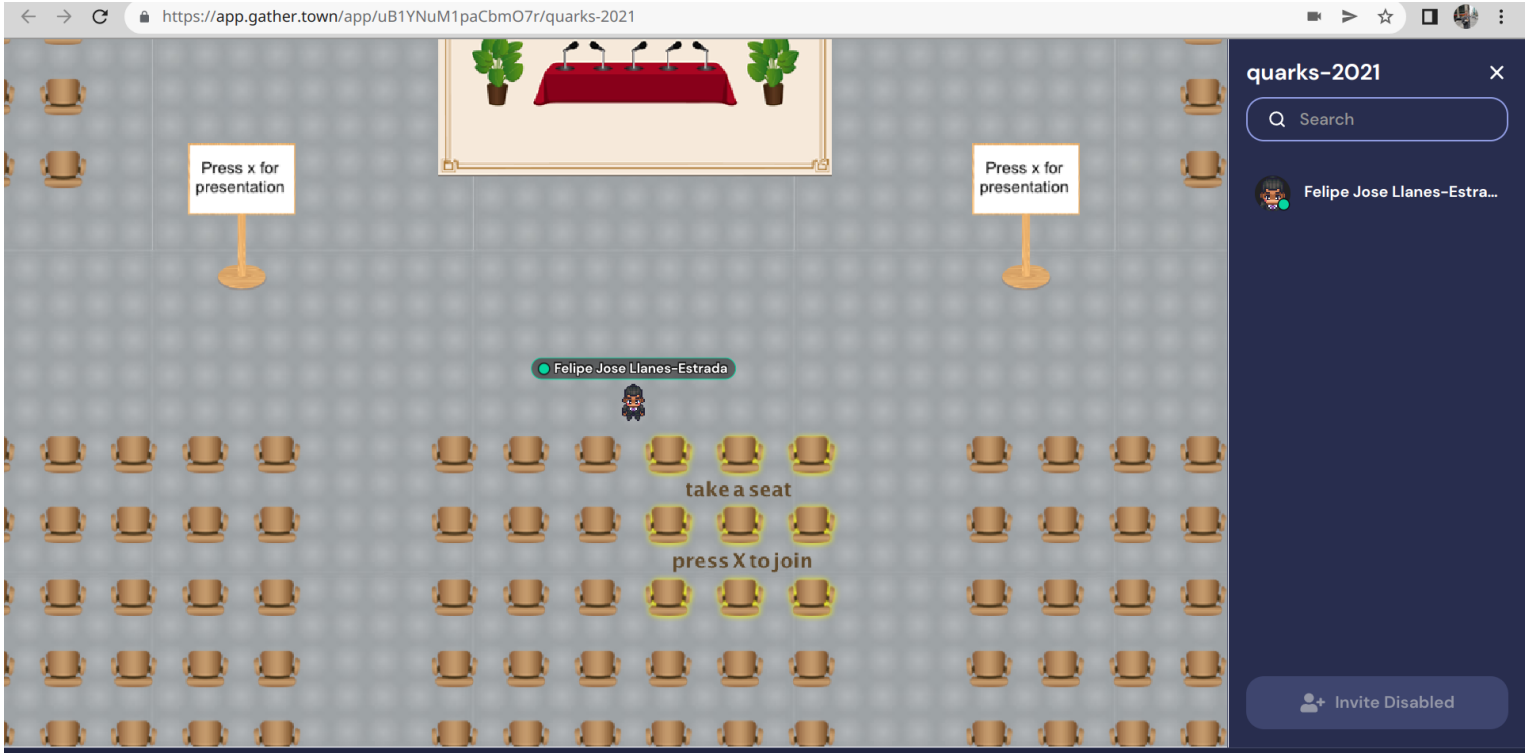
Universidad Complutense de Madrid, Departamento de Física Teórica and IPARCOS

1/Aug/2022 XV Quark Confinement & Hadron Spectrum

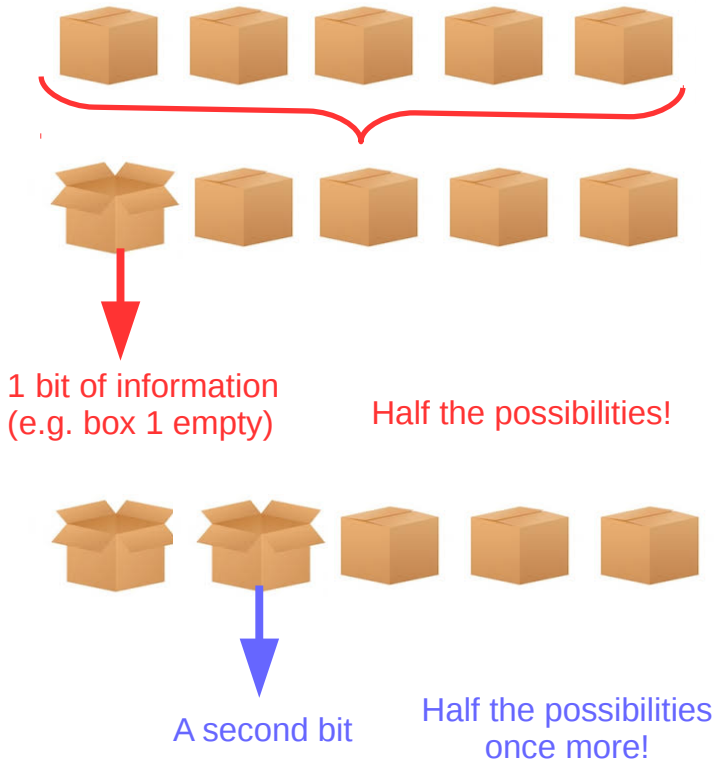
Remember the meeting in Gathertown?



Remember the meeting in Gathertown?



Information entropy



Probability of a configuration in terms of the number of bits of information obtained:

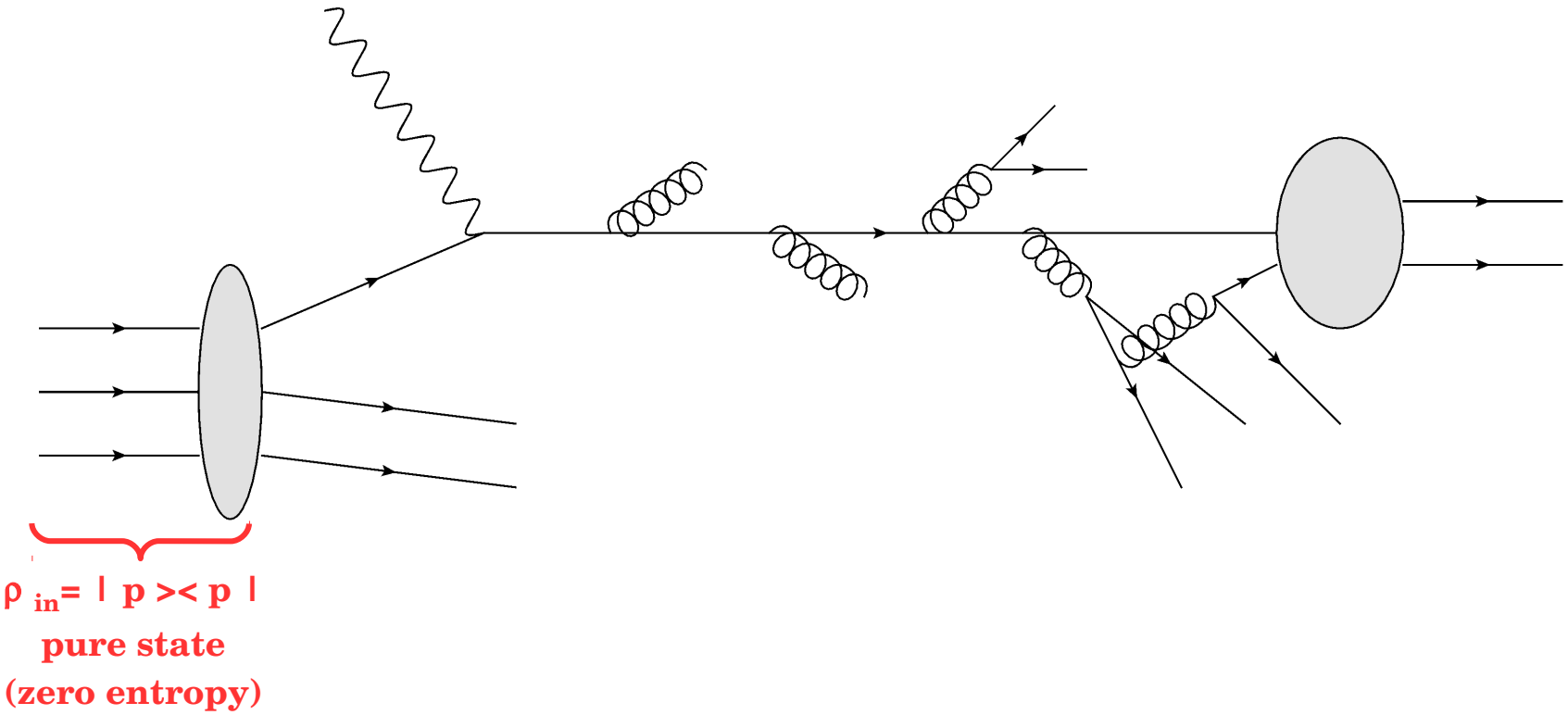
$$p = \left(\frac{1}{2}\right)^I \implies I = -\log_2(p)$$

Shannon entropy: expectation value of the information in a probability distribution

