PDFs and machine learning

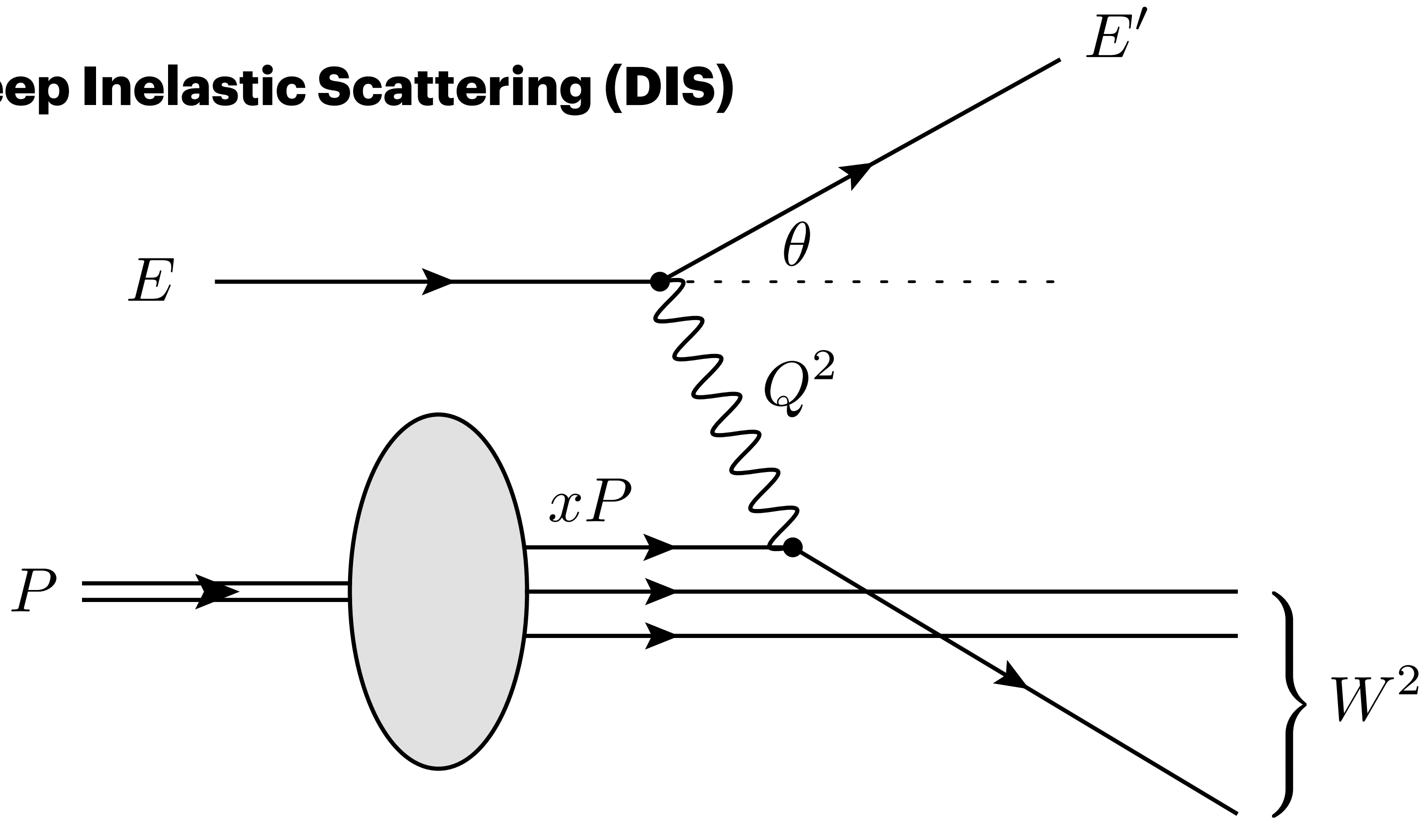
P. Zurita

in collaboration with **D. Rentería-Estrada**, **R. Hernández-Pinto** and **G. Sborlini**

Universität Regensburg

About Parton Distribution Functions (PDFs)

Deep Inelastic Scattering (DIS)



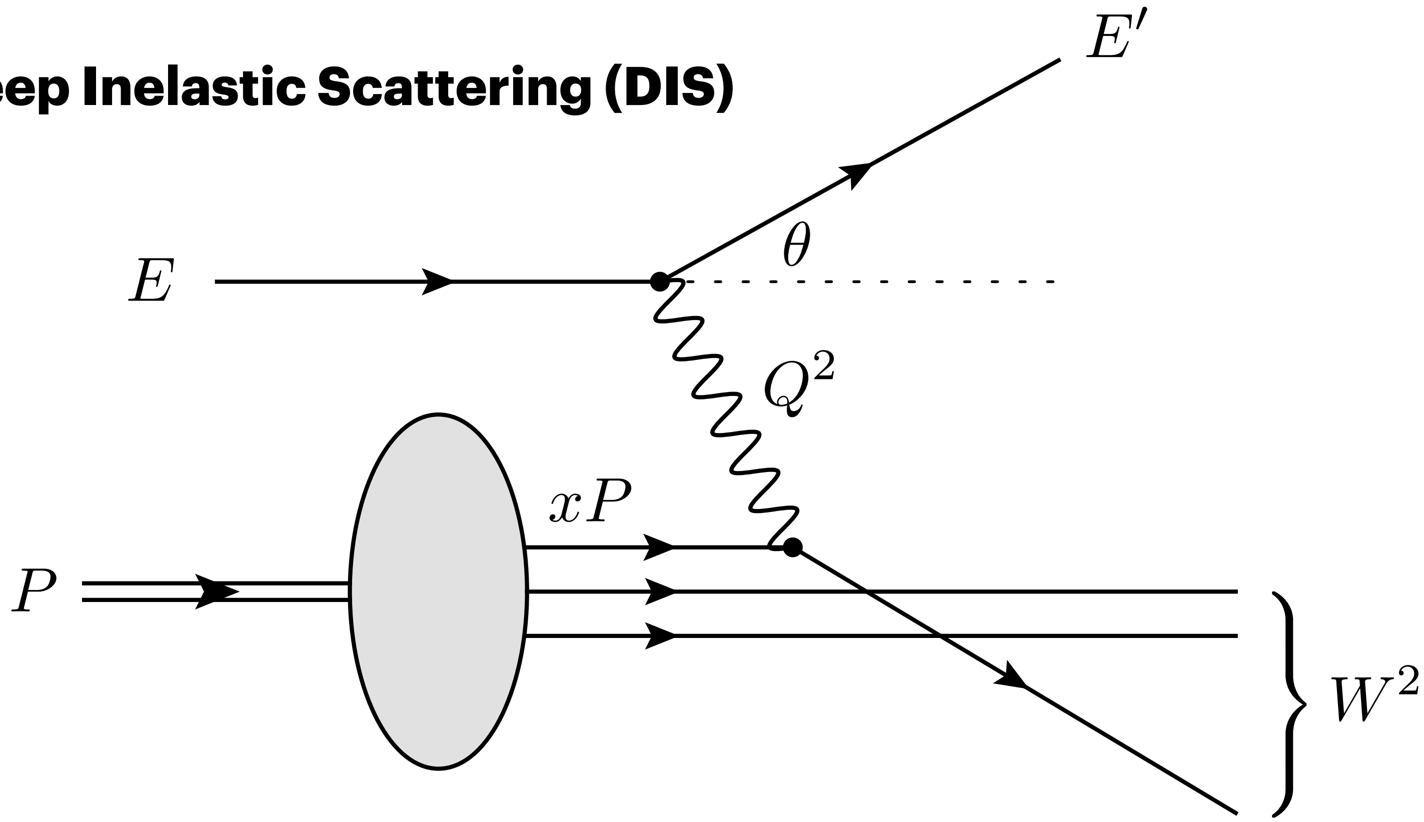
$$x = \frac{Q^2}{2P \cdot q}$$

$$Q^2 = -(k - k')^2$$

$$d\sigma^{DIS} = \sum_i d\sigma^{l+i \to l'} \otimes f_i$$

About Parton Distribution Functions (PDFs)

Deep Inelastic Scattering (DIS)



$$x = \frac{Q^2}{2P \cdot q} \qquad Q^2 = -(k-k')^2$$

$$d\sigma^{DIS} = \sum_i d\sigma^{l+i\to l'} \otimes f_i$$

- At LO $f_i(x, \mu^2)$ is the **probability** density of finding the parton ***i*** inside the proton, carrying a fraction ***x*** of the proton's longitudinal momentum (in the Breit frame) when we look at it with scale μ^2.

