Reconstructing parton collisions with machine learning techniques

G. F. R. Sborlini,[a,b,*] **D. F. Rentería-Estrada,**[c] **R. J. Hernández-Pinto**[c] **and P. Zurita**[d]

[a]*Departamento de Física Fundamental e IUFFyM, Universidad de Salamanca, E-37008 Salamanca, Spain*

[b]*Escuela de Ciencias, Ingeniería y Diseño, Universidad Europea de Valencia, Paseo de la Alameda 7, 46010 Valencia, Spain*

[c]*Facultad de Ciencias Físico-Matemáticas, Universidad Autónoma de Sinaloa, Ciudad Universitaria, CP 80000, Culiacán, Sinaloa, México*

[d]*Institut für Theoretische Physik, Universität Regensburg, 93040 Regensburg, Germany*

E-mail: german.sborlini@desy.de

Having access to the parton-level kinematics is important for understanding the internal dynamics of particle collisions. Here, we present new results aiming to an efficient reconstruction of parton collisions using machine-learning techniques. By simulating the collider events, we related experimentally-accessible quantities with the momentum fractions of the involved partons. We used photon-hadron production to exploit the cleanliness of the photon signal, including up to NLO QCD-QED corrections. Neural networks led to an outstanding reconstruction efficiency, suggesting a powerful strategy for unveiling the behaviour of the fundamental bricks of matter in high-energy collisions.

---

*Speaker

1. Introduction and motivation

In current high-energy experiments, bunches of particles collide, allowing their fundamental constituents to interact and produce new particles. The detectors identify the product of the collisions, but it is not straightforward to understand what is really going on during the process. In our recent work [1], we explore the application of Machine-Learning (ML) techniques to model the partonic momentum fractions in terms of experimentally-accessible variables. In first place, we calculated the differential hadronic cross-section for photon-hadron production including up to Next-to-Leading Order (NLO) QCD and Leading-Order (LO) QED corrections. Using a code based on Monte-Carlo (MC) integration [2, 3], we simulated the collisions and analyzed the events to determine the correlations among measurable and partonic quantities. Then, we applied ML algorithms that allow us to find the momentum fractions of the partons involved in the process in terms of suitable combinations of the final state momenta.

We tested our proof-of-concept (PoC) with photon-hadron production at colliders because the photon provides a clean probe to access the parton kinematics. The aim was to reconstruct the momentum fraction x and z of the partons coming from the protons and undergoing the fragmentation into a hadron, respectively. Due to the fact that parton kinematics are not physically-defined (i.e. it is a model), we provided a quantitative estimation of their most probable values. For this purpose, we trained Neural Networks (NN) to predict the MC partonic momentum fractions in terms of external momenta.

2. Computational setup

In order to carry out the calculations, we relied on the well-known factorization theorem, which implies that the cross-section is described by the convolution between PDFs, FFs and the partonic cross-section. On top of that, we took advantage of the smooth-cone isolation [4] with the purpose of avoiding the introduction of the photon fragmentation (whose accuracy is considerably lower than the other FFs). So, our starting point for the cross-section calculation was

$$d\sigma_{H_1 H_2 \to h\gamma} = \sum_{a_1 a_2 a_3} \int dx_1 \, dx_2 \, dz \, f_{H_1/a_1}(x_1, \mu_I) \, f_{H_2/a_2}(x_2, \mu_I) \, D_{a_3/h}(z, \mu_F) \, d\hat{\sigma}_{a_1 a_2 \to a_3 \gamma} \,, \quad (1)$$

where we included up to NLO QCD and LO QED corrections. For the NLO part, we applied the FKS method [5] to cancel the infrared singularities. Once the IR-finite NLO differential cross-section was defined, we established a *binning strategy* to discretize the phase-space. In this way, we managed to combine the real (2-to-3), virtual (2-to-2) and counter-terms (2-to-2) contributions by defining bins in the measurable variables, i.e. $\bar{\mathcal{V}}_{\text{Exp.}} = \{\bar{p}_T^\gamma, \bar{p}_T^\pi, \bar{\eta}^\gamma, \bar{\eta}^\pi, \overline{\cos}(\phi^\pi - \phi^\gamma)\}$. Then, we integrated the fully differential cross-section from Eq. (1) in each bin, obtaining σ_j for $p_j \in \bar{\mathcal{V}}_{\text{Exp.}}$. This step was crucial, since the higher-order corrections to the cross-sections involve contributions living in different phase-spaces.

3. Parton kinematics reconstruction

For the process that we considered, i.e. $p + p \to \pi + \gamma$, the partonic momentum fractions x_1, x_2 and z are unambiguously fixed by external particle's kinematics only at LO. When computing