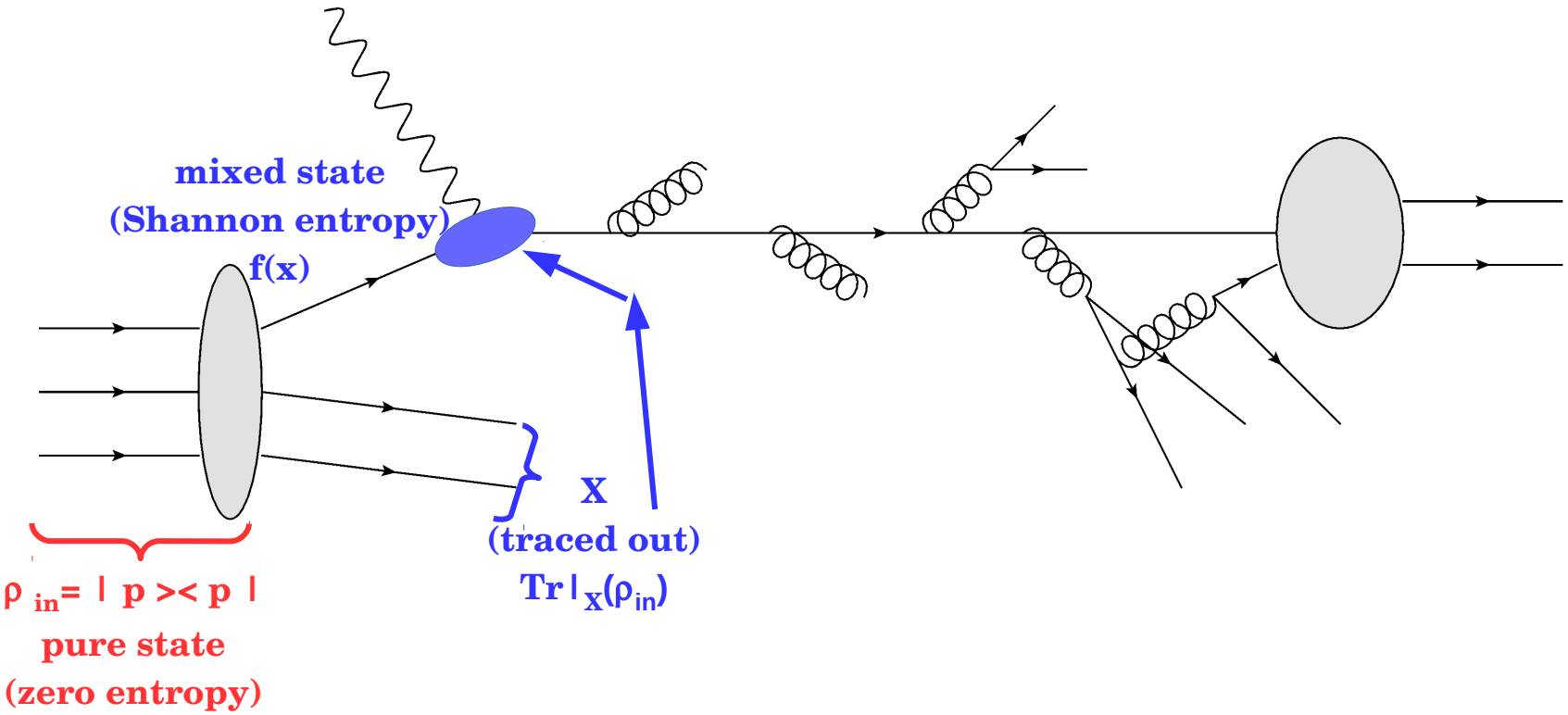
$$\boxed{S = -\sum_i p_i \log p_i}$$

Quantum Von Neumann entropy $S = -\mathrm{Tr}(\rho \log \rho)$

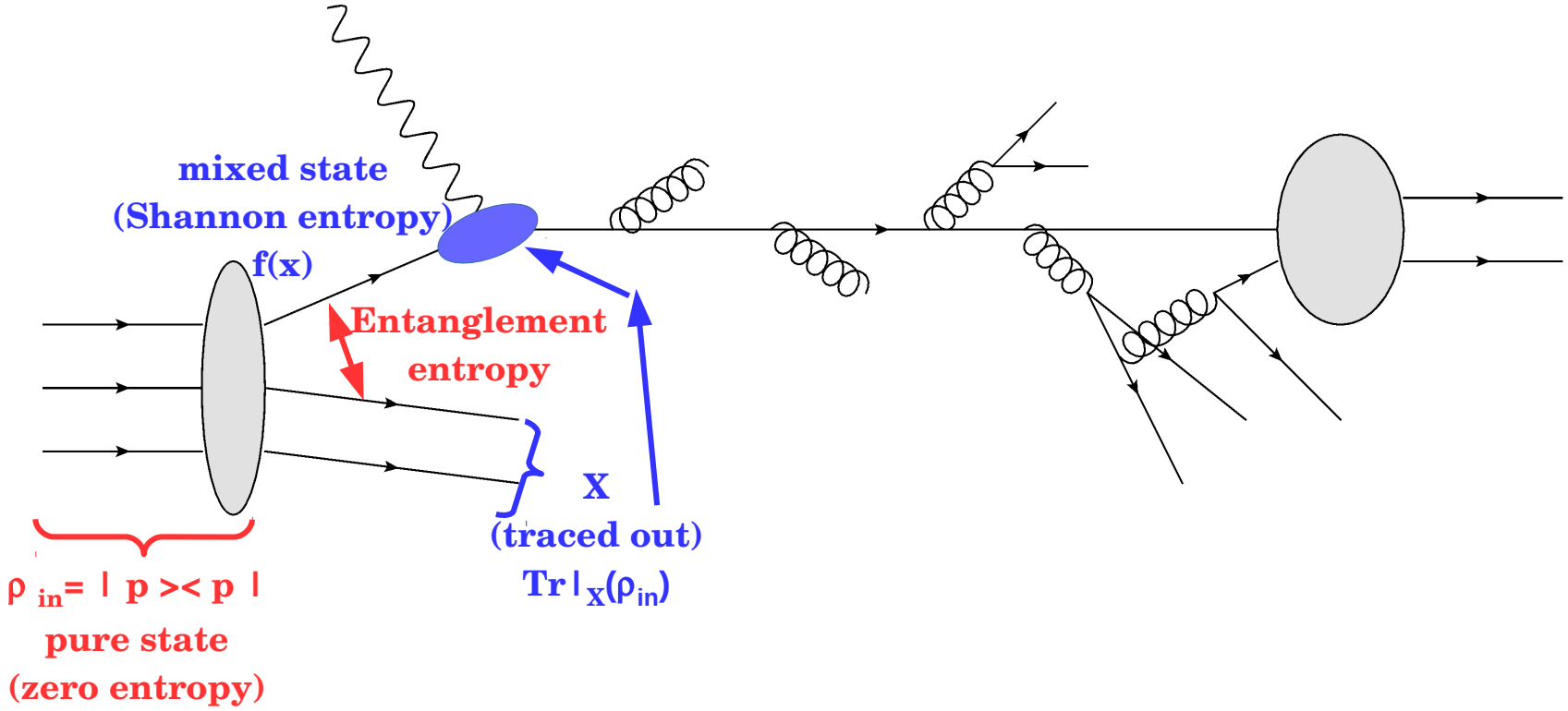
Entropy in high-energy physics



Entropy in high-energy physics

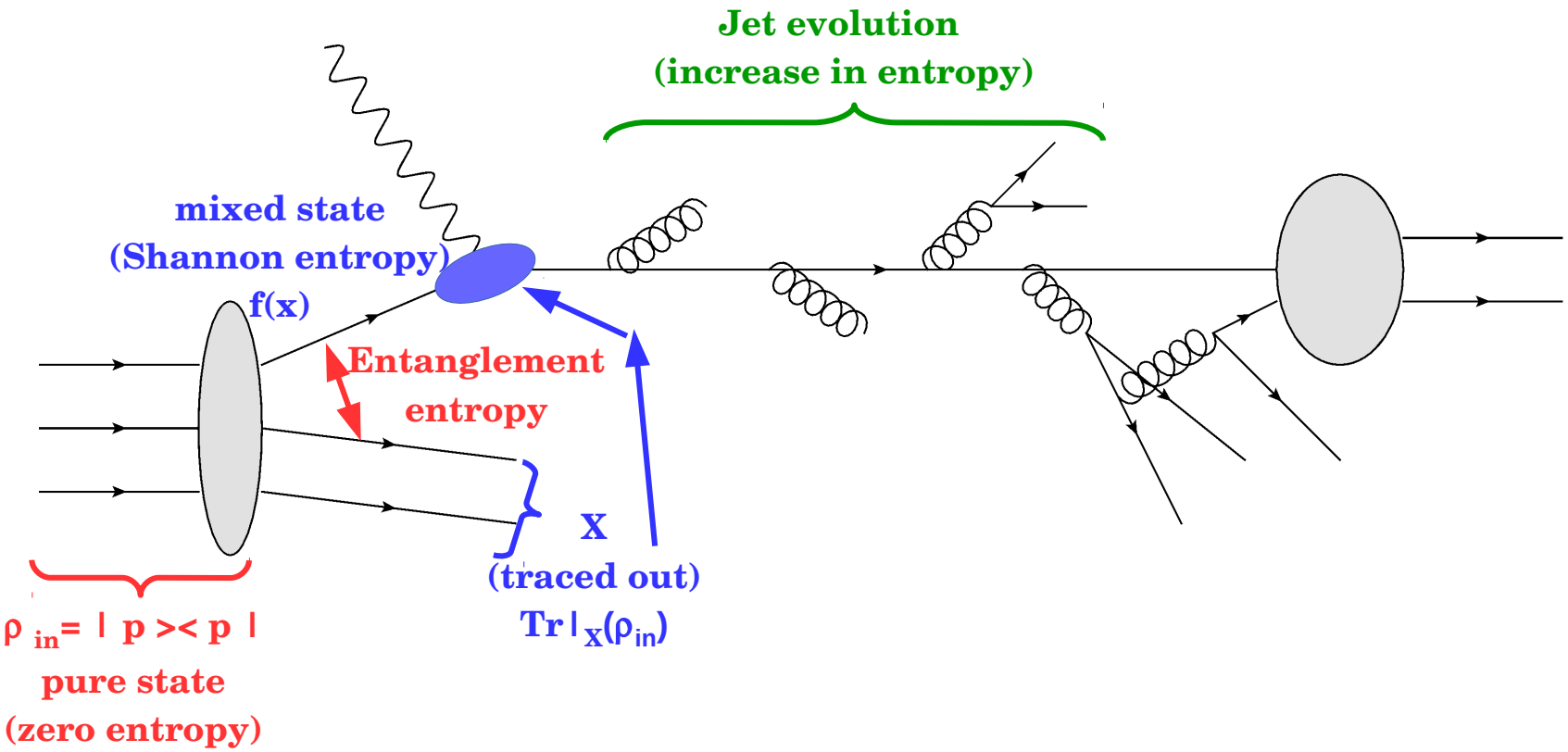


Entropy in high-energy physics



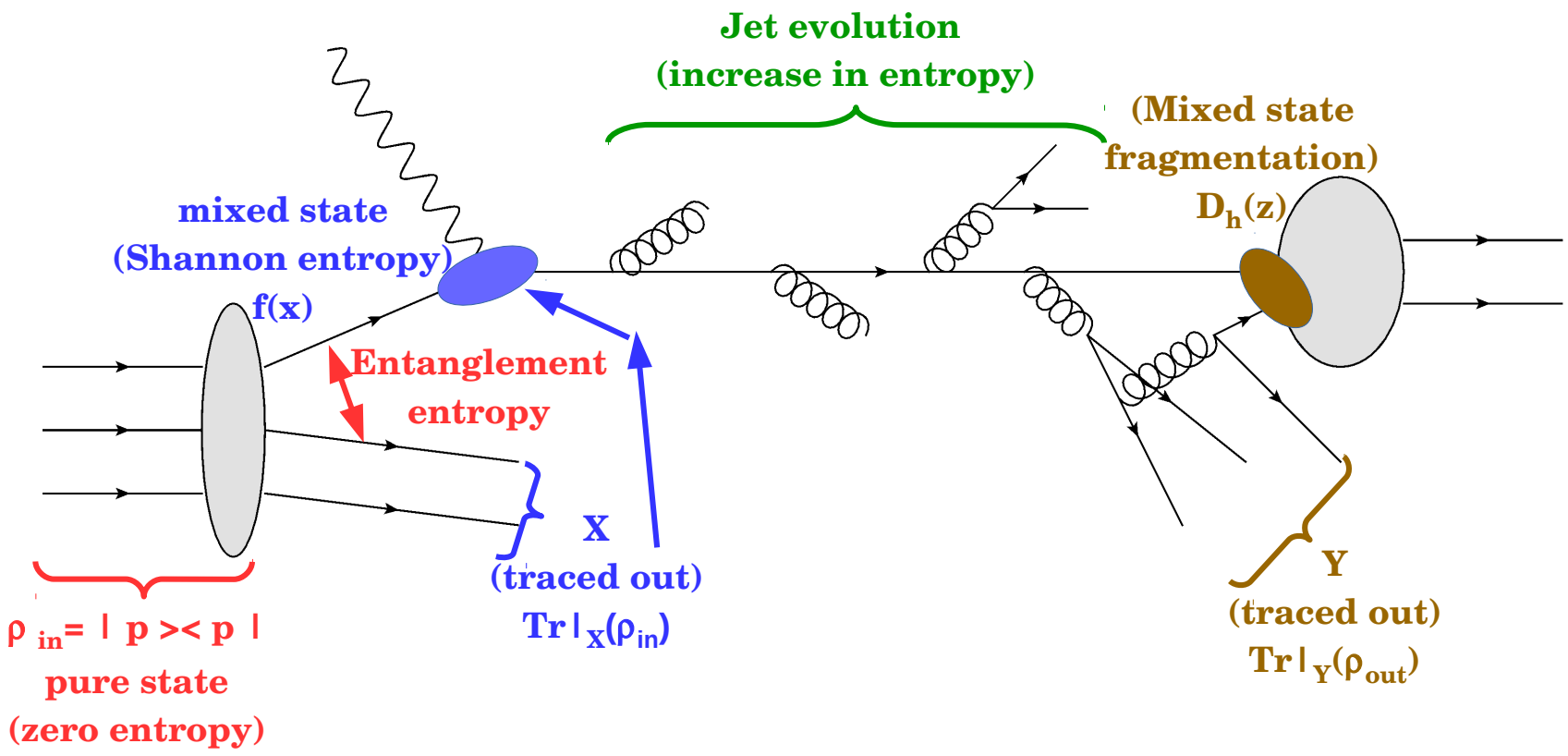
Kharzeev and Levine, Phys.Rev.D 104 (2021) 3, L031503

Entropy in high-energy physics

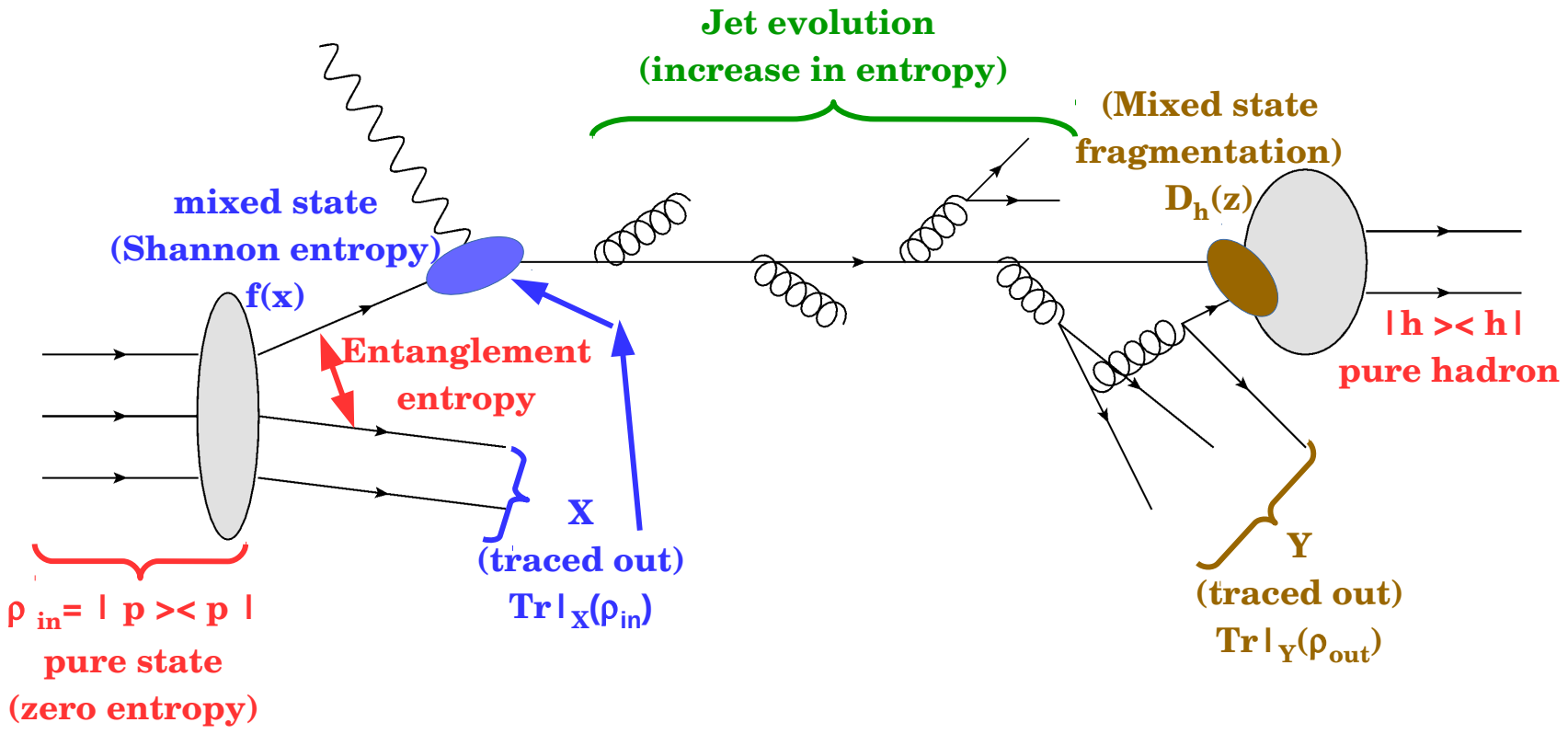


Neill and Waalewijn, Phys.Rev.Lett. 123 (2019) 14, 142001

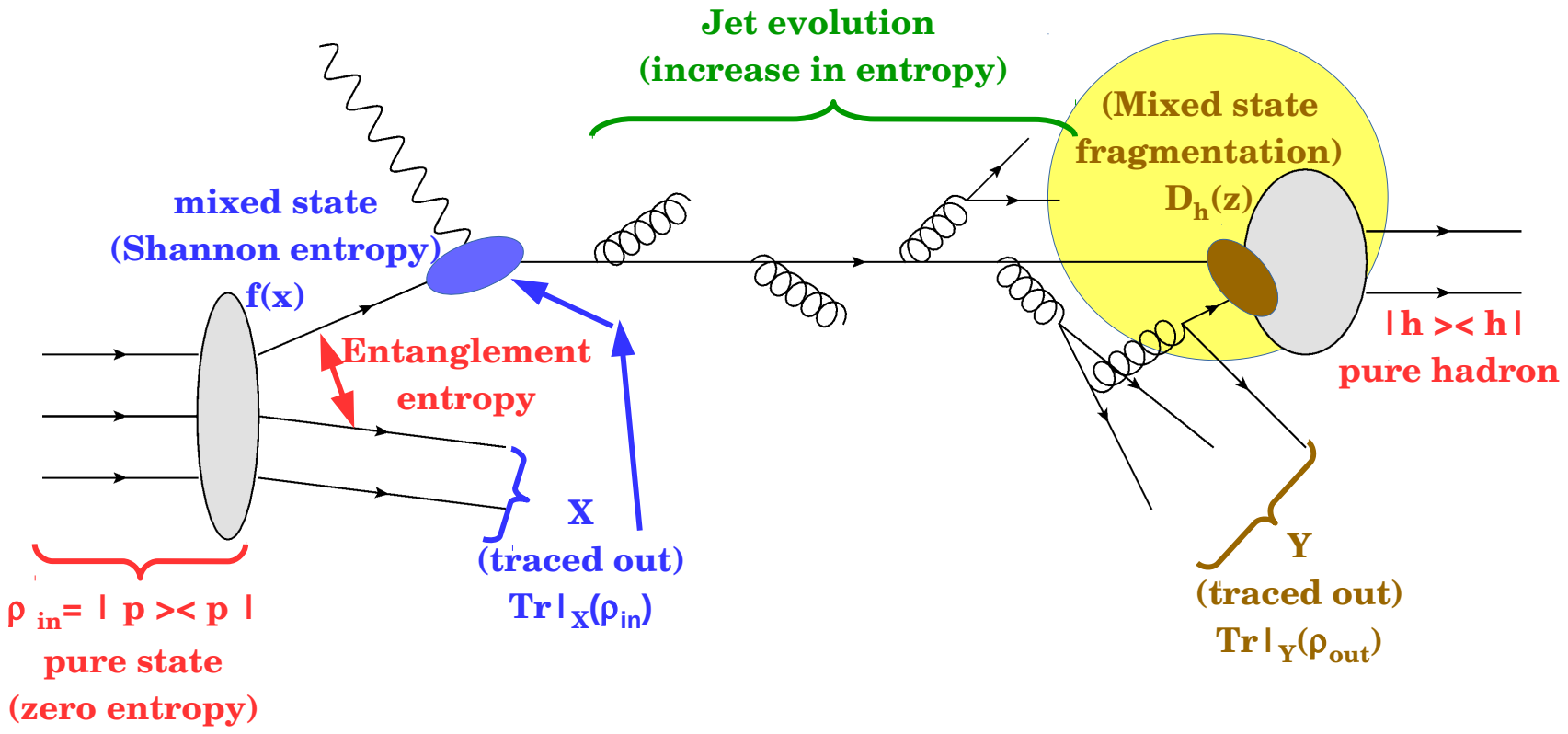
Entropy in high-energy physics



Entropy in high-energy physics



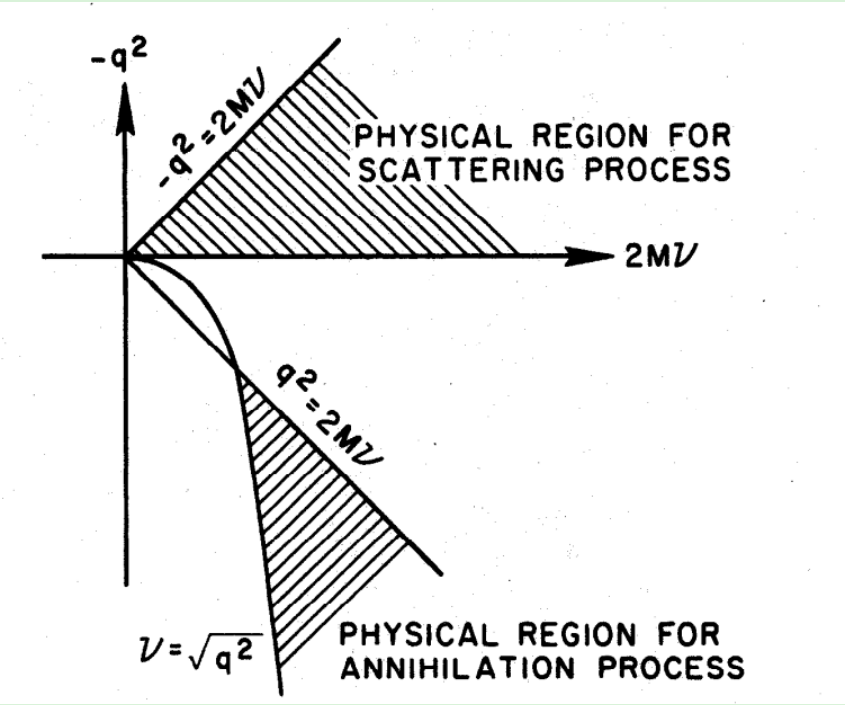
This work: entropy in fragmentation functions



Pdfs vs FFs

Parton distributions	Fragmentation functions
$\frac{d\sigma(e^- h \to e^- X)}{dx dy_1} = \frac{2\pi\alpha^2}{Q^4} s[1 + (1 + y_1)^2 F_2(x)]$	$\frac{d\sigma(e^- e^+ \to hX)}{dz} = \frac{4\pi\alpha^2}{3Q^2} z^2 \hat{F}_2(z)$
$F_2(x) = \sum_i e_i^2 x[f_i(x) + \bar{f}_i(x)]$	$\hat{F}_2(z) = -3 \sum_q e_q^2 \frac{D^q(z) + D^{\bar{q}}(z)}{z^2}$
$\sum_i \int x f_i(x) dx = 1$	$\sum_h \int_0^1 z D_h^q(z) dz = 1$

DIS and fragmentation in e^-e^+ related by continuation



Drell, Levy, Yan, Phys. Rev. D 6, 1617 (1970)

Proposed relations between pdfs and FFs

- Drell-Levy-Yan relation $D(z) = z\, f\left(\frac{1}{z}\right)$

  Problem: if $z \in [0,1]$, $\frac{1}{z} \in [1,\infty)$
  (unphysical except at $z \to 1$)

- Gribov-Lipatov "reciprocity" relation $3D(z) = f(z)$

  theoretical status: rather a model

- For $z > 0.6$ Barone-Drago-Ma eq. $D(z) \simeq zf\left(2 - \frac{1}{z}\right)$

  Approximate save $z \to 1$; but both have physical argument

Proposed relations between pdfs and FFs

- Drell-Levy-Yan relation $D(z) = z\, f\left(\frac{1}{z}\right)$

  Problem: if $z \in [0,1]$, $\frac{1}{z} \in [1,\infty)$
  (unphysical except at $z \to 1$)

- Gribov-Lipatov "reciprocity" relation $3D(z) = f(z)$

  theoretical status: rather a model

- For $z > 0.6$ Barone-Drago-Ma eq. $D(z) \simeq z f\left(2 - \frac{1}{z}\right)$

  Approximate save $z \to 1$; but both have physical argument

Proposed relations between pdfs and FFs

- Drell-Levy-Yan relation $D(z) = z\, f\left(\frac{1}{z}\right)$

  Problem: if $z \in [0,1]$, $\frac{1}{z} \in [1,\infty)$
  (unphysical except at $z \to 1$)

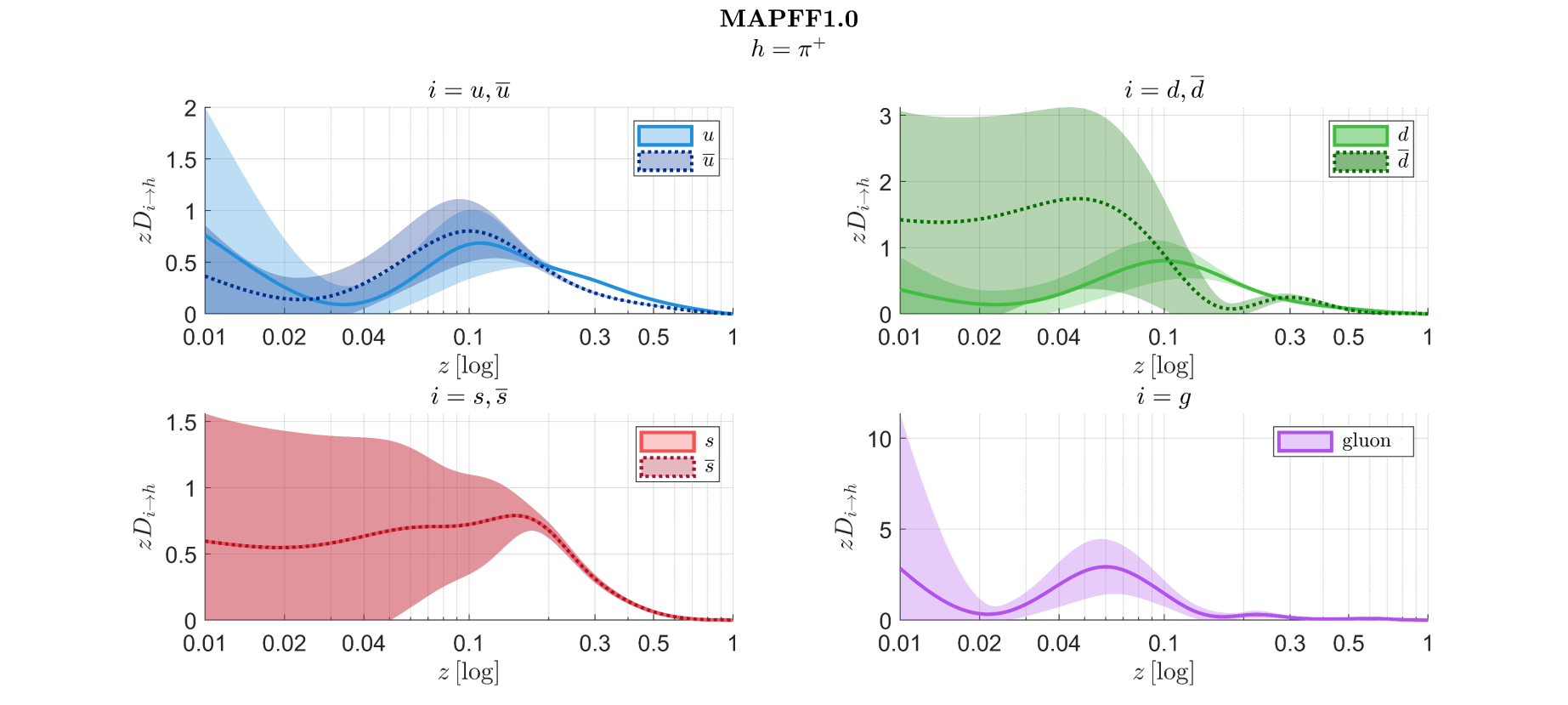
- Gribov-Lipatov "reciprocity" relation $3D(z) = f(z)$

  theoretical status: rather a model

- For $z > 0.6$ Barone-Drago-Ma eq. $D(z) \simeq zf\left(2 - \frac{1}{z}\right)$

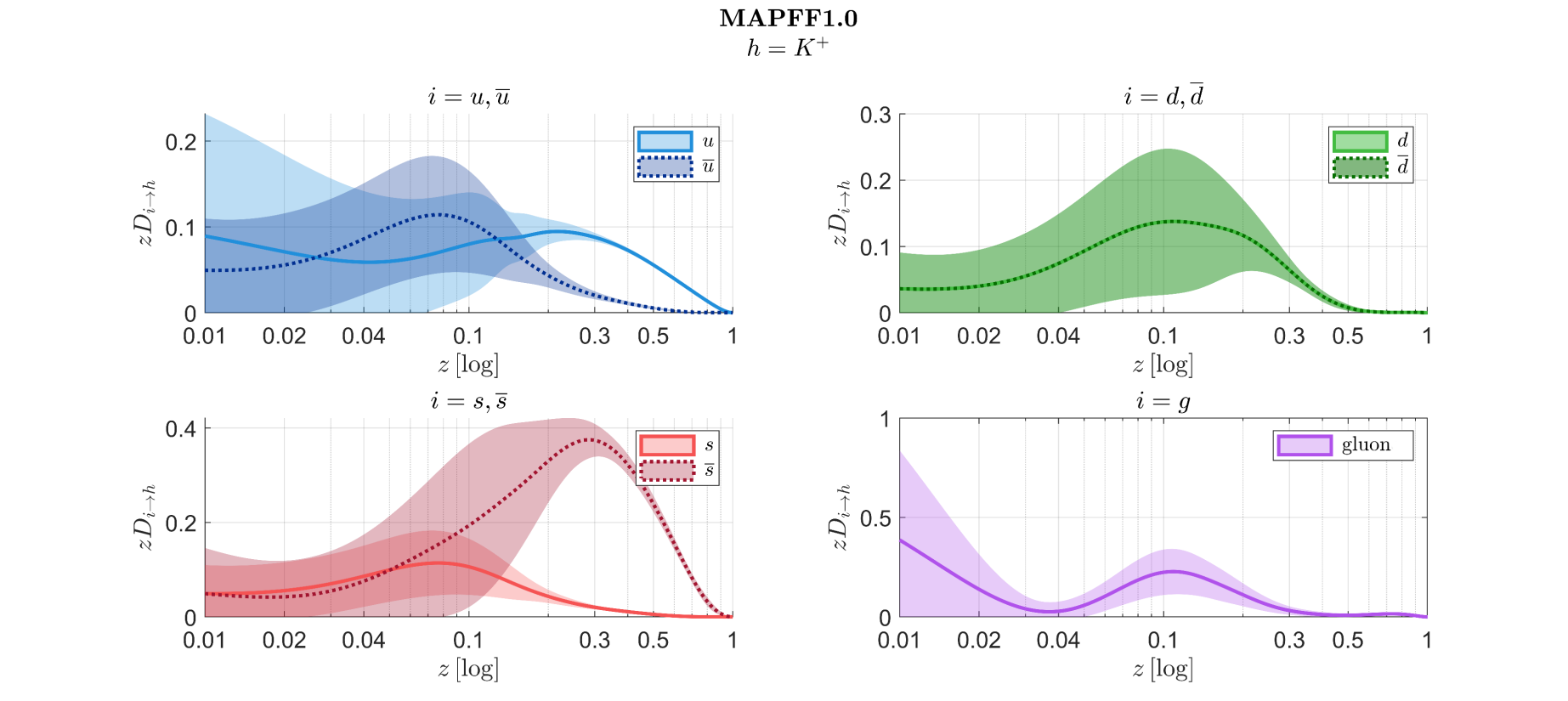
  Approximate save $z \to 1$; but both have physical argument

FFs from neural network fits in the literature



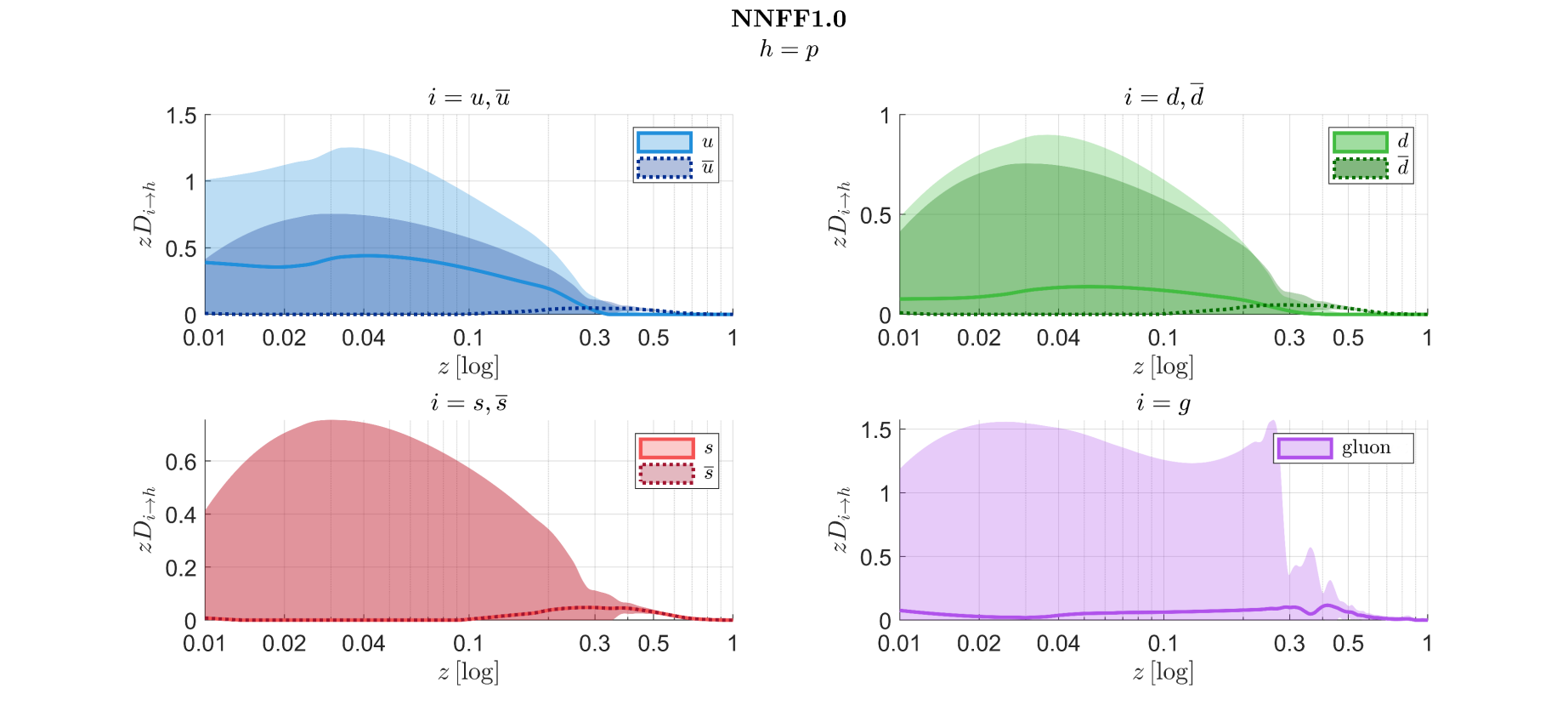
NNPDF coll. EPJC 77, 516 (2017)

FFs from neural network fits in the literature



NNPDF coll. EPJC 77, 516 (2017)

FFs from neural network fits in the literature



NNPDF coll. EPJC 77, 516 (2017)

Entropy for a continuous probability distribution

Simple discretization fails:

$zD(z) \to \{p_1, p_2 \dots p_i \dots p_N\}; \qquad \sum_i p_i = 1$

$S = -\sum_{i=1}^{N} (p_i \log p_i) \xrightarrow{N \to \infty} \log N \to \infty$

Generalization to a continuum distribution:

$S(F) := -\int f(x) \ln f(x) \mathrm{d}x$

Entropy for a continuous probability distribution

Simple discretization fails:

$zD(z) \to \{p_1, p_2 \dots p_i \dots p_N\}; \qquad \sum_i p_i = 1$

$S = -\sum_{i=1}^{N} (p_i \log p_i) \xrightarrow{N\to\infty} \log N \to \infty$

Generalization to a continuum distribution:

$S(F) := -\int f(x) \ln f(x) \mathrm{d}x$

Entropy for a continuous probability distribution

Simple discretization fails:

$zD(z) \to \{p_1, p_2 \dots p_i \dots p_N\}; \qquad \sum_i p_i = 1$

$S = -\sum_{i=1}^{N} (p_i \log p_i) \xrightarrow{N\to\infty} \log N \to \infty$

Generalization to a continuum distribution:

$S(F) := -\int f(x) \ln f(x) \mathrm{d}x$

Kullback-Leibler divergence

$$D_{\mathrm{KL}}(P\|Q) := \sum_x p(x)\log\frac{p(x)}{q(x)} \to$$

How different are the distributions P and Q

Continuum generalization:

$$D_{\mathrm{KL}}(F\|G) = \int f(x)\log\frac{f(x)}{g(x)}\mathrm{d}x$$

$S(F) := -\int f(x)\ln f(x)\mathrm{d}x$ is then the KL divergence to the uniform distribution (that with least information)

What to do with insufficient data?

$$\underbrace{\sum_h}_{\text{Need them all!}} \int_0^1 z D_h^q(z) dz = 1$$

In practice, have $\pi^\pm$, $K^\pm$, p and little more...

Split: $D_q(z) = D_q^{\text{measured}}(z) + D_q^{\text{unknown}}(z)$

$$\underbrace{\sum_h}_{\text{measured}} \int_0^1 z D_h^q(z) dz = p < 1 \qquad \underbrace{\sum_h}_{\text{unknown}} z D_h^q(z) := 1 - p$$

(maximize entropy: uniform distribution remainder)

What to do with insufficient data?

$$\underbrace{\sum_h}_{\text{Need them all!}} \int_0^1 z D_h^q(z) dz = 1$$

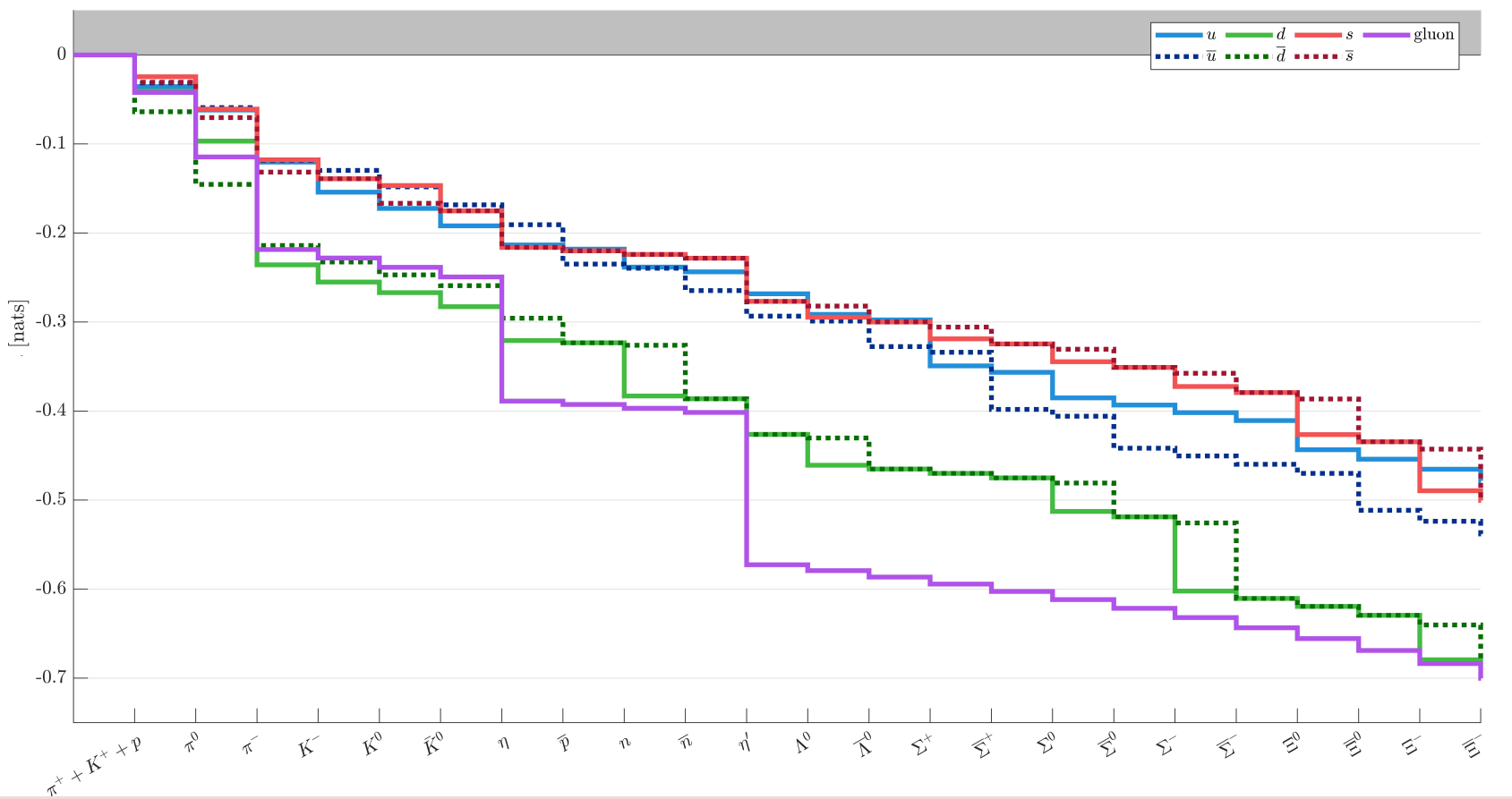
In practice, have $\pi^\pm$, $K^\pm$, p and little more...

Split: $D_q(z) = D_q^{\text{measured}}(z) + D_q^{\text{unknown}}(z)$

$$\underbrace{\sum_h}_{\text{measured}} \int_0^1 z D_h^q(z) dz = p < 1 \qquad \underbrace{\sum_h}_{\text{unknown}} z D_h^q(z) := 1 - p$$

(maximize entropy: uniform distribution remainder)

What would happen with future data? Upper bound on S



What would happened with future data?

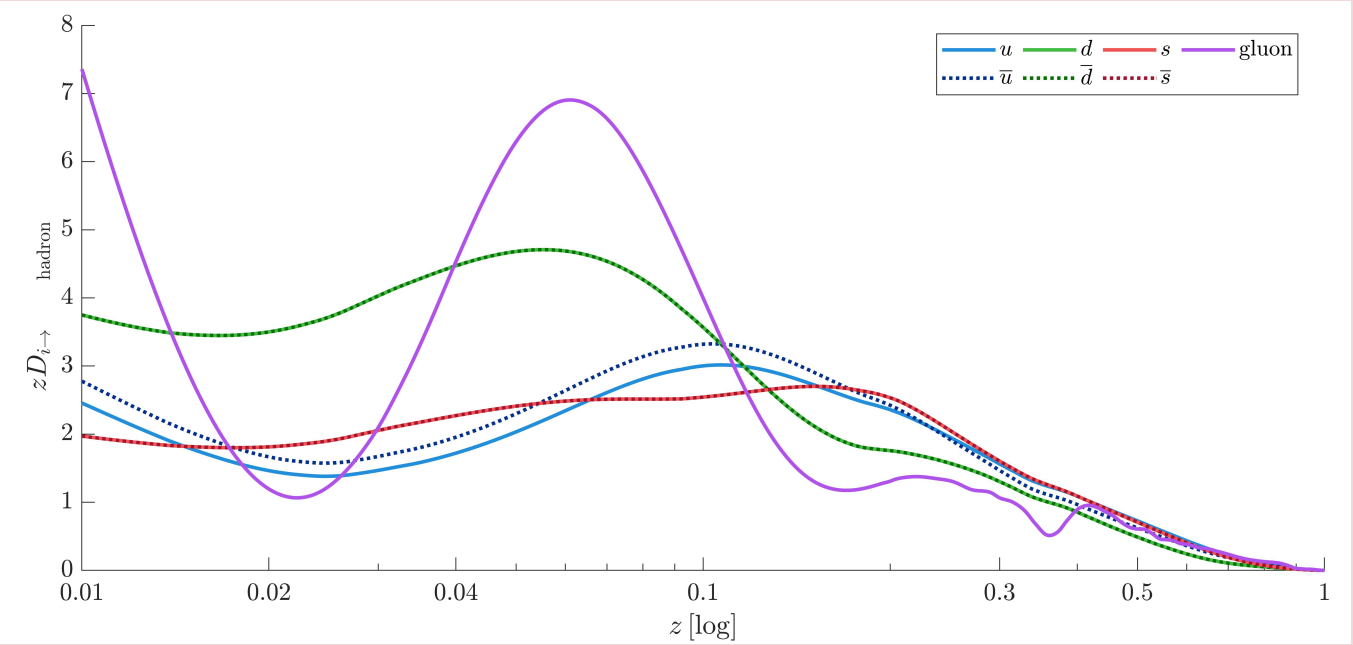
- From $SU(3)$ relations + valence/sea separation as customary *e.g.* $D_{d\to K^0}$ (unknown) set $=$ $D_{u\to K^+}$ (already extracted)
- $\theta(\eta,\eta') = -15.5^{\circ}$
- S: one number that quantifies progress on Fragmentation Functions

What would happened with future data?

- From $SU(3)$ relations + valence/sea separation as customary *e.g.* $D_{d\to K^0}$ (unknown) set $=$ $D_{u\to K^+}$ (already extracted)
- $\theta(\eta,\eta') = -15.5^{\rm o}$
- S: one number that quantifies progress on Fragmentation Functions

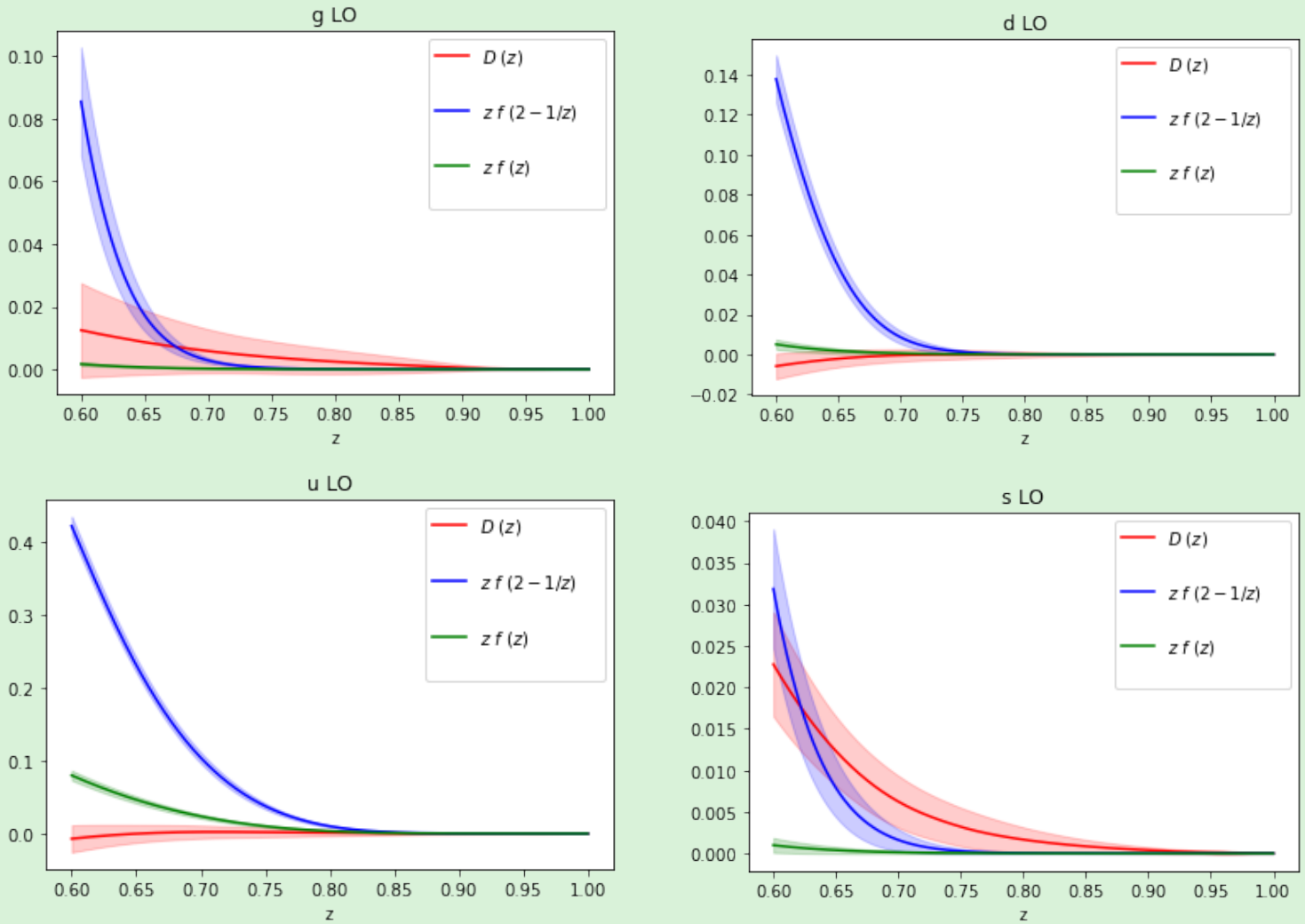
What would happened with future data?

For completeness, the looks of the "reconstructed" FFs:



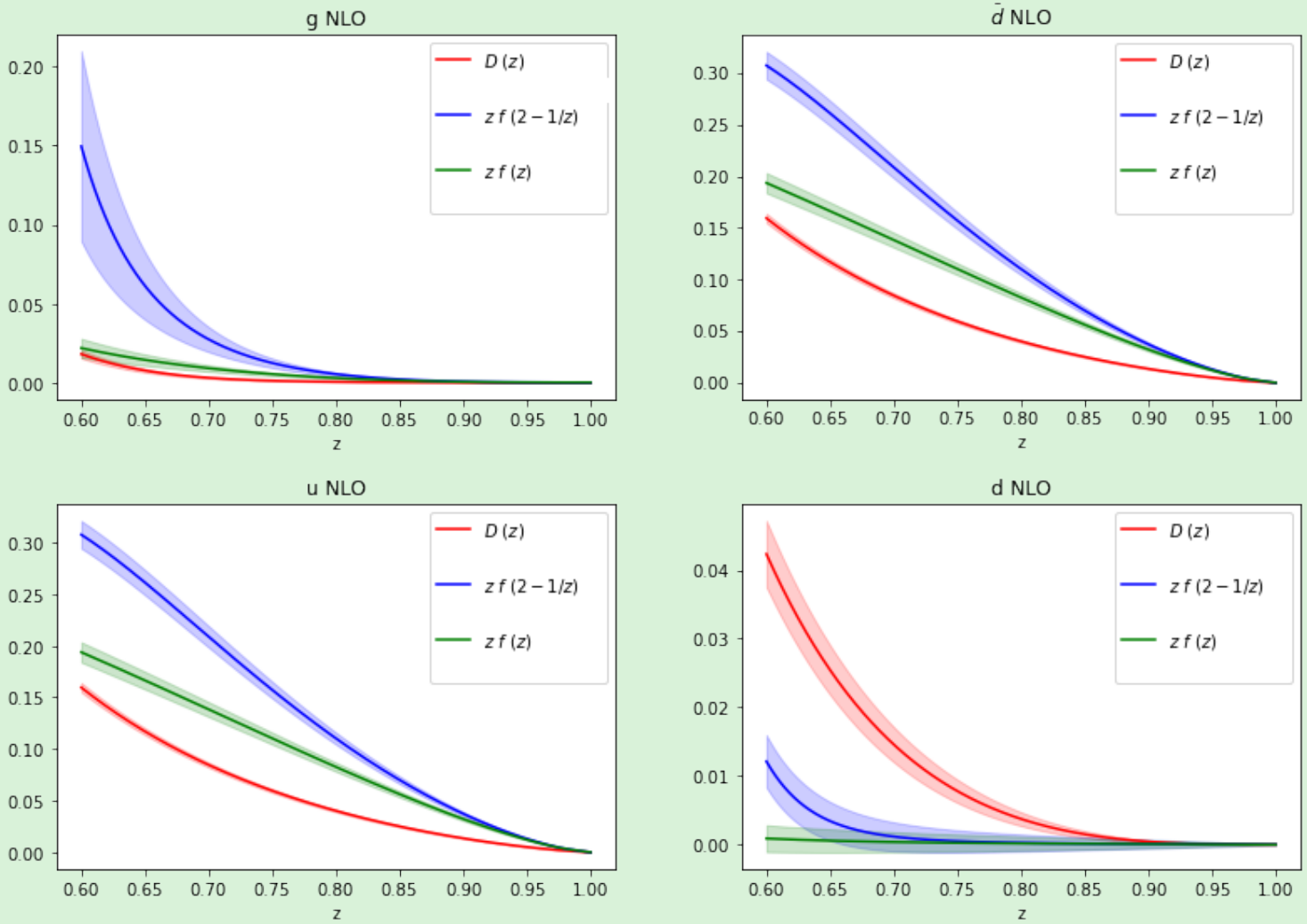
Click here for animation

Comparison between *D*s and *FF*s



(Using the proton's NNPDF parametrizations, $Q^2 = 100\mathrm{GeV}^2$)

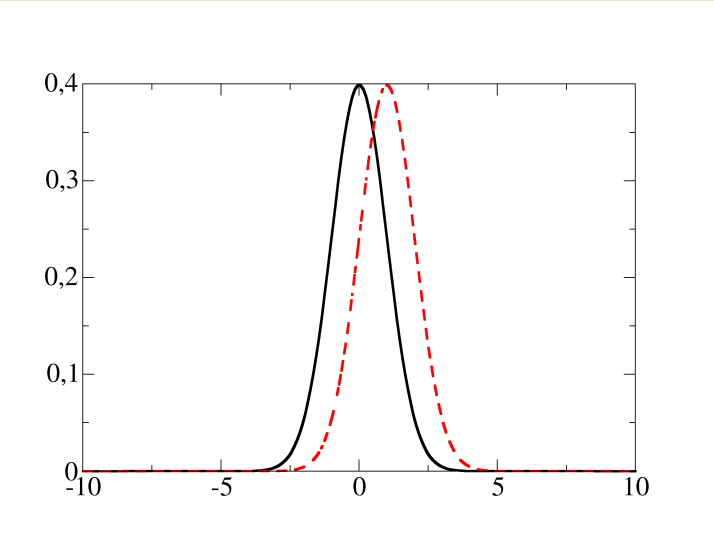
Comparison between *D*s and *FF*s



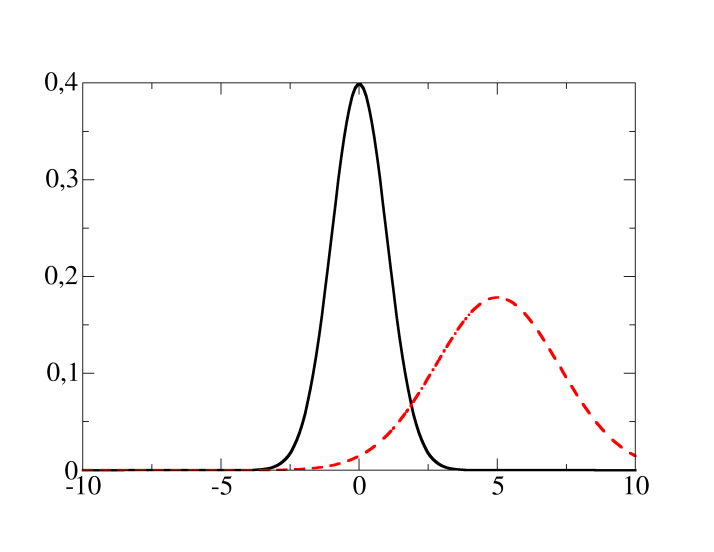
(Traditional template fit to π^+ data, $Q^2 = 100\mathrm{GeV}^2$)

Jefferson Lab Angular Momentum Collaboration, Phys.Rev.D 104 (2021) 1, 016015

Kullback-Leibler divergence: typical cases



$$D_{\mathrm{KL}} = 0.12$$



$$D_{\mathrm{KL}} = 1.63$$

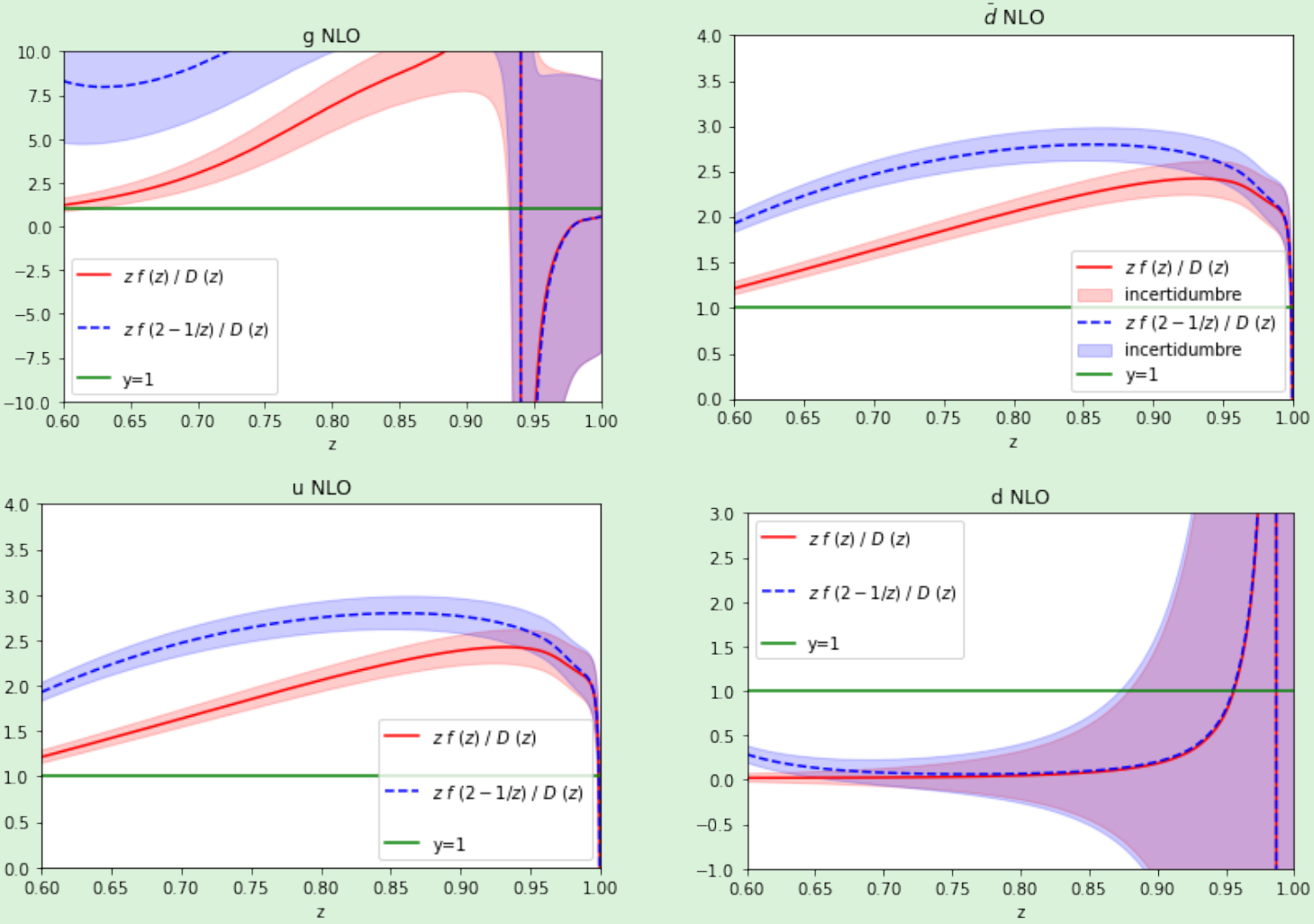
Kullback-Leibler divergence for two relations

Parton	$D_{KL}(zf_\pi(z)\|\|D^\pi(z))$	$D_{KL}(zf_\pi(2-1/z)\|\|D^\pi(z))$
gluon	2,28	7,70
up	1,19	2,36
anti-down	1,19	2,36
down	6,64	5,28

(contribution from $z \in (0.7, 1)$)

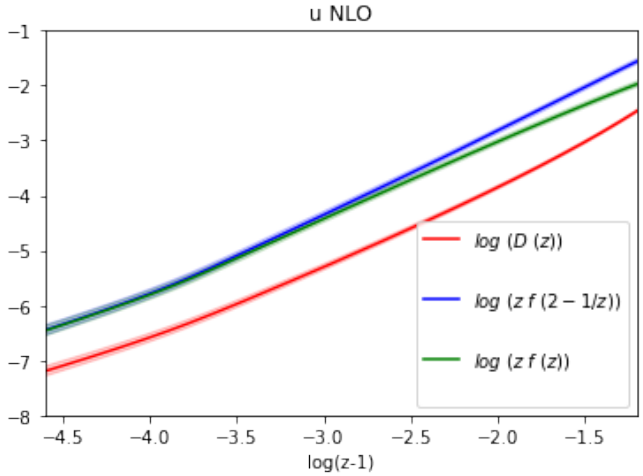
Not well satisfied!

Ratios between *D*s and *FF*s: not 1!

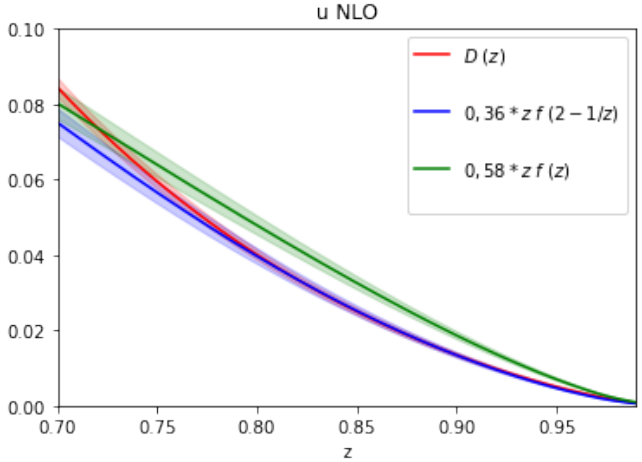


(NLO parametrizations of JAM coll. π^+ data, $Q^2 = 100\mathrm{GeV}^2$)

Relations between $f_\pi^u(x)$ and $D_u^\pi(z)$



Parallel lines
in a log-log plot



Barone-Drago-Ma relation off by a multiplicative constant only!

Further approximating $f(2-1/z) \to f(z)$ not so bad either

Slopes (power-law): $D(z)$, 1.48; $zf(2-\frac{1}{z})$, 1.52; $zf(z)$, 1.37

Conclusions

- Information entropy and KL divergence deployed for fragmentation functions
- S is one number to quantify progress on FF knowledge
- Relations $pdfs \leftrightarrow FFs$ not well satisfied with current sets
- Guess at what's happening? $f(x)$ measured on the proton; but $N^*, \Delta^* \to p \ldots$ contribute to $D(z)$, not only p
- Or Barone-Drago-Ma relation trivial and good only at $z = 1$ exactly?

Conclusions

- Information entropy and KL divergence deployed for fragmentation functions
- S is one number to quantify progress on FF knowledge
- Relations $pdfs \leftrightarrow FFs$ not well satisfied with current sets
- Guess at what's happening? $f(x)$ measured on the proton; but $N^*, \Delta^* \to p \ldots$ contribute to $D(z)$, not only p
- Or Barone-Drago-Ma relation trivial and good only at $z = 1$ exactly?

Conclusions

- Information entropy and KL divergence deployed for fragmentation functions
- S is one number to quantify progress on FF knowledge
- Relations *pdfs* $\leftrightarrow$ *FFs* not well satisfied with current sets
- Guess at what's happening? $f(x)$ measured on the proton; but $N^*, \Delta^* \to p \ldots$ contribute to $D(z)$, not only p
- Or Barone-Drago-Ma relation trivial and good only at $z = 1$ exactly?

Conclusions

- Information entropy and KL divergence deployed for fragmentation functions
- S is one number to quantify progress on FF knowledge
- Relations $pdfs \leftrightarrow FFs$ not well satisfied with current sets
- Guess at what's happening? $f(x)$ measured on the proton; but $N^*, \Delta^* \to p \ldots$ contribute to $D(z)$, not only p
- Or Barone-Drago-Ma relation trivial and good only at $z = 1$ exactly?

Conclusions

- Information entropy and KL divergence deployed for fragmentation functions
- S is one number to quantify progress on FF knowledge
- Relations *pdfs* $\leftrightarrow$ *FFs* not well satisfied with current sets
- Guess at what's happening? $f(x)$ measured on the proton; but $N^*, \Delta^* \to p \ldots$ contribute to $D(z)$, not only p
- Or Barone-Drago-Ma relation trivial and good only at $z = 1$ exactly?

Information entropy and fragmentation functions

Felipe J. Llanes-Estrada
Guillermo Benito Calviño & Javier García Olivares

Universidad Complutense de Madrid, Departamento de Física Teórica and IPARCOS

1/Aug/2022 XV Quark Confinement & Hadron Spectrum

Funding acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824093; grants MICINN: PID2019-108655GB-I00, PID2019-106080GB-C21 (Spain); UCM research group 910309 and the IPARCOS institute.