- Beyond LO the probabilistic interpretation is no longer clear.

- The PDFs can’t be computed from first principles in pQCD.

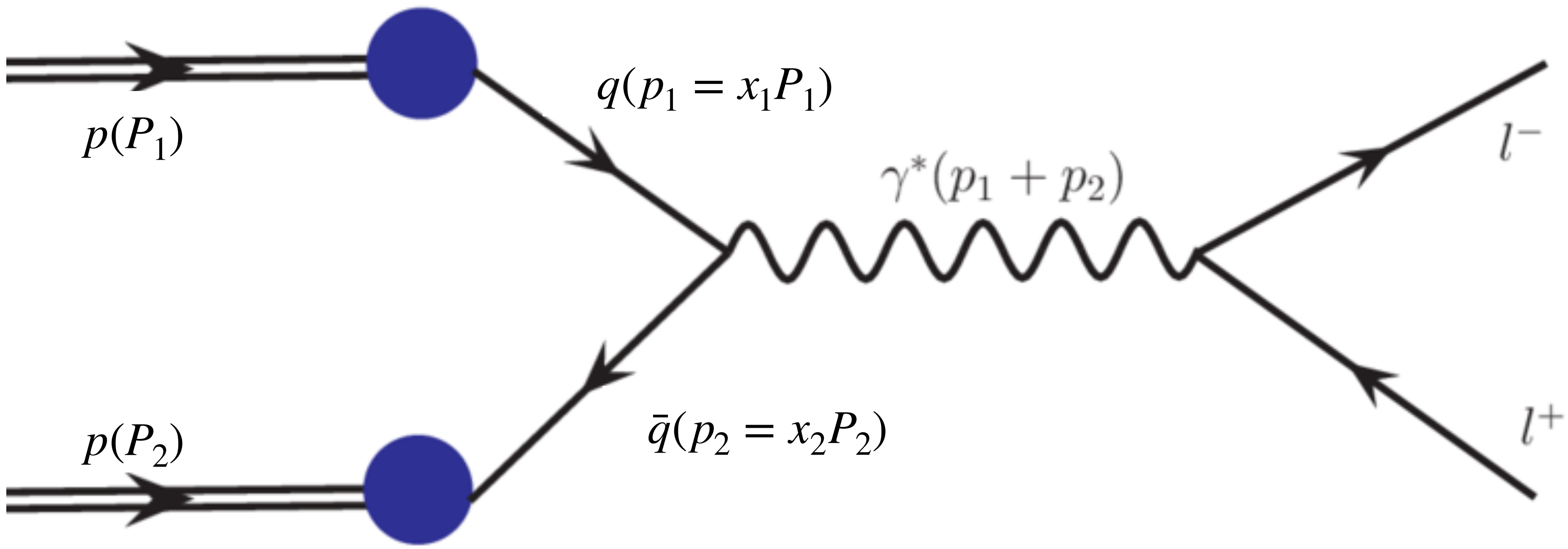
- We do know how they evolve with the scale, via the DGLAP evolution equations.

- The PDFs can't be computed from first principles in pQCD.
- We do know how they evolve with the scale, via the DGLAP evolution equations.
- To obtain the PDFs we do **global** fits to data.

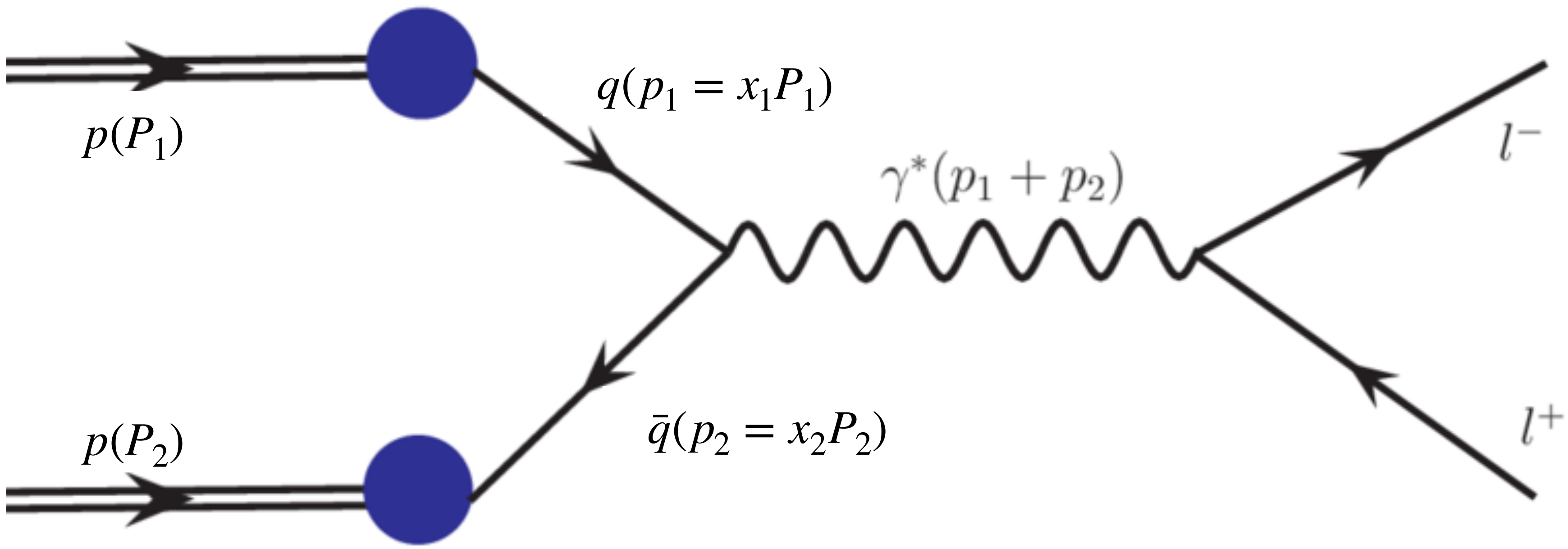
- The PDFs can't be computed from first principles in pQCD.
- We do know how they evolve with the scale, via the DGLAP evolution equations.
- To obtain the PDFs we do **global** fits to data.
- We have one PDF per light flavour (up, down, strange), their anti-particles and the gluon.

- The PDFs can't be computed from first principles in pQCD.
- We do know how they evolve with the scale, via the DGLAP evolution equations.
- To obtain the PDFs we do **global** fits to data.
- We have one PDF per light flavour (up, down, strange), their anti-particles and the gluon.
- In DIS, by measuring the outgoing electron we know everything about the kinematics.
- But we need more observables.

- Ex. Drell-Yan

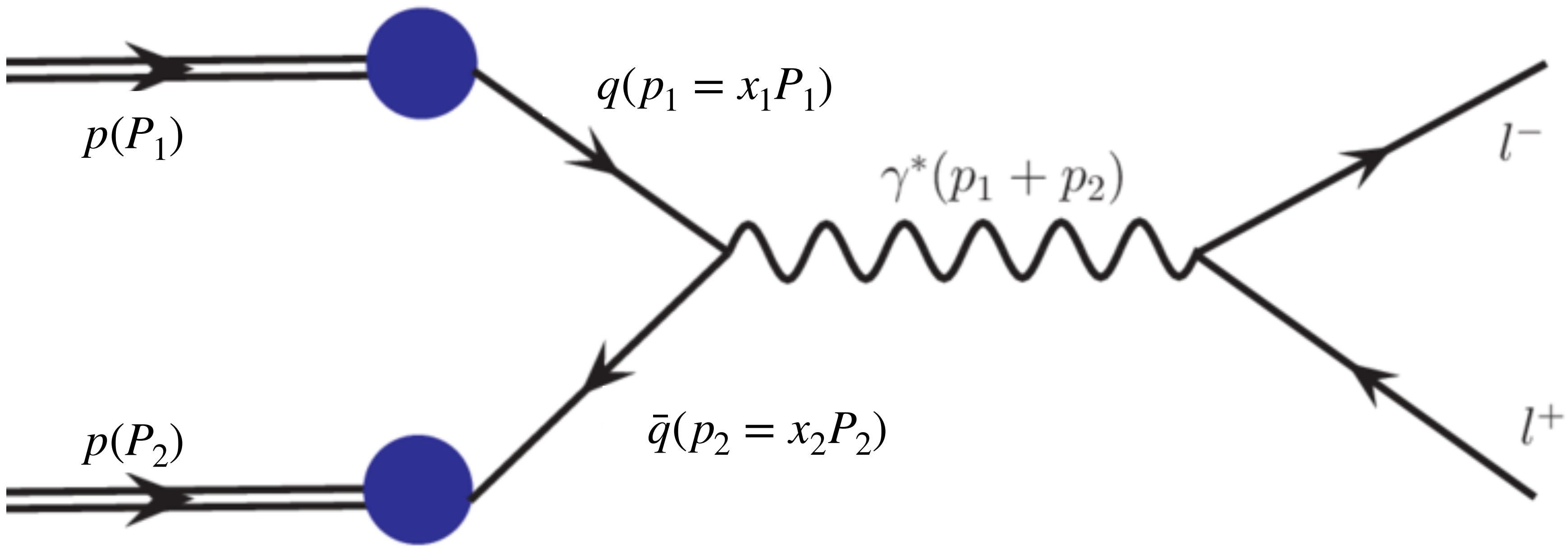


- Ex. Drell-Yan



- At LO $x_1, x_2 = \sqrt{M^2/s}\ e^{\pm y}$. Beyond LO this is no longer true.

- Ex. Drell-Yan



- At LO $x_1, x_2 = \sqrt{M^2/s}\, e^{\pm y}$. Beyond LO this is no longer true.

- Possible impact in :
    - p+A collisions (QGP benchmarking).
    - polarised collisions (spin).
    - BSM searches.

Accessing the kinematics using ML

Renteria-Estrada et *al*.,
arXiv:2112.05043 [hep-ph]

- We looked at one particular process: $p + p \to \pi^+ + \gamma$
- Reconstructed x_1, x_2 and z from the momenta of π^+, γ
- For RHIC kinematics, so we could compare with previous results.

D. de Florian and G. Sborlini,
Phys.Rev.D 83 (2011) 074022

Accessing the kinematics using ML

Renteria-Estrada et *al.*, arXiv:2112.05043 [hep-ph]

- We looked at one particular process: $p + p \to \pi^+ + \gamma$

- Reconstructed x_1, x_2 and z from the momenta of π^+, γ

- For RHIC kinematics, so we could compare with previous results.

  D. de Florian and G. Sborlini, Phys.Rev.D 83 (2011) 074022

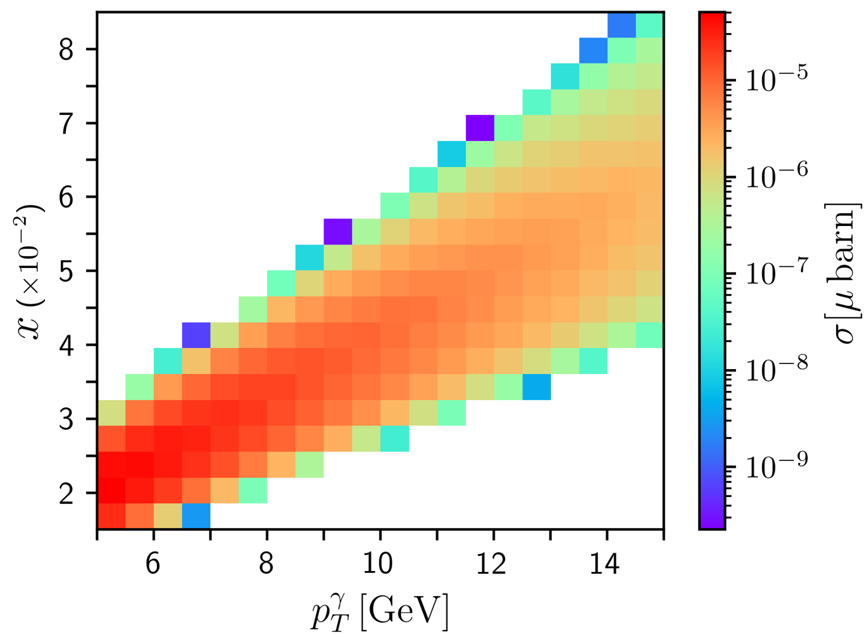
- This type of calculation is done with Vegas.

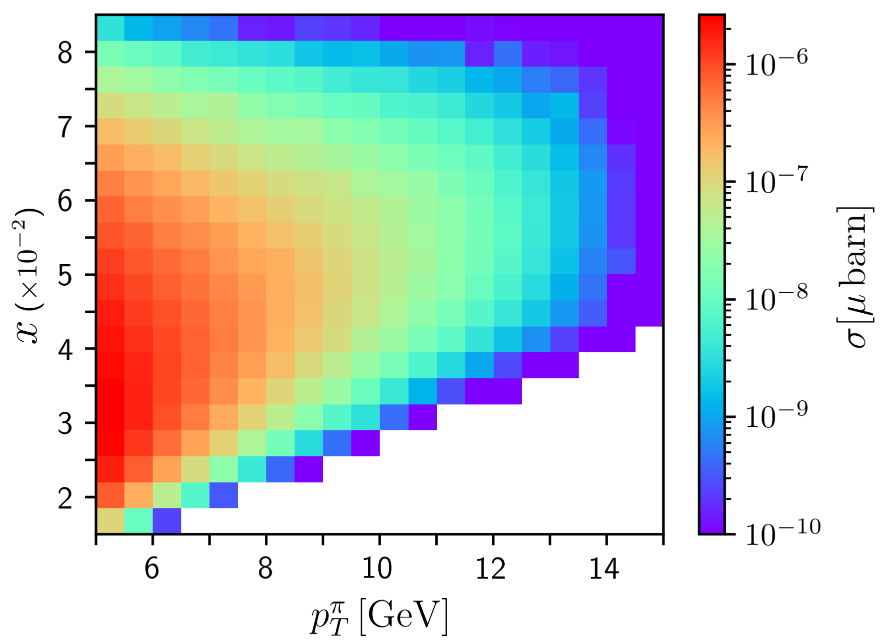
- Full check of kinematic dependences.

Second: check correlations

LO Kinematics

$$x_{1,2} = \frac{p_T^\gamma}{\sqrt{s}} \left(e^{\eta^{\pm\pi}} + e^{\eta^{\pm\gamma}} \right)$$





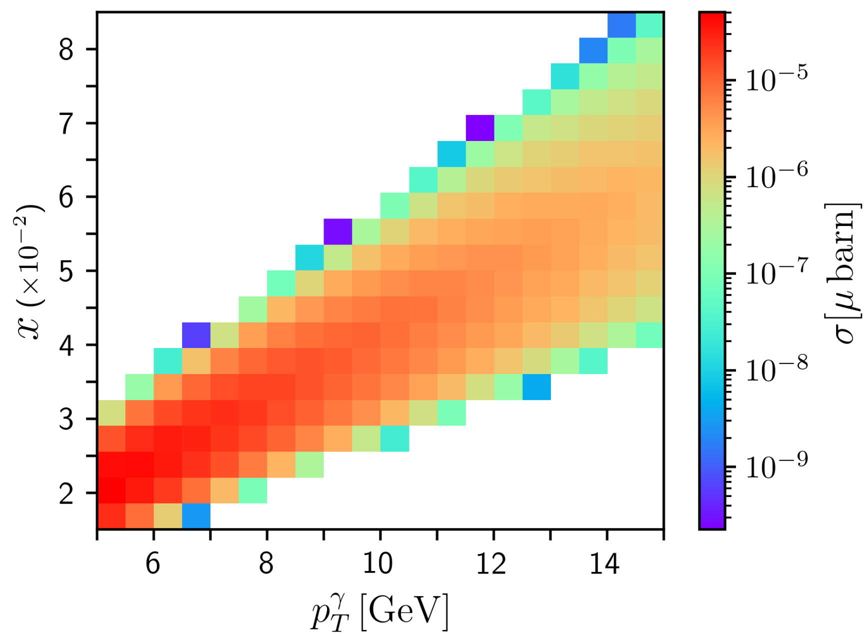
Second: check correlations

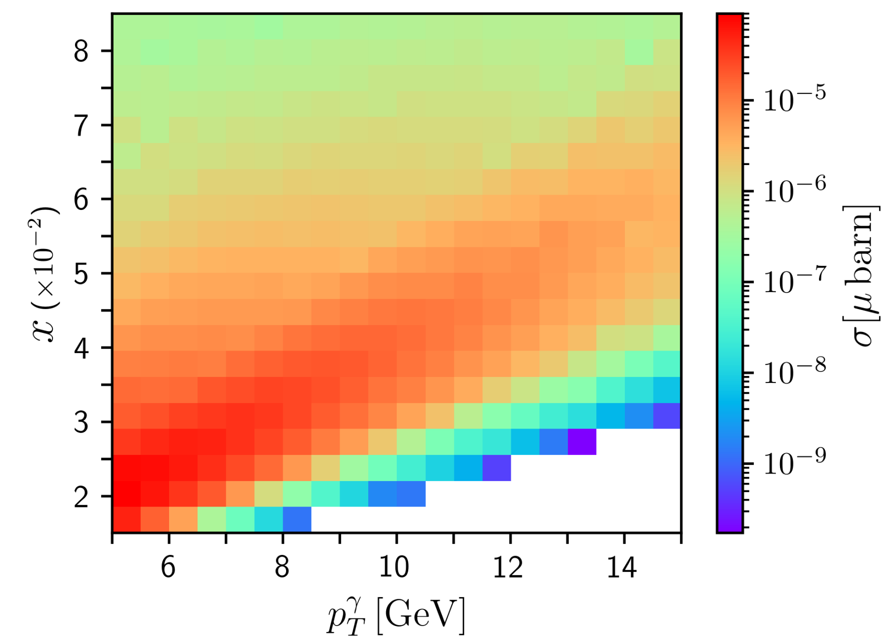
LO Kinematics

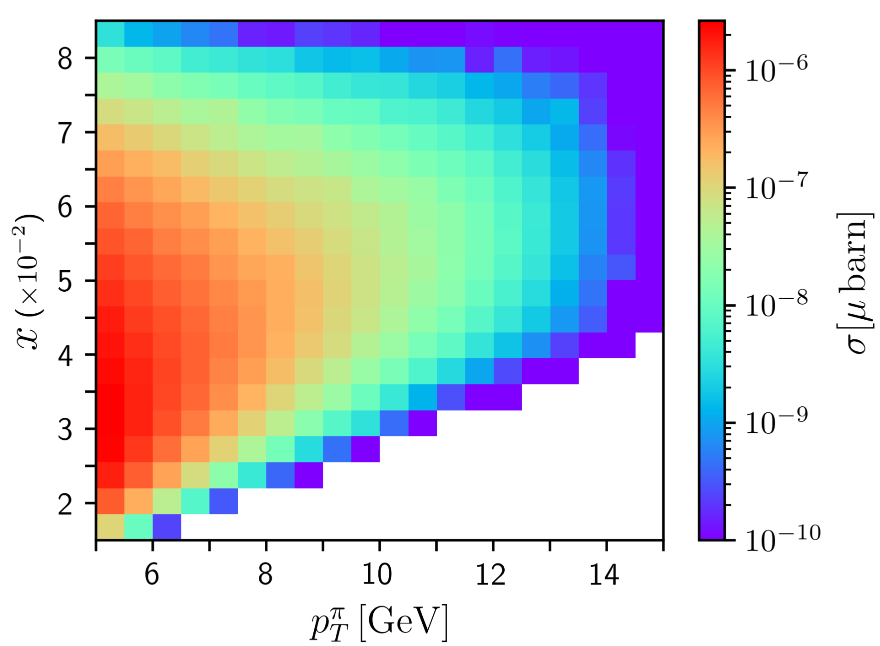
$$x_{1,2} = \frac{p_T^\gamma}{\sqrt{s}}\left(e^{\eta^{\pm\pi}} + e^{\eta^{\pm\gamma}}\right)$$

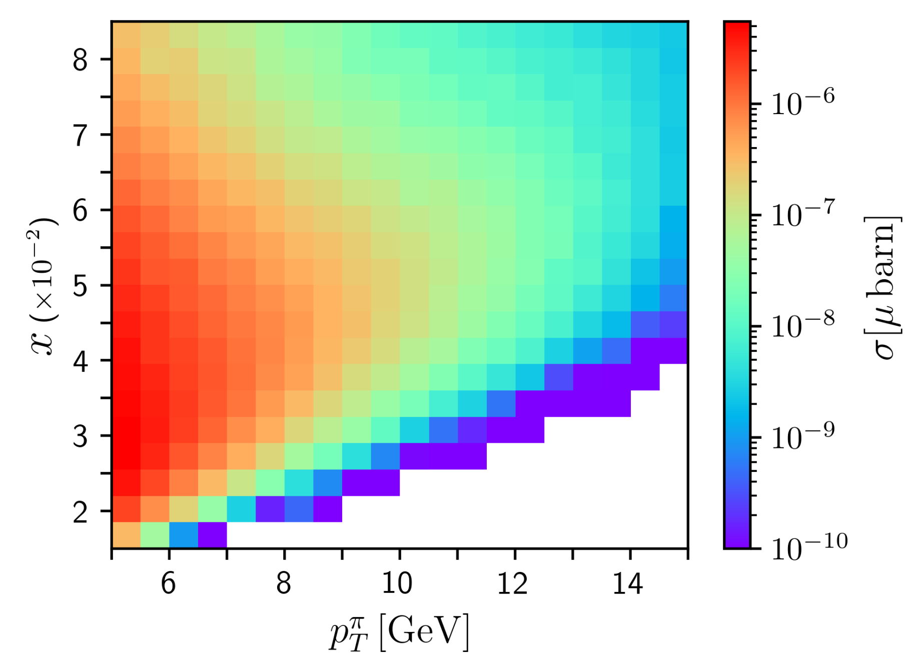
NLO Kinematics

$$x_{1,2} = ?$$









Kinematics: LO

$$x_{1,2}^{rec.} = \frac{p_T^{\gamma}}{\sqrt{s}} \left(e^{\pm\eta^{\pi}} + e^{\pm\eta^{\gamma}} \right)$$

Kinematics: NLO

$$x_{1,2}^{rec.} = ?$$

Kinematics: LO

$$x_{1,2}^{rec.} = \frac{p_T^\gamma}{\sqrt{s}} \left(e^{\pm\eta^\pi} + e^{\pm\eta^\gamma} \right)$$

Kinematics: NLO

$$x_{1,2}^{rec.} = ?$$

- Experimental collaborations used

$$x_{1,2}^{rec.} = \frac{p_T^\gamma e^{\pm\eta^\pi} - \cos(\phi^\pi - \phi^\gamma) p_T^\gamma e^{\pm\eta^\gamma}}{\sqrt{s}}$$

Kinematics: LO

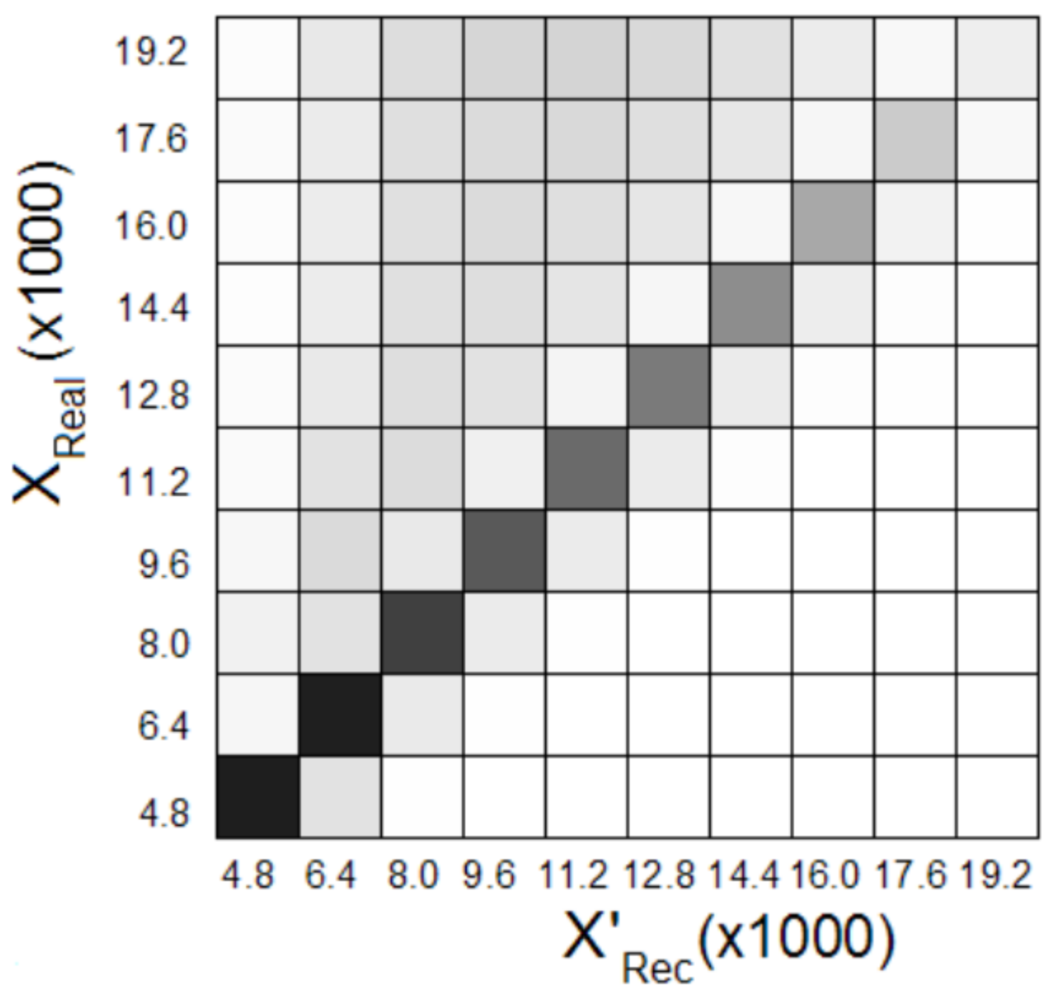
$$x_{1,2}^{rec.} = \frac{p_T^\gamma}{\sqrt{s}}\left(e^{\pm\eta^\pi} + e^{\pm\eta^\gamma}\right)$$

Kinematics: NLO

$$x_{1,2}^{rec.} = ?$$

- Experimental collaborations used

$$x_{1,2}^{rec.} = \frac{p_T^\gamma e^{\pm\eta^\pi} - \cos(\phi^\pi - \phi^\gamma) p_T^\gamma e^{\pm\eta^\gamma}}{\sqrt{s}}$$



D. de Florian, G. Sborlini, PRD 83, 074022

- At NLO we have real $(2 \to 3)$ and virtual $(2 \to 2)$ contributions and counterterms $(2 \to 2)$.

- Cancellations can only happen in the MC integration when histograming.

$$\left\{\bar{p}_T^{\gamma}, \bar{p}_T^{\pi}, \bar{\eta}^{\gamma}, \bar{\eta}^{\pi}, \overline{\cos}(\phi^{\pi} - \phi^{\gamma})\right\} \in \bar{\mathscr{V}}_{EXP}$$

$$\sigma_j(\bar{p}_T^{\gamma}, \bar{p}_T^{\pi}, \bar{\eta}^{\gamma}, \bar{\eta}^{\pi}, \overline{\cos}(\phi^{\pi} - \phi^{\gamma})) = \int_{(p_T^{\gamma})_{j,MIN}}^{(p_T^{\gamma})_{j,MAX}} dp_T^{\gamma} \int_{(p_T^{\pi})_{j,MIN}}^{(p_T^{\pi})_{j,MAX}} dp_T^{\pi} \int dx_1 dx_2 dz \; d\bar{\sigma}$$

- We weight the momentum fractions from the MC with the per-bin cross-section

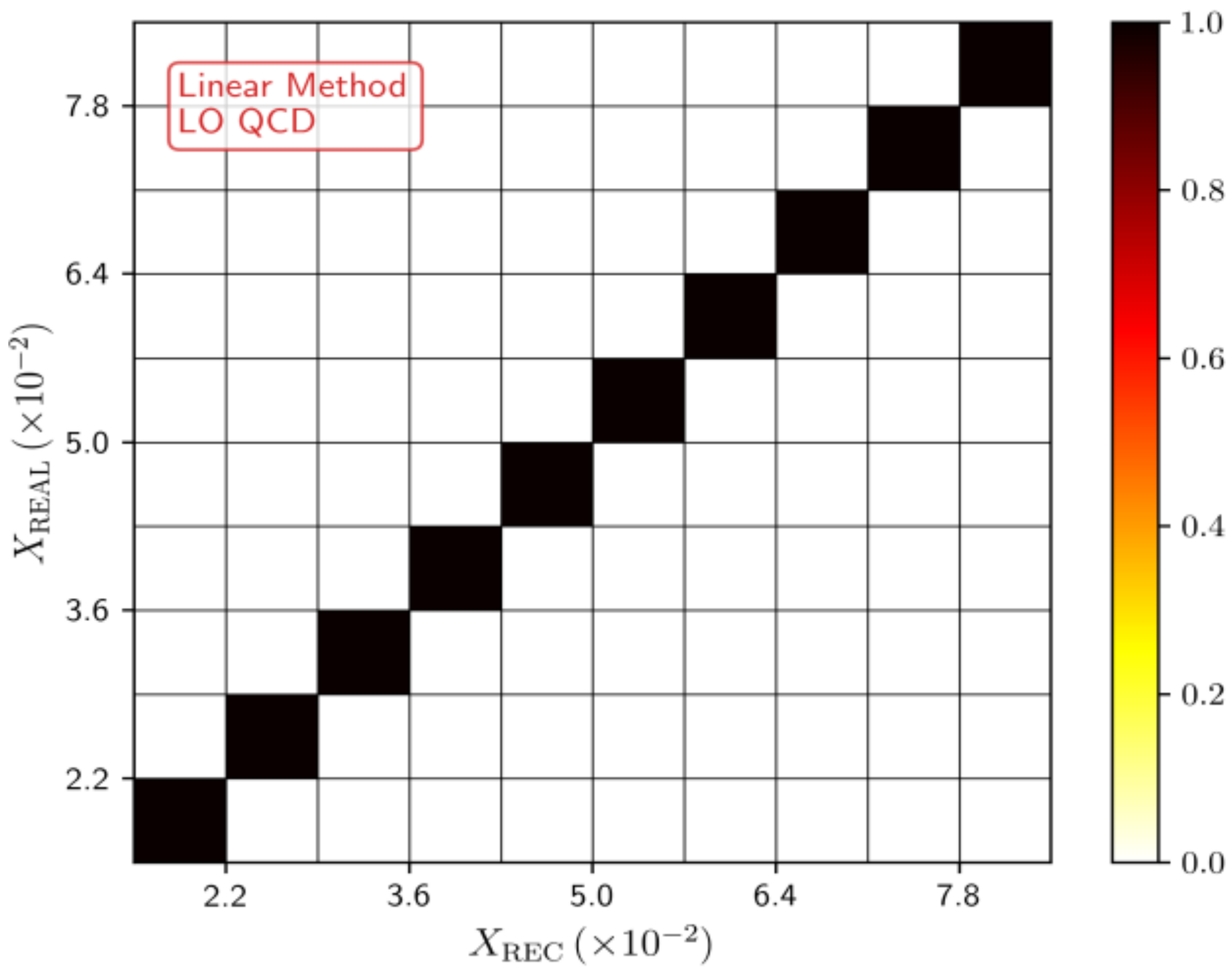
$$(x_1)_j = \sum_i (x_1)_i \frac{d\sigma_j}{dx_1}(p_j; (x_1)_i)$$

- With this we search for the mapping

$$X_{1,REC} : \; \bar{\mathscr{V}}_{EXP} \to \overline{X}_{1,REAL} = \{(x_1)_j\}$$

- Let us start with LO and use linear regression:
- Basis: $$\mathscr{B}_{LO} = \left\{ \frac{p_T^\gamma e^{\eta^\pi}}{\sqrt{s}}, \frac{p_T^\gamma e^{\eta^\gamma}}{\sqrt{s}}, \frac{p_T^\gamma e^{-\eta^\pi}}{\sqrt{s}}, \frac{p_T^\gamma e^{-\eta^\gamma}}{\sqrt{s}}, \frac{p_T^\pi}{p_T^\gamma} \right\}$$

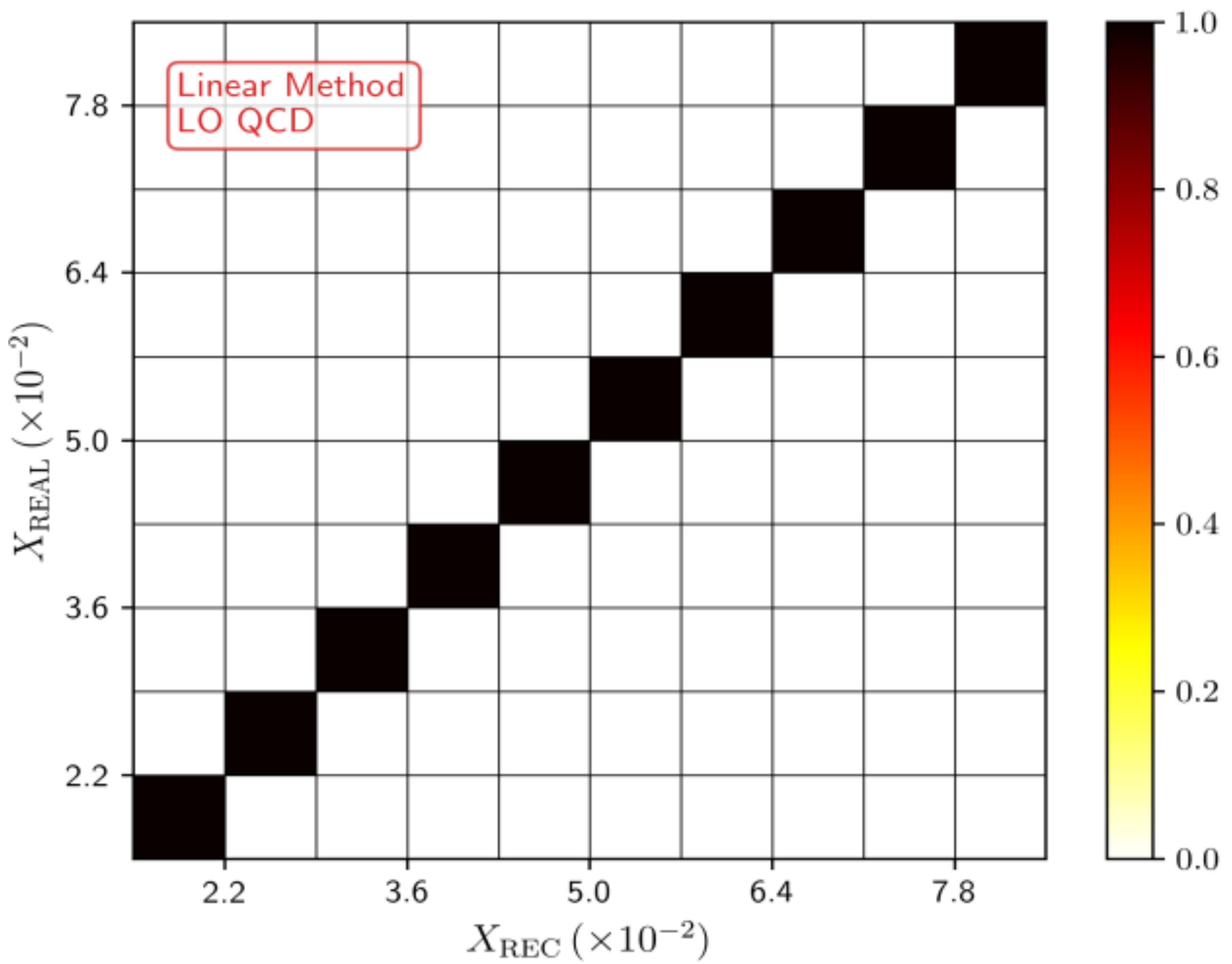
$$x_{1,2}^{rec.} = \frac{p_T^\gamma}{\sqrt{s}} \left(e^{\eta^{\pm\pi}} + e^{\eta^{\pm\gamma}} \right)$$



- Let us start with LO and use linear regression:

- Basis: $$\mathscr{B}_{LO} = \left\{ \frac{p_T^\gamma e^{\eta^\pi}}{\sqrt{s}}, \frac{p_T^\gamma e^{\eta^\gamma}}{\sqrt{s}}, \frac{p_T^\gamma e^{-\eta^\pi}}{\sqrt{s}}, \frac{p_T^\gamma e^{-\eta^\gamma}}{\sqrt{s}}, \frac{p_T^\pi}{p_T^\gamma} \right\}$$

$$x_{1,2}^{rec.} = \frac{p_T^\gamma}{\sqrt{s}} \left(e^{\eta^{\pm\pi}} + e^{\eta^{\pm\gamma}} \right)$$

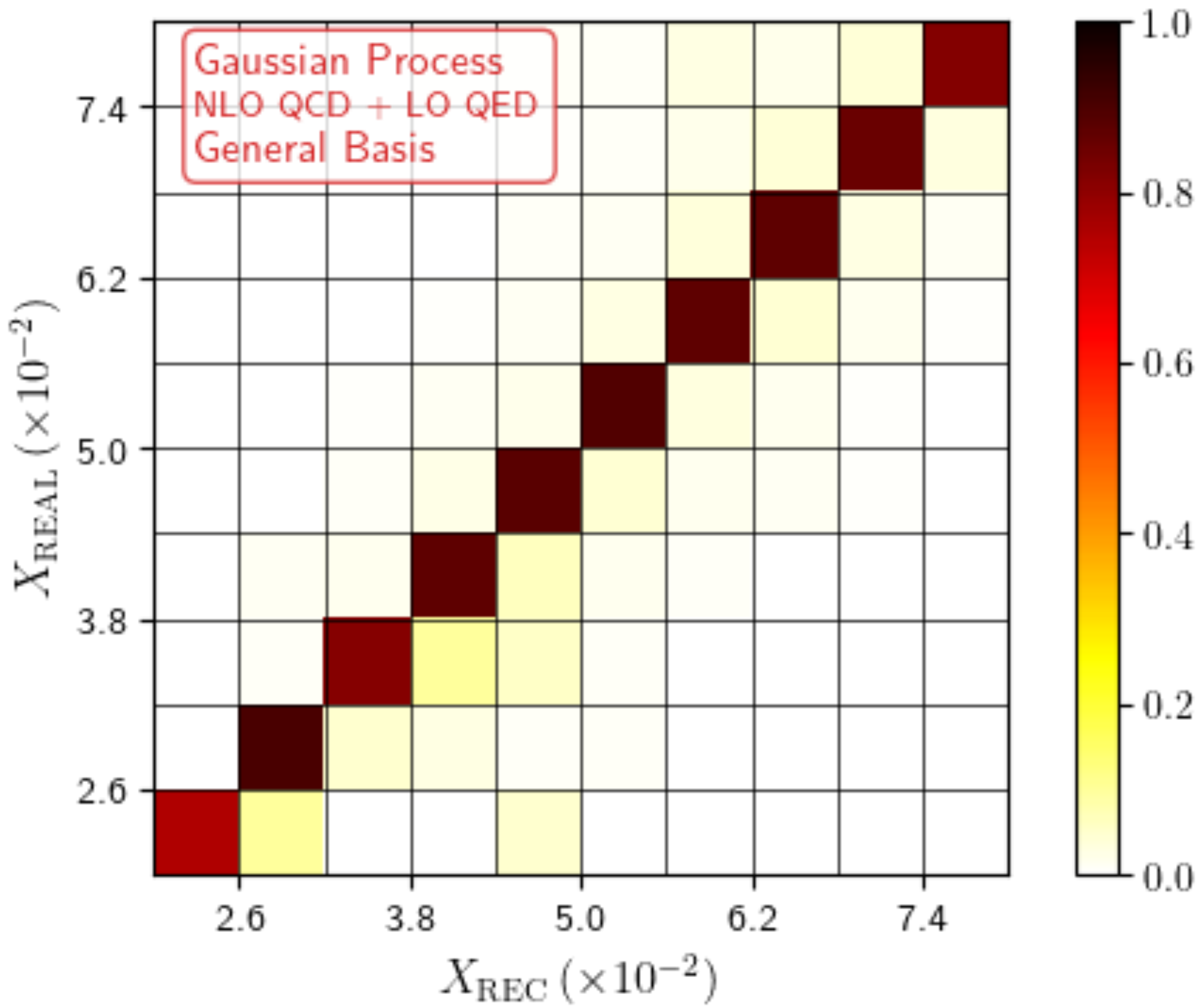


- At NLO we created three bases and tried to reconstruct x.

- Let us start with LO and use linear regression:

- Basis: $$\mathscr{B}_{LO} = \left\{ \frac{p_T^\gamma e^{\eta^\pi}}{\sqrt{s}}, \frac{p_T^\gamma e^{\eta^\gamma}}{\sqrt{s}}, \frac{p_T^\gamma e^{-\eta^\pi}}{\sqrt{s}}, \frac{p_T^\gamma e^{-\eta^\gamma}}{\sqrt{s}}, \frac{p_T^\pi}{p_T^\gamma} \right\}$$

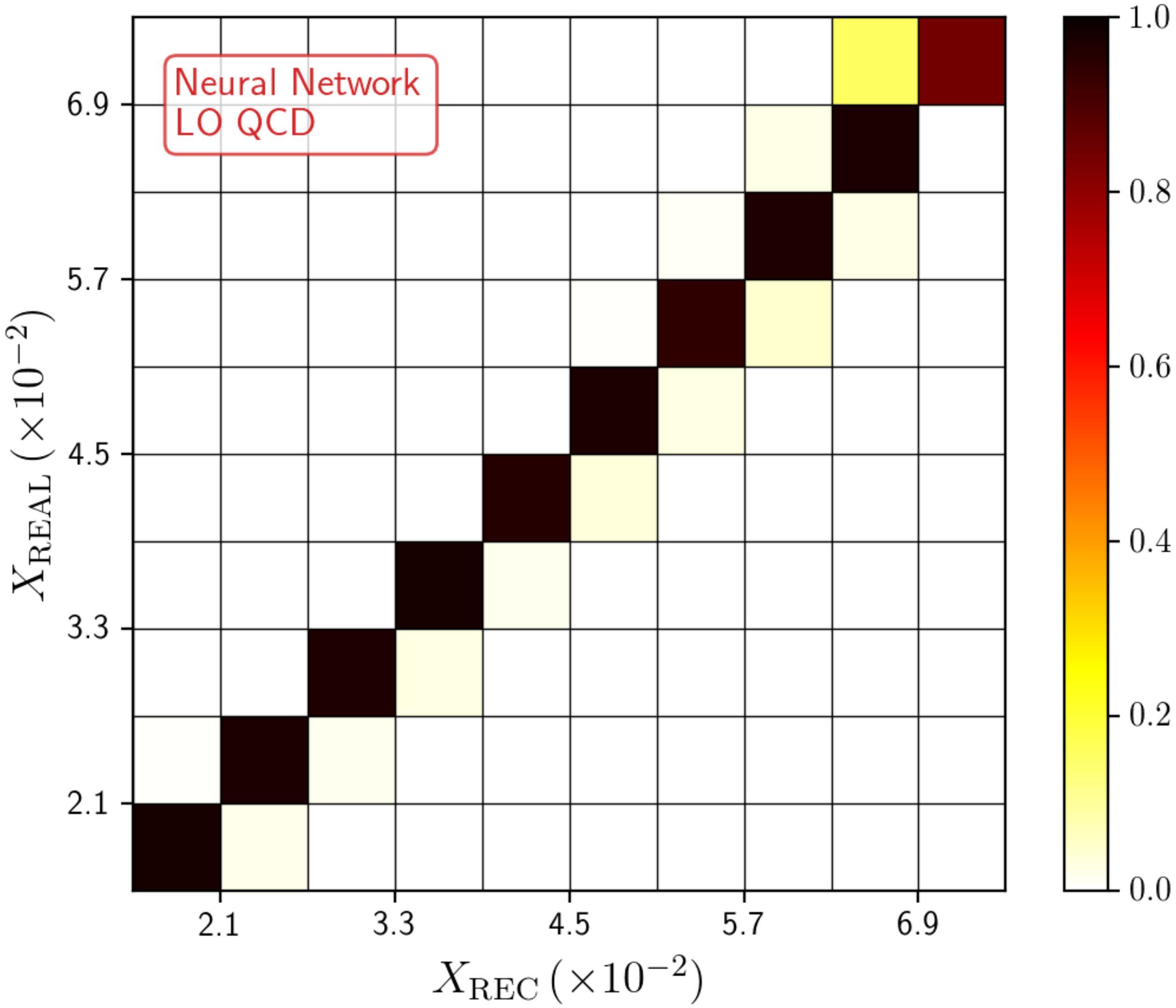
$$x_{1,2}^{rec.} = \frac{p_T^\gamma}{\sqrt{s}} \left(e^{\eta^{\pm\pi}} + e^{\eta^{\pm\gamma}} \right)$$



- At NLO we created three bases and tried to reconstruct x.

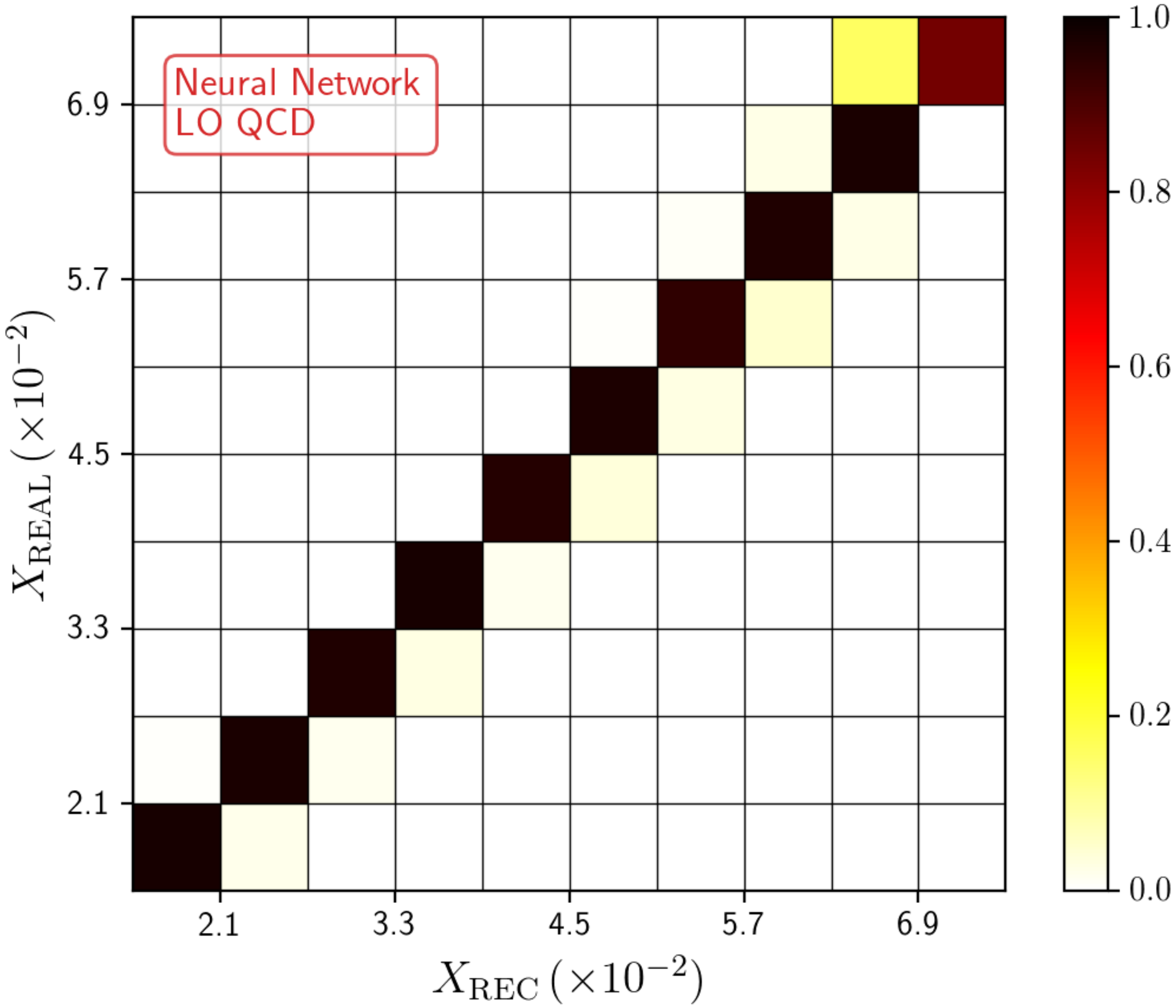
- We also tried RBF with the elements of the bases.

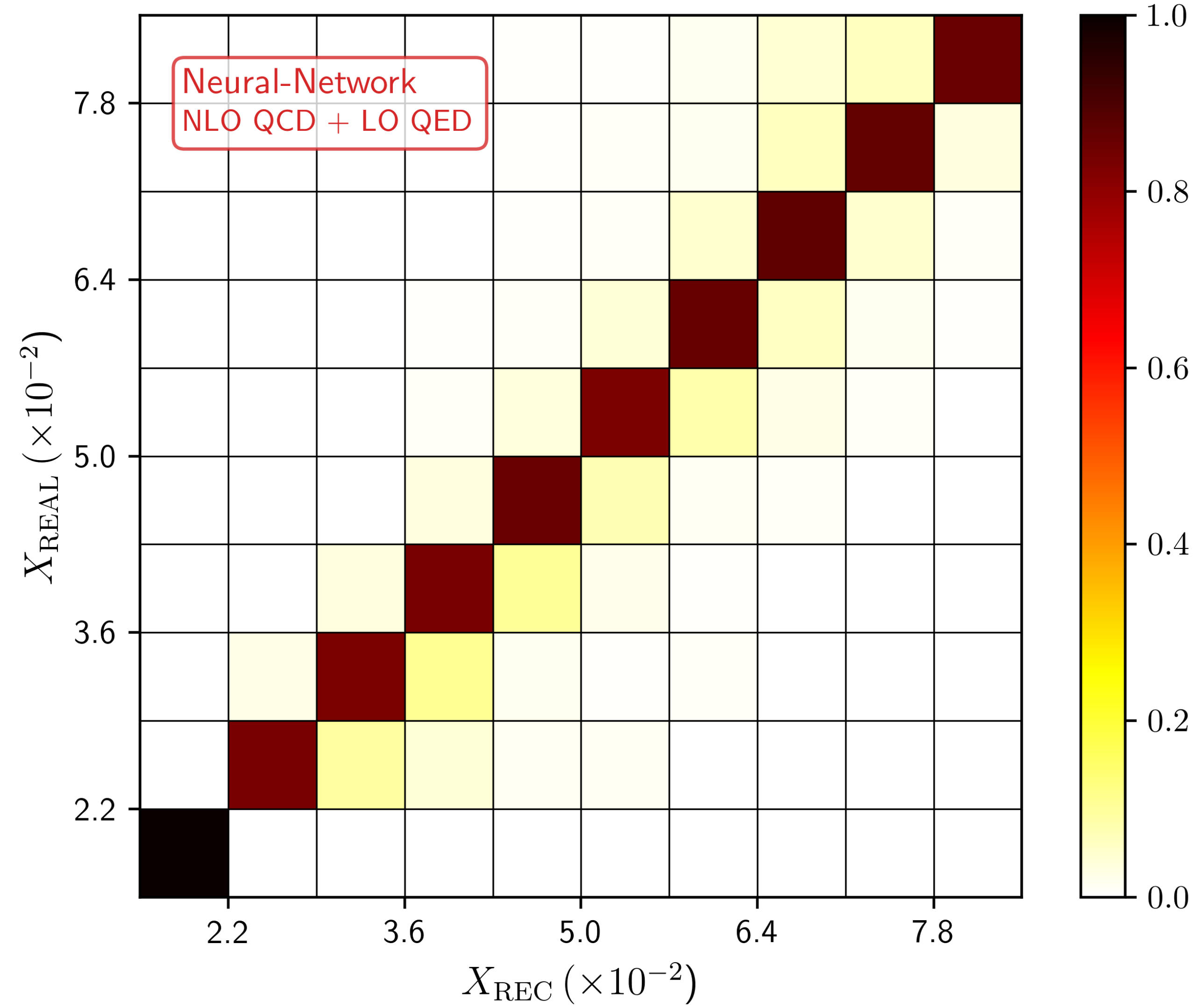
Neural networks



- For LO the complexity of the NN greatly surpasses the complexity of the problem.

Neural networks





- For NLO NN provide very good reconstruction.