higher-order corrections, real radiation including extra particles must be included and momentum conservation equations differs from LO kinematics. In order to obtain an approximation, we relied on the LO expressions and we found a decent agreement with the MC momentum fractions, as reported in Ref. [2]. However, in our recent work [1], we decided to tackle the problem by using more flexible ML techniques.
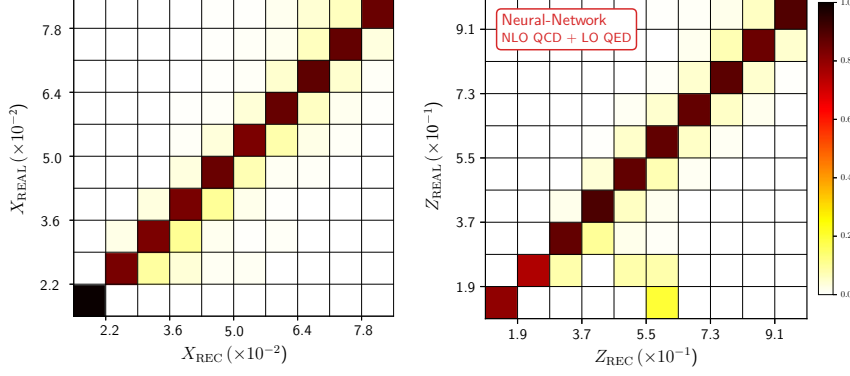
Figure 1: Comparison of the real momentum fractions (y-axis) and the reconstructed ones using Neural Networks (x-axis), for x (left) and z (right), respectively.

The first step was to introduce the concept of *effective momentum fractions* at higher-orders. Given a point in the grid, $p_j = \{\bar{p}_T^\gamma, \bar{p}_T^\pi, \bar{\eta}^\gamma, \bar{\eta}^\pi, \overline{\cos}(\phi^\pi - \phi^\gamma)\} \in \bar{\mathcal{V}}_{\text{Exp}}$, we defined

$$(x_{1,2})_j = \sum_i (x_{1,2})_i \frac{d\sigma_j}{dx_{1,2}}(p_j; (x_{1,2})_i), \quad (z)_j = \sum_i z_i \frac{d\sigma_j}{dz}(p_j; z_i), \tag{2}$$

with the purpose of identifying a mean value for the MC momentum fractions associated to the sum of all the real-emission topologies contributing to the same measurable final-state configuration (i.e. the same point p_j). Then, we used ML to find mappings connecting the points in the grid with the average value of the momentum fraction per bin, i.e.

$$X_{\text{REC}} := \bar{\mathcal{V}}_{\text{Exp}} \longrightarrow \bar{X}_{REAL} = \{(x)_j\}, \quad Z_{\text{REC}} := \bar{\mathcal{V}}_{\text{Exp}} \longrightarrow \bar{Z}_{REAL} = \{(z)_j\}. \tag{3}$$

We explored different approaches, being NN among them. The flexibility of the NN provided the best reconstruction with less constraints, as shown in the correlation plot in Fig. 1.

Finally, we studied the propagation of the scale uncertainties by defining different datasets, varying the renormalization and factorization scales by a factor 2 up and down. With these datasets, we trained estimators (i.e. $X_{\text{REC}}^{(\xi)}$, with ξ the energy scale) and evaluated them in each point of the grid. As reported in Ref.[1], the average reconstruction error turned out to be 7% for x and 5% for z. In Fig. 2, we show the error bands in the correlation plots, for different ML approaches. The bands are narrower for smaller values of x and z, since the cross-section is higher and the training is more accurate.

4. Conclusions

In this work, we presented a proof-of-concept (PoC) to reconstruct parton-level kinematics by using ML techniques. The results are in agreement with previous findings [2], but they required
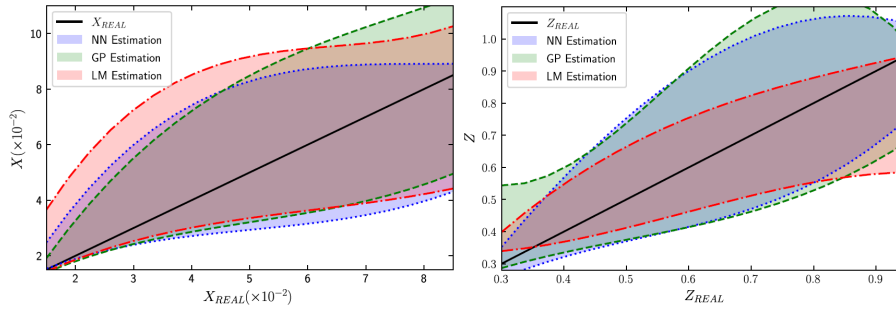
Figure 2: Correlation plots for x_1 (left) and z (right), with the associated error bands due to the propagation of the scale uncertainties in the reconstruction. The Linear Method (LM), Gaussian Regression (GP) and Neural Networks (NN) were considered.

much less human intervention. In particular, NN did not even need to select a function basis (which is an important choice in other approaches, such as Gaussian Regression or the Linear Method). Our results indicate that this PoC could be applied to other processes, and could be used to impose stricter constraints for PDF/FF determination.

Acknowledgments

References

[1] D.F. Rentería-Estrada, R.J. Hernández-Pinto, G.F.R. Sborlini and P. Zurita, *Reconstructing partonic kinematics at colliders with Machine Learning*, 2112.05043.

[2] D. de Florian and G.F.R. Sborlini, *Hadron plus photon production in polarized hadronic collisions at next-to-leading order accuracy*, *Phys. Rev. D* **83** (2011) 074022 [1011.0486].

[3] D.F. Rentería-Estrada, R.J. Hernández-Pinto and G.F.R. Sborlini, *Analysis of the Internal Structure of Hadrons Using Direct Photon Production*, *Symmetry* **13** (2021) 942 [2104.14663].

[4] S. Frixione, *Isolated photons in perturbative QCD*, *Phys. Lett. B* **429** (1998) 369 [hep-ph/9801442].

[5] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys. B* **467** (1996) 399 [hep-ph/9512